Reflection Across AI-based Music Composition

COREY FORD, Queen Mary University of London, UK
ASHLEY NOEL-HIRST, Queen Mary University of London, UK
SARA CARDINALE, Queen Mary University of London, UK
JACKSON LOTH, Queen Mary University of London, UK
PEDRO SARMENTO, Queen Mary University of London, UK
ELIZABETH WILSON, Queen Mary University of London, UK
LEWIS WOLSTANHOLME, Queen Mary University of London, UK
KYLE WORRALL, University of York, UK
NICK BRYAN-KINNS, University of the Arts London, UK

Reflection is fundamental to creative practice. However, the plurality of ways in which people reflect when using AI Generated Content (AIGC) is underexplored. This paper takes AI-based music composition as a case study to explore how artist-researcher composers reflected when integrating AIGC into their music composition process. The AI tools explored range from Markov Chains for music generation to Variational Auto-Encoders for modifying timbre. We used a novel method where our composers would pause and reflect back on screenshots of their composing after every hour, using this documentation to write first-person accounts showcasing their subjective viewpoints on their experience. We triangulate the first-person accounts with interviews and questionnaire measures to contribute descriptions on how the composers reflected. For example, we found that many composers reflect on future directions in which to take their music whilst curating AIGC. Our findings contribute to supporting future explorations on reflection in creative HCI contexts.

CCS Concepts: • **Applied computing → Sound and music computing**; • **Computing methodologies → Artificial intelligence**; • **Human-centered computing → User studies**.

Additional Key Words and Phrases: reflection, first-person, creativity, music, music composition, artificial intelligence, AI, generative AI, music generation

ACM Reference Format:
Corey Ford, Ashley Noel-Hirst, Sara Cardinale, Jackson Loth, Pedro Sarmento, Elizabeth Wilson, Lewis Wolstanholme, Kyle Worrall, and Nick Bryan-Kinns. 2024. Reflection Across AI-based Music Composition. In *Creativity and Cognition (C&C '24), June 23–26, 2024, Chicago, IL, USA*. ACM, New York, NY, USA, 24 pages. https://doi.org/10.1145/3635636.3656185

1 INTRODUCTION

When we create art, we reflect. Human-Computer Interaction (HCI) literature includes many reports on people's reflection in creative user experiences, from sketching [62] to music [92] to dance [39]. However, there are few explorations on the plurality of ways artists engage in reflection within a creative user experience. This may be because capturing reflections involves articulating thoughts which are tacit and hard to express [84]. We suggest that a deeper

Manuscript submitted to ACM

understanding of a variety of ways in which people reflect across a single creative practice could offer fresh insights for Creativity Support Tool (CST) research [45, 86].

In this paper, we explore reflection across different AI music tools and approaches to music composition. Like many creative domains, music composition presents an open-ended challenge where technology (such as AI) mediates the potential creative possibilities [67]. Interest in AI Generated Content (AIGC) has grown significantly in the creative industries [21] and, alongside recent calls for human-centred AI research [85] including in music [58] and the Arts [13], we see a timely opportunity to explore reflection in this domain.

This paper presents a collection of **six first-person [52, 66] reflective accounts** from composers on their experience composing a piece of music, each using a different AI tool. Our methodological novelty is to purposefully ask the composers to pause and reflect back on their music making, documenting their thoughts using a *reflection board* (see Section 3.4.2). The subjective accounts are triangulated with findings from a Thematic Analysis [9] of the first-person accounts and interviews, and questionnaire measures of reflection [44]. To be clear, our focus is not to compare whether AI-based music composition offers a more reflective experience than non-AI based music composition tools, rather to use AI music making as an area for *exploring* reflection across different practices. In summary, we offer the following contributions:

- **Contribution:** Our primary contribution is our first-person accounts of the AI-based composition processes of six artist-researchers, representing personal viewpoints on their practice.
- **Contribution:** Descriptions of how participants reflected when using AI in their composition processes triangulated from their first-person accounts, interviews and questionnaire measures. For example, we describe how composers reflected in-the-moment [17] using their instincts to curate AIGC synthesised in real-time.
- **Contribution:** A novel method where people pause to reflect back on their last hour of creative practice, supporting our primary contribution of the first-person accounts. Whilst we demonstrate the method for AI-based music composition, our method has potential for use across CST contexts.

The paper is structured as follows. First, we contextualise our research question by characterising reflection for creative user experiences and summarising literature on AI-based creativity support and AI music (Section 2). We then introduce our novel study method (Section 3), before presenting our findings (Section 4). We close by discussing our findings in the context of our literature, and suggest how our insights could inform future CST research (Section 5).

2 BACKGROUND

There is no consensus definition of reflection [3, 41, 72]. A common understanding of reflection is: moments where people sit back in quiet contemplation [72, 93]. In HCI, influential characterisations of reflection have guided research [3] such as Schön's [84] reflection-in-action (reflecting in-the-moment on actions) and reflection-on-action (reflecting back on past action). Some HCI research on designing technology for reflection focuses on the domain of personal informatics, designing technology for the functional goal of supporting self-improvement [6, 37]. This contrasts the goals of creative user experiences, which we define as human-computer interactions with tools that support open-ended tasks with no concrete metric of success [59].

2.1 Reflection in Creative Experience

Tools have been designed to support reflection for creative practice e.g. for documenting design projects [31, 89] or prompting reflection in children's storytelling [55]. However, CSTs for reflection to date have tended to leverage characterisations of reflection from HCI instead of characterisations specific to *creative* HCI. Table 1 shows characterisations of reflection from two key works for creativity-related HCI: Candy [17] suggested types of reflection drawn from interviews with creative practitioners; Ford and Bryan-Kinns [44] suggested types of reflection for their RiCE (Reflection in Creative Experience) questionnaire, consulting creativity experts and users of creative technology. We focus on these types of reflection in this paper and explore them through examples of AI-based creativity support, introduced below.

Table 1. Characterisations of Reflection specific to Creativity-Related HCI

Reflection	Definition
Reflection-for-action [17]	Reflecting on the possible actions to take in preparation for creating
Reflection-in-the-making-moment [17]	Reflecting on decisions during interaction with materials
Reflection-at-a-distance [17]	Taking an objective step back to evaluate one's art
Reflection-on-surprise [17]	Reflecting on unexpected occurrences
Reflection-on-current-process [44]	Reflecting on alternative directions to take an artwork
Reflection-on-self [44]	Reflecting on personal learning in the experience
Reflection-through-experimentation [44]	Reflecting on hypothesises through comparisons in a system

2.2 AI-based CSTs and Reflection

Early CST research [45] explored ways to enhance human creativity with technology, from automating menial tasks to developing fully collaborative digital partners [64]. Recent advances in AI techniques have led to AIGC being integrated into CSTs to act more like the latter, generating novel media indistinguishable from human creations. There are over 50 documented AI-based CSTs [88] where AIGC contributes to a shared product with the artist [80], often referred to as co-creative AI [33] or mixed-initiative systems [34]. When interacting with AI-based CSTs, people typically generate ideas with AI, curate these ideas, and then assemble the ideas into a cohesive whole [73]. However, the ways that reflection occurs when people use AI-based CSTs is underexplored, beyond a few speculative examples (e.g. [43, 93]).

To illustrate how reflection *might* occur when AIGC is used in creative practice, we briefly introduce examples of research on how artists have used AI-based CSTs. First, Caramiaux and Fdili Alaoui [20] found that pioneering creators of AI artworks leverage the ambiguity of AI-outputs by making glitches central to their process. We suggest that the ambiguity introduced by AI provides opportunities for reflection [43, 47, 93], such as to reflect on surprises [17] in the AI-output. Second, Yurman and Reddy [97] explored using Generative Adversarial Networks [48] in their watercolor practice, finding that they needed to reflect on their own perspectives to assign meanings to ambiguous AIGC. Third, Lewis [61] found that ChatGPT [78], when acting like an art teacher, would provide suggestions influencing their drawing style – sparking their reflection on the ownership of the data used by ChatGPT.

2.3 AI-based CSTs and Reflection for Music Composition

As with CSTs, there is a rich history of research on digital interfaces for music making e.g. at the New Instruments for Musical Expression conference [56]. In these music contexts, similar opportunities for reflection may arise as described in the preceding section. For example, glitches arise when people are improvising or organising musical material, reflecting on imperfections or ambiguity as an aesthetic choice [32, 50]. The composition process also varies across genres. For example, rock bands often construct and test ideas through jamming [7], live coders make edits on the fly and build patterns [70], whilst orchestral composers are noted to develop themes or motifs [68].

Despite its history, human-centred research on AI music is nascent [58]. AI music research has focused more on modelling musical aspects such as melody or timbre [51, 53], with most AI tools affording limited interactivity [12]. We illustrate a few suggestions as to how people *might* reflect with music AI tools, mirroring our suggestions for co-creative AI above. Huang et al. [53] found most musician/developer teams participating in the international AI songwriting contest[1] generated vast quantities of AIGC for later curation. We suggest that reflection likely occurs in the curation phase once material has been generated and varies across different stages of the composition process. For example, Sturm [92] reflected on AIGC from their FolkRNN AI system to identify and curate expressions for musical ideas which they could not personally formulate.

2.4 Summary & Research Question

Above we introduced characterisations of reflection for creative HCI research (Table 1) and suggested how reflection could occur during interaction with AI-based CSTs and AI music tools. With reflection largely underexplored in AI-based music composition [43, 93], we aim to further current understanding by asking the research question: **How do composers reflect when using an AI music tool in their music composition practice?** For our research question, we consider a range of AI-based music composition tools and approaches to music composition, as described below. Our question is purposefully open-ended to generate subjective insights for a breadth of AI music tools, as opposed to giving focus to a particular category of AI music tool.

3 METHOD

To showcase individual insights and identify commonalities in AI-based music making, we collected both qualitative and quantitative data. Our study was inspired by ethnographic approaches [4, 23, 71] to allow us explore a range of composition practices in their usual locations of happening e.g. at home [4, 93]. The study was approved by the Queen Mary University of London ethics committee. The participants in this study provided written consent and were reimbursed with a £100 (GBP) voucher following UK Musician's Union rates[2]. Each participant is acknowledged as co-authors of this paper for their first-person accounts and contributing compositions.

3.1 Participants

To recruit participants, we sent e-mails to research groups in the UK interested in music and AI. Our criteria for participation was to: i) be a PhD student; ii) have developed a way of integrating AI into your music making; and iii) be aged eighteen or older. Our participants are thus composers and artist-researchers [91] with a unique perspective in that they think about AI music in their everyday work life, have technical skills to use state-of-the-art AIGC within a music practice, and have academic writing skills for the first-person accounts. Seven composers were recruited, with two collaborating on a single composition as a band. Their characteristics are shown in Table 2, drawn from a questionnaire (see Appendix) which included the Goldsmith's Musical Sophistication Index [74] to quantify musical expertise – all score 75% and over indicating their strong musical skills; and the Self-Reflection and Insight Scale [49] to quantify natural capacity for reflection – all score over 70% indicating they are naturally reflective. Further details are introduced throughout Section 4.1.

[1] https://www.aisongcontest.com/
[2] https://musiciansunion.org.uk/

Table 2. Brief overview of the composers' characteristics.

ID	Age	Gender	Self-described Music Experience	Musical Sophistication Score [74]	Self-Reflection Score [49]
P1	26	Male	Performed in electronic and contemporary music ensembles for 10 years. Masters in Sonic Arts. Plays guitar and drums. Previous experience writing contemporary and minimalist music for chamber groups, jazz, indie and popular acts.	90%	71%
P2	25	Female	Undergraduate degree in Creative Music Technology. Media composer writing music for published video games, short movies and media companies. Also worked writing music for dance performances.	91%	88%
P3	26	Male	Guitarist for 15 years. 6-7 years music composition and production experience. Has released 5 original albums and produced/mixed music. Played in rock bands on guitar, bass, drums and vocals. **Note:** in a band with P4.	89%	83%
P4	32	Male	20+ years experience composing music, from classical guitar pieces to progressive metal. Experience as a solo classical guitarist and in 5 people ensembles (drums, two guitars, bass, keyboards). **Note:** in a band with P3.	87%	87%
P5	31	Female	Writing music for 15 years using conventional instruments e.g. guitar and piano. 5+ years experience as a live coder, making experimental electronic music, actively gigging.	84%	91%
P6	29	Male	Classically trained composer, writing both as a traditional composer and working with various small ensembles. Also a performer/improviser. Actively gigs. Writes experimental and computer music, and contemporary classical.	96%	88%
P7	29	Male	BA (Hons) in Creative Music Production; MSc in Sound and Music for Interactive Games. Specialised in composing for games. IMDb credit for a feature length horror film. 15+ years experience as a performer in death metal bands.	75%	88%

3.2 AI Tools

Each music composition in our study was made with a different AI tool which the participants self-selected and decided how to integrate into their workflow. We considered the AI tools sufficient for use if they had either been used in music making previously or published at an academic conference – all participants selected AI tools they had used at least once before to make music. Three key AI model architectures were present in the selected tools:

- **Markov Chains** model a probabilistic sequence of events, where events could be music data e.g. chords or melody notes [22]. Markov Chains are good at modelling small datasets compared to deep learning approaches but struggle to model long-term musical variations [81].
- **Transformers** are a deep-learning architecture able to generate musical output with an awareness of long-term structure [54]. They follow natural language approaches using music in a textual format as input [11]. Interest in transformers has arguably been influenced by the media emphasis on models used by ChatGPT [78] which can generate text [79], code [57], and music [2], and can be integrated into UI contexts e.g. Github Co-Pilot[3]. However, transformer models are difficult to control, other than through prompting [11, 24], as its inner processing is complex.
- **Variational Auto-Encoders** (VAEs) are an architecture where neural networks encode a dataset into a smaller compressed *latent* representation, in turn decoded by another set of neural networks [60]. This allows the internal model to be controlled more easily as users can tweak the values of the compressed latent representation and parse this to the decoder [81].

The AI tools used in the study are summarised in Table 3, with screenshots in Figure 1. We give further detail on each tool prior to each first-person account in Section 4.1 for context.

[3]https://github.com/features/copilot

Table 3. Summary of AI tools used by composers. Labelling (a) through (f) refers to the image labelling in Figure 1.

AI Tool	Composer	Architecture	Input	Output	Integration
(a) RAVE	P1	VAE	Audio	Audio with modified timbre	Plugin for Max
(b) Neural Resonator	P6	Neural Network	Audio excitation & UI	Audio of a synthesized drum	Plugin for music software
(c) CFEP	P7	Transformer	Text (MIDI)	Humanized MIDI as Text	Manual import MIDI
(d) Mark of Markov	P2	Markov Chain	Manual parameters in code	MIDI Notes & Chords	Records to music software
(e) ProgGP	P3 & P4	Transformer	Text (Guitar Tab)	Text (Guitar Tab)	Manual import MIDI
(f) Tidal-Fuzz	P5	Markov Chain	Text (Music Code)	Text (Music Code)	Manual import MIDI

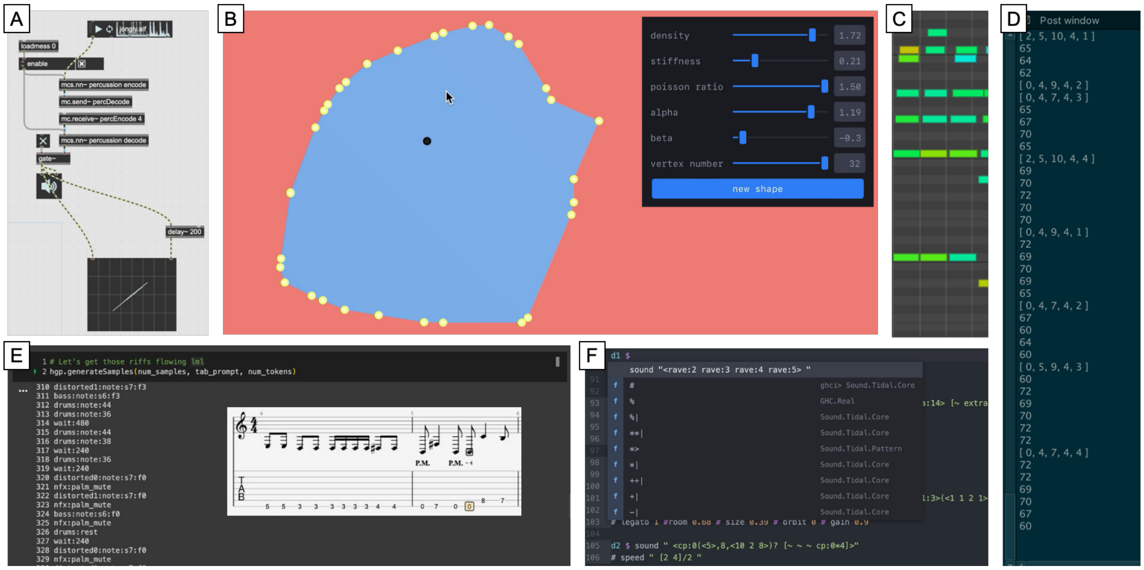


Fig. 1. AI tools integrated into the composers' practices. **(A)** Real-time Audio Variational Auto-Encoder (RAVE) [15]; **(B)** Neural Resonator [35]. **(C)** Cue-Free Express + Pedal (CFEP) [96]; **(D)** Mark of Markov (MoM); **(E)** ProgGp [63]; **(F)** Tidal-Fuzz [94].

3.3 Procedure

We asked participants to create a music composition with a minimum length of one minute in their chosen genre, using their chosen AI. They were asked to complete four sets of one hour long composition sessions, pausing to reflect back on their composing after every hour – in pilot tests we found four hours sufficient for a full composition cycle from ideation to completion for one minute of composed music. Coincidentally, all participants requested to complete the sessions within one day to balance with their other time commitments. The study was completed remotely to allow participants to be located in typical environments for their music making [4, 93]. We created moments for the composers to pause and reflect every hour, instead of the composers self-selecting moments to reflect, to ensure that we captured sufficient data on people's reflection whilst being mindful of time-constraints. This contrasts methodologies for CST studies on qualities such as feelings of flow [29] which is not the focus of this paper. The procedure is outlined in Table 4.

Table 4. The steps of the procedure. P = participant; R = researcher. For more detail, see the corresponding sections in this paper.

Step	Who	Task	Section	Length (Mins)
(1)	P	Complete the pre-test questionnaire	§3.1	10
(2)	P	Compose with AI whilst recording the computer screen.	-	60
(3)	R	Notify participants that the 1 hour session is over	-	1
(4)	P	Complete the RiCE [44] questionnaire for the recent session	§3.4.1	5
(5)	P	Complete a reflection board	§3.4.2	30
(6)	P & R	Complete an interview on Microsoft Teams	§3.4.3	10
(7)	P & R	Repeat steps 2 through 6 until 4 sessions are completed	-	-

3.4 Data Collection

We collected reflection questionnaires, reflection boards and interview data after each one hour music making session. These were followed by a first-person account after all sessions were completed.

3.4.1 Reflection Questionnaire. To explore possible trends in reflection throughout the composition processes, we collected metrics from the most recent version of the RiCE questionnaire (v2) [44][4]. RiCE is designed to measure post-hoc how much of a certain type of reflection a person self-reports to have experienced, based on a set of statements scored on ordinal scales. Averages from these statements are calculated to obtain scores for the following types of reflection (described in Table 1): reflection-on-current-process, reflection-on-self, reflection-through-experimentation and a total RiCE score.

3.4.2 Reflection Boards. Participants were given a template for the online collaborative whiteboard Miro[5] (see Figure 2). The template posed questions at the top of a set of columns based on the three factors of RiCE [44] to prompt and organise the participants' thinking. Participants were instructed to add 6-10 screenshots from their composition session that best represented their creative process, organising them in chronological order (from top to bottom) in the leftmost column – this shows how the composition unfolds as inspired by studies on the composition process [27, 42]. We used screenshots as they offer insights into the composers' personal decisions at specific points in time [46]. Then, using the post-it note feature, they were asked to document their reflections and thoughts on their composition process, using the guiding questions at the top of each column. The reflection boards were used instead of other retrospective protocols [18] as we wanted the composers to be self-sufficient in their documentation and able to quickly refer to the data later for their first-person accounts.

3.4.3 Interview. A short interview was undertaken in which participants were asked to i) talk through what they did in the preceding hour, and ii) talk through the reflections in Miro. Our approach was semi-structured to give the researcher opportunities to probe unexpected lines of discussion. We aimed to elicit descriptive accounts of the participant's experience, avoiding leading questions so that the researcher had minimal impact on the discussion [10]. The interviews also served as a contingency in case the composers could not complete their first-person accounts.

3.4.4 First-person Account. After all sessions were completed, participants were asked to write an 800-1000 word **first-person account** with the following instructions:

[4]Full details on RiCE can be found at https://ricequestionnaire.github.io/.
[5]https://miro.com/

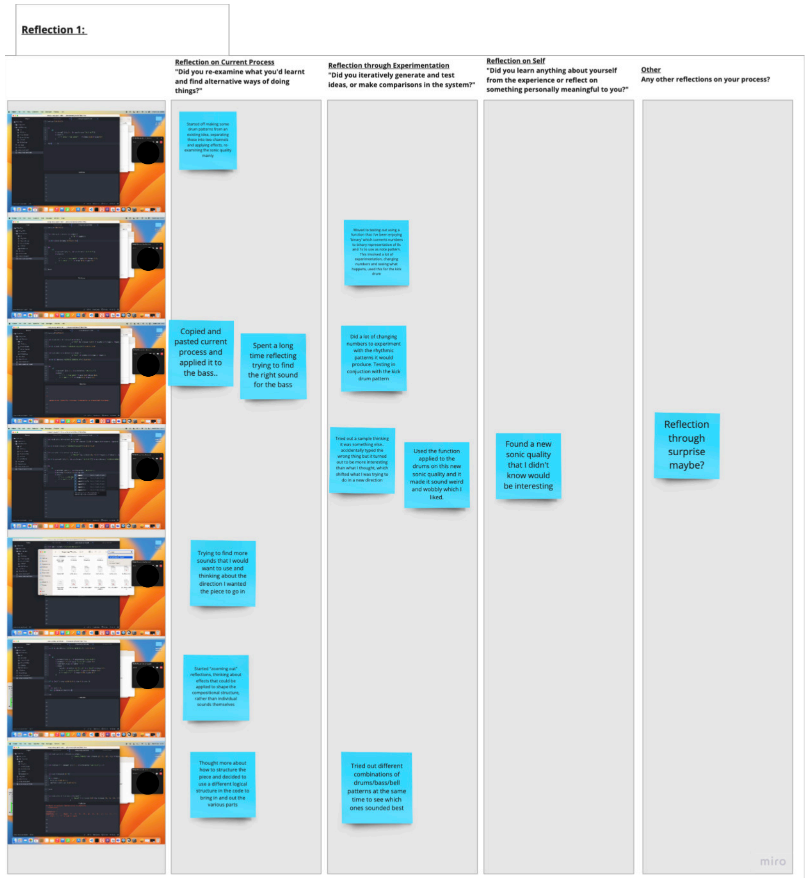


Fig. 2. An example reflection board from the first session completed by P5.

> "Write an account of how you composed with your chosen AI and what you reflected upon, looking over your Miro boards. We only expect first-draft quality. Please include all the key points you would like to talk about that you think are important, using your own voice."

We suggest these first-person accounts provided opportunities for the composers to articulate and clarify the reflections they captured in their composition sessions, and bring insights into their subjective viewpoints on their practice [14, 16, 19]. We also wish to give voice to the personal ways that our composers reflect to generate knowledge which might resonate with readers [38, 40, 90]. We were inspired by first-person approaches [52, 66] across HCI such as collaborative pieces [62, 91], autoethnographies [61, 65, 76, 87] and vignettes [5, 24], demonstrating the value of first-person approaches in giving insights into personal experience.

3.5 Data Analysis

We collated our reflection boards, interview transcripts, and the first-draft of the first-person accounts into one document. We then performed an inductive *Thematic Analysis* [9] on the document which involved iteratively: i) generating short descriptive codes for passages in the data, ii) refining codes, iii) organising the codes into themes, and iv) re-applying themes to test their fit. The first author consulted with their supervisor in regular meetings to ensure the rigor and consistency of their Thematic Analysis. Due to the small sample size, we do not conduct statistical analysis on our quantitative measures.

4 FINDINGS

We report in the following subsections our findings from: the first-person accounts, the RiCE questionnaire measures, and our Thematic Analysis.

4.1 First-person Accounts

For each first-person account below, we recap each composer's expertise and give details on their chosen AI tool. We present the first-person accounts from composers as edited extracts which we believe best relates to our research question whilst retaining the artists' voices (see Appendix for full accounts and the procedure used for editing). We invite readers to listen to each music composition at https://codetta.codes/reflection-across-AI-music/ whilst reading the accounts below attentively to immerse themselves into the composers' worlds.

4.1.1 P1: Ash. Ash composes music with a Glitch aesthetic, recording improvisations with interfaces they create using the visual programming language Max[6]. They chose the VAE model named RAVE [15] (see Figure 1a). RAVE can take an audio clip as input and change its timbre. For example, a recording of a person singing can be transformed to sound like a trumpet following the same melody. RAVE can generate high quality 48kHz audio signals and be used with a standard laptop CPU [15]. RAVE can also be controlled by varying values of the latent space in its VAE architecture and feeding this into its decoder.

> Typically, I like to get output as soon as possible. But I was surprised by how little I initially got from RAVE. I started becoming aware of the artifice of the technology, becoming increasingly aware that there is not even an idea in the way yet – just the technology. Do I even need to have this problem? What is the aim? What am I trying to do? The 8-dimensional input of RAVE and its non-deterministic output made me re-evaluate the structure of my typical process. I considered ideas from John Croft [28] such as what layer of abstraction (or the level of complexity) I wanted?
>
> Later on, when looking at Max documentation, I accidentally sidetracked into intermodulation. I started thinking about using FM synthesis (controlling the intermodulation of sine waves) and using this with RAVE for more nuanced control over its latent space. Through various signal processing techniques, I ended up with a way to control both RAVE and a non-AI FM synth. This allowed me to negotiate between the AI and non-AI sounds, where you can decide which to dominate whilst improvising. The combination of predictable and unpredictable, semantic and black box, brings a similar level of expectation with pleasant surprise as I had experienced being in Jazz ensembles. I regained control over some aspects whilst accepting I have little control over others. However, I still couldn't think of my composition in a deterministic way, like in FM synthesis where you have a good idea of what will happen when a parameter is changed (see John Chowning's Stria [26]). I can't control the model and know what it's doing, so I handed off control to the AI.

[6]https://cycling74.com/products/max

> I was surprised at sounds feeding back into RAVE. Some surprises came from the generation of sound outside the audible range of RAVE, as it interfered both with audible sounds and their latent representation. Welcoming the surprise of the AI again was productive and even fun at times.

4.1.2 P2: Sara. Sara is a media composer with experience working for video game companies, mainly writing orchestral music. Their chosen AI *Mark of Markov* (MoM; see Figure 1d) uses Markov Chains to output notes and chords that switch between modes (scales offering different musical moods) based on various probabilities[7]. Each chord output is a bar in length. On compiling MoM, its output is synthesised in real-time and can be recorded as MIDI.

> The initial material generated by MoM was boring – too quantized and not human at all. This was kind of restraining. To be less boring, I tried using an arpeggiator, but eventually re-recorded the chords myself. Furthermore, the generated melody was a bit hard to work with because it had weird rhythms and wasn't consistent. Because MoM spits out MIDI based on its previous music, I couldn't copy and paste parts from the melody and stitch them together, because there is a chance the chords could be in a different key. Instead, I changed the rhythm in the melody to make it less weird.
>
> I felt really bad changing the stuff MoM created – I wanted to use all of it so it did not go to waste. I thought that if I kept changing the system output, was I really using it to its full potential? Was I just taking over?
>
> Whilst composing, it was interesting that I kept making comparisons to a composition I previously wrote using MoM, which I was really proud of. I also would compare myself to people such as John Williams[8], and think, "well if I am going for a similar style to his, I cannot even get close to the quality of his compositions". This can get very demoralising and add a lot of pressure. I found taking small chunks of the output and trying to make them work together helped to take off the pressure. It also naturally allowed the composition to go in a different direction to the composition I previously made with MoM. Using the generated material in this way felt more like a game and gave me more freedom. If I didn't like the result, I could just try make something else by re-arranging parts or maybe changing them.
>
> Overall, MoM felt like a composition buddy! It felt like I was collaborating with another composer. Although I didn't write the material, I can say "hey, I took this generative melody and made it work within the composition", which I feel is a skill in itself.

4.1.3 P3 & P4: HEL900. Jack (P3) and Pedro (P4) create progressive metal music using AI as the band HEL9000[9]. They chose ProgGP [63], a transformer model [30] trained on the DadaGP dataset [82] – a dataset of 26k rock and metal guitar tablatures[10] – and fine-tuned further on a set of progressive metal guitar tablatures. The notation software Guitar Pro[11] is used by HEL9000 to write guitar tablatures, which are converted to text and fed as a prompt to ProgGP in a Google Colab notebookto generate continuations of rock and metal songs [83] (see Figure 1e). Notably, outputs contain not only guitar sections, but also bass and drums alongside the guitar, and are converted to MIDI to be added to music software for editing.

[7]The probabilities are described at https://saracardinalemusic.com/project/mark-of-markov/
[8]https://www.imdb.com/name/nm0002354/
[9]https://twitter.com/HEL9000ismetal
[10]Guitar tablatures are a music notation system designed specifically for guitarists.
[11]https://www.guitar-pro.com/

> Due to its characteristics, the interaction with ProgGP was mostly dictated by an initial need for isolated riffs, or musical ideas, that could be put together to form a full song. The process started with Jack experimenting on guitar to compose a riff. We started with a human-written riff to bring a bit of our own musical personalities to the generated riffs.
>
> After Pedro notated the initial riff into tabulature manually, we input the riff to our AI. We divided our workflow: Pedro took care of filtering continuations and feeding them back into the model to get variations; Jack started recording the initial riff on guitar to the computer, and adding drums and bass digitally. After Pedro filtered ideas, we both listened to the AI-outputs together and curated a few riffs we felt could be put together coherently. This step was particularly important because it was at this point that we envisioned an overall structure for the song based on the AI ideas.
>
> We then focused on recording these ideas. To enrich the song, we added extra layers using samples or new lead guitar parts. One particular AI-output had a distinctive drum beat generated alongside the guitar riff, which prompted us to explore samples that we wouldn't usually use for [the band's] music. Another section made us reflect on The Ocean's[12] aesthetics, prompting us to include a marimba and glockenspiel over a lead guitar part. Inspired by Periphery[13], we added a piano mimicking the melodic line of the guitar – it seemed like it might fit well with one of the AI-generated riffs.

4.1.4 P5: Lizzie. Lizzie creates experimental electronic dance music as a live coder – a genre where code is executed in real-time to produce sound and music. They use the domain-specific programming language Tidal Cycles [70], itself an extension of the functional programming language Haskell. Their chosen AI, Tidal-Fuzz, is a Markovian agent which outputs sequences of code by randomly walking through and choosing Tidal Cycles functions that form musical patterns [94]. These are integrated into the UI as suggestions to add to the music code cf. GitHub Co-Pilot (see Figure 1f).

> In some composition sessions, I started by generating ideas from Tidal-Fuzz. Other times I started from my own ideas.
>
> Where patterns were solely created by the human, some reflection came through errors made. For example, at one point, I was looking for a specific sample and typed the wrong number, which prompted me to explore a sample that I might've not considered. With patterns solely created by Tidal-Fuzz, reflection materialised in a few separate ways. Firstly, the agent's patterns were evaluated against my aesthetic preferences. The generation process of the agent has inbuilt metrics of modelling human aesthetic choices, however, a lot of reflection still occurred around evaluating whether these matched *my* aesthetic preferences. Which elements of the machine generated code were creating misalignment with my intentionality? In understanding the affective states driving my internal aesthetic evaluation function – through considered, deep listening (see Oliveros [77]) – I also was forming understandings of myself in relation to the music.
>
> Tidal-Fuzz's patterns sometimes prompted me to recursively hybridise, blending the machine's computational creativity with my own artistic insights, leveraging the strengths of both entities to

[12] https://www.theoceancollective.com/
[13] https://periphery.net/

> produce compositions that wouldn't be conceptualised by the human live coder alone. The creative impetus the agent provided, although helped steer the composition in new directions, also meant that there is some relinquishing of control. Perhaps, humans need to learn to accept some lack of control, viewing it not as a loss, but instead an exchange for new creative ideas.

4.1.5 P6: Lewis. Lewis is a composer and performer, including in the band Julia Set[14]. They typically create experimental computer music and contemporary classical. They chose the neural resonator plugin [35] (see Figure 1b) which uses neural networks to predict co-efficients for a resonant filter bank [36]. An audio or MIDI excitation can be input to the plugin and used to trigger feedback which propagates throughout the filter bank to synthesise different drum sounds. Moving parameters on the plugin's interface changes the shape of the drum (i.e. the filter bank co-efficients).

> In the first session, I felt that I was already too keen to think of structure and form. I questioned my relationship to the material I initially generated. As much as I was familiar with the Neural Resonator already, I was not able to clearly audiate (meaning to imagine sounds mentally) its product. This brings forth differences between my mentality as a composer (how do my actions affect my future self and what is my creative idea) and as an improviser (how do my actions affect my *present* self and what is the *performed* idea/instinct). As an improviser, I'm encouraged to respond to the material itself – the form of a work can then emerge without projecting expectations. Leveraging this mentality, I spent the second composition session generating material by improvising with the AI, using instinct. This enabled me to generate many threads of ideas from which to develop a composition.
>
> In the third composition session, I started stitching together and refining early ideas. My mentality shifted away from the instinctual and towards the considered. My creative decisions were no longer influenced by the AI, but were instead imposing themselves onto the material it had just generated.
>
> I find working in response to material a rewarding and successful methodology, although not without its detriments. I feel that this method is also important when working with new instruments, or instruments whose outcome is not always what is expected, where I do not have the same somatic or determinist relationship with them – my creative ideas can more easily arise through listening/interpreting than conceiving/enacting. One composes with material generated through immediate intuition, and attaches to that immediacy a reflective and cohesive narrative, which defines the composed expression.

4.1.6 P7: Kyle. Kyle is a media composer specialised in game audio. They chose the AI, CFEP [96], which transforms MIDI recordings to sound more human and expressive, based only on the musical features of pitch and note timing to support expressiveness when richer musical data is not available [96]. It combines transformer models trained on piano datasets to predict the velocity, timing and tempo for input MIDI files, outputting more human sounding adjustments of the input music (see Figure 1c).

> I began my composition. Although I couldn't use CFEP without having written initial material, it was interesting how the AI coloured my initial choices. I call this a butterfly effect where CFEP's design had unexpected knock-on impacts on my creative workflow. The first butterfly effect was in choosing piano – although common to my chosen genre, I also chose the instrument knowing that

[14] https://juliaset.bandcamp.com/

CFEP is trained on a piano dataset so would perform well on this type of data. Similarly, knowing that CFEP needs quantised values as input, I was neater in how I wrote the music than I would have been without the AI.

In session two, I began to experience the AI as a pseudo-co-producer, in the sense that the inclusion of it in the project influenced decisions that you make creatively as a mix engineer or composer. For example, I added staccato piano notes and drums, however, eventually disregarded these ideas because, in addition the vibe/feeling of the music not being correct, I knew CFEP would ultimately not work well on drums.

Once I introduced CFEP in session 3, I ran the piano through the algorithm and compared the expressive AI-output to the non-expressive AI-input. Surprisingly, I thought that the AI-output was good enough that I felt moved. I really did not expect to be moved by the piece, and I do not know why I found this quite moving. Perhaps, as I cannot play the piano well myself, but found the AI was playing it to an acceptable level, I thought "oh that is it, it's realising what I want to hear from my music" – buying into my ego as a composer [95].

4.2 Reflection Questionnaire

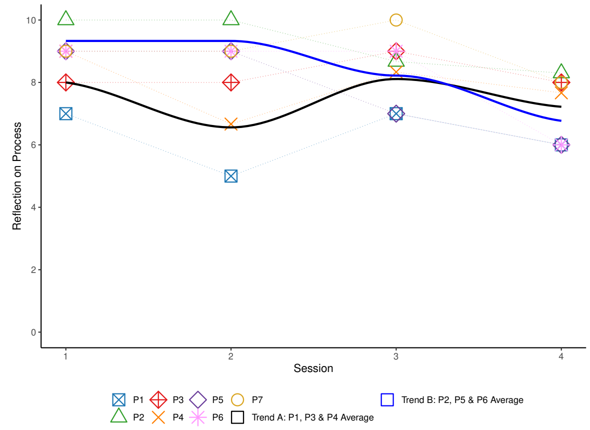
Figure 3 shows plots for the three RiCE [44] metrics retrospectively reported by our participants after each hour of composing: reflection-on-process (Figure 3a), reflection-on-self (Figure 3b) and reflection-through-experimentation (Figure 3c). To illustrate the changes in reflection over time, we plot curved mean average trend lines for subsets of participants in which we noted similar trends in their scores for Figure 3a and Figure 3b, and a linear trend line for Figure 3c.

For **reflection-on-process** (Figure 3a), P1, P3 and P4 show peaks in their scores in Session 1 and Session 3 (Trend A) – this suggests to us that they considered different directions to take their music at the start of their composing and before finalising their compositions. In contrast, we observed that P2, P5 and P6's scores generally decreased after Session 2 (Trend B) – this suggests to us that these participants considered alternative directions for their music early in their composition process. We also note that all the reflection-on-process scores are high, with none falling below four, suggesting the consideration of where to take a piece of music occurred throughout the music making process.

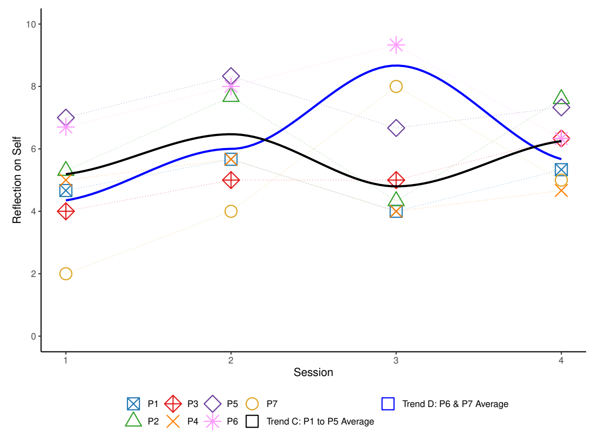
For **reflection-on-self** (Figure 3b), we observed that participants 1 through 5 showed peaks in Session 2 and Session 4 (Trend C). This suggests a temporal fluctuation in how the participants reflect on their personal experience. However, P6 and P7 gradually increase to a peak in Session 3 (Trend D), suggesting that they reflected-on-self at different points in their composition process to the other participants.

For **reflection-through-experimentation** (Figure 3c), we tentatively observe a decline over time (Trend E). Whilst this might be driven by the outlier P3 in Session 4 (who mostly took on production duties at this moment), we see a clear decline across sessions from P2 and P5 also. In Session 2, we note that the scores converge, and then diverge by Session 3. P6 and P7 annotated on the plot show high scores in Session 3; P1, P3 & P4 annotated on the plot show low scores in Session 3. This suggests that the changes in participants activity between Session 2 to Session 3 might have sparked or reduced reflection-through-experimentation.

To offer illustrative context for our RiCE scores, we compare the average of all participants' scores across sessions with RiCE scores calculated from the open-source dataset in Ford and Bryan-Kinns [44]. The open-source dataset is comprised of RiCE scores for CSTs including Photoshop, Word, and some DAWs (e.g. Cubase, Garageband, FL Studio, Ableton and Logic Pro). Table 3 compares the open-source scores along with the average of participant RiCE



(a) Reflection-on-process.



(b) Reflection-on-self.

(c) Reflection-through-experimentation.

Fig. 3. Plots of the participants' responses for the metrics from RiCE [44] for each hour of their composing.

Table 5. Mean average RiCE scores from open-source data in Ford and Bryan-Kinns [44] (top) and our study (bottom). DAWs scored in the RiCE survey [44] include Cubase (n=2), Garageband (n=2), Ableton (n=2), Logic and FL Studio.

Dataset	Process	Self	Experiment	RiCE
Open-source dataset [44]				
All CSTs (n=300)	7.4	6.1	6.9	6.8
MS Word Subset (n=43)	7.2	6.4	6.6	6.7
Photoshop Subset (n=42)	7.4	5.9	7.1	6.8
Visual Studio Subset (n=15)	8.0	7.1	7.3	7.5
DAWs Subset (n=8)	7.6	5.7	7.1	6.8
This paper				
Participant average (n=7)	8.1	5.8	6.6	6.8

scores in this paper. Overall, our participants have the same RiCE score as the open-source data (mean=6.8). Our participants also seem to show higher reflection-on-process scores, including for the subset of RiCE scores for DAWs. Reflection-on-process is also lower for the open-source DAW scores (mean=7.6) than our data (mean=8.1), whilst reflection-through-experimentation is higher for the open-source DAW scores (mean=7.1) than our data (mean=6.6).

4.3 Thematic Analysis

We generated six themes from our Thematic Analysis [9] across the participants' reflections: Theme 1) Reflection on Past Instincts; Theme 2) Reflection on Direction and Surprises; Theme 3) Reflection for AI; Theme 4) Reflection on Feelings; Theme 5) Reflection on Influences; and Theme 6) Reflection on Technical Challenges. We describe these below.

4.3.1 Theme 1: Reflection on Past Instincts. P1 and P6 curated AIGC by reflecting "in the moment" (P6) and using their instincts, creating environments where they could listen to AIGC in real-time – in a way that was more "improvisatory" (P6) and "instinctual" (P6). P6 found this "felt quite familiar", similar to how they would improvise in their music practice, whilst P1 said the process reminded them of playing in Jazz ensembles. Furthermore, P6 described this process as deliberate: they split their process into choosing material based on their instinct in Session 2 and then reflected on their *past decisions* when organising this material in Session 3. Indeed, P6 describes this as reflection on their *past self* in their first-person account. Similarly, P5 also describes reflection-on-self when curating material by listening in real-time and live coding. They said they were "forming understandings of [themselves] in relation to the music" (P5), and reflected on how AIGC matched their aesthetic: "the things that [the AI] was producing weren't necessarily in my aesthetic, so then it was a case of refining what it was that I actually wanted" (P5).

4.3.2 Theme 2: Reflection on Direction and Surprises. All our composers reflected on the direction in which to take their music. For example, P5 found the AI "pushes me in different directions or gets me thinking about doing things in a different way that I haven't thought about myself". The direction to take a piece was also sparked by reflection-on-surprise, from both the AI and other aspects of composers' workflows. For example, P1 found their AI "really worked as like a surprise prompt" helping them to continue "taking risks and experimenting". P5 tested a Cowbell sample they usually wouldn't use in their practice, noting that "this surprise moment was[...] crucial for building [their music]".

4.3.3 Theme 3: Reflection for AI. P1 and P7 reflected on how their current actions might integrate with their chosen AI tool. P7 described their AI as a producer – musical material fed to the AI would either work/not work. P1 explained that, similar to when you compose for performers and shape your composition to what people can physically play on their instruments, you shape your composition to the AI and its affordances – "you have to take into consideration things like what people can physically play[...] so they kind of shape your composition cause of the limitations[...] I think it's quite similar [with AI]" (P1). P7 went so far as to describe this as a butterfly effect where their compositional decisions were limited to those that would work well from the start with the AI they chose.

4.3.4 Theme 4: Reflection on Feelings.

P2 and P7 reflected on their feelings of using the AI in their practice. P2 felt bad about changing outputs from their AI system, describing self-awareness and feelings of imposter syndrome, e.g., when comparing themselves against the famous composer John Williams in their first-person account. A different emotional response was from P7, where their chosen AI, CFEP, transformed their music to sound more humanistic, and they were "surprisingly moved" that the AI could play their music in a way that they could not.

4.3.5 Theme 5: Reflection on Influences.

From our first-person accounts, we identified several inspirations which the composers reflected on to inform their creative practice. There were references to literature from musicians and philosophers such as John Croft [28], John Chowing [26], Pauline Oliveros [77] and Periphery (see P3 & P4's first-person account). It was possible to trace the musicians' creative influences to ideas implemented into their practice. For example, P1's negotiation between AI and non-AI mirrors Croft's [28] philosophising on levels of control to afford in musical improvisations, whilst P3 & P4's use of Piano was inspired by Periphery.

4.3.6 Theme 6: Reflection on Technical Challenges.

Unsurprisingly, the composers reflected on challenges they came across to be able to integrate their chosen AI into their workflow, either: needing to format data to move between their instruments and the AI inputs and outputs (P1, P3, P4, P7), using templates to avoid complex setup (P1, P2, P6), or tweaking code whilst music making (P2, P6, P7).

5 DISCUSSION

In this exploratory study, we gathered first-person accounts, questionnaire measures and interviews about the processes of creating six music compositions, each written by composers integrating a unique AI tool into their typical music making practice. These first-person accounts helped us in understanding a plurality of AI tools explored in this paper – offering insights on multiple AI tools to complement studies on individual AI tools [75].

There are many variations in the data collected. For example, our composers all have different practices and approaches to music composition, work in different genres, and used different AI tools. The set of AI tools used are heterogeneous and nested within a complex ecosystem of software and hardware [69] – we acknowledge conflation between various types of tools in our findings. Nevertheless, we see our work as generative, suggesting directions for future work. We do not claim our findings generalise without further investigation but suggest that they capture qualities of a plurality of our artist-researchers' [91] real-world practice. Our approach of first-person accounts also suits making practices where a heterogeneity of different tools is the norm – for example, see the range of tools across the live coding community [1, 70].

Below, we first discuss trends across our data (Section 5.1). Second, we discuss the unique ways our participants reflected (Section 5.2). We then reflect on our method and its limitations (Section 5.3). Key takeaways are offered in Section 5.4. Throughout, we situate our findings within our literature on reflection and co-creative AI (see Section 2).

5.1 Discussion on Trends

Our exploration identified fresh insights in relation to the RiCE metrics and the temporal nature of our participants' reflection. Some findings we suggest extend current understandings on AI music making (Section 5.1.1), whilst others are confirmatory (Section 5.1.2).

5.1.1 Novel Trends. We identified possible explanations for our observed trends by comparing the subplots in Figure 3 with the first-person accounts. For example, we observed that when reflection-on-process was high (Figure 3a), participants were listening to music in real-time – such as P1 improvising to select AIGC in Session 3, or P3 & P4 improvising to select musical layers in Session 3.

In contrast, higher reflection-on-self scores (Figure 3a) occurred at moments when participants were *arranging* their, sometimes AI-generated, ideas. Notably, P6's high reflection-on-self score in Session 3 connects with their description of self-reflection in Theme 1 (Reflection on Past Instincts) – they reflected on the instinctual decisions that their *past self* had created in the previous session, learning about themselves by analysing their choices retrospectively. This connects to Candy's reflection-at-a-distance [17] as, perhaps, P6 was purposefully distancing themselves from their earlier decisions to assess their work from a more objective viewpoint. Or, a different perspective could be that P6 was analysing their perspectives to assign meaning to their AIGC, corroborating Yurman and Reddy [97].

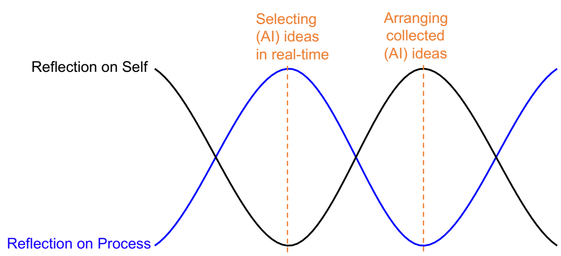


Fig. 4. A speculative model showing the trade-off relationship we observed for reflection-on-process and reflection-on-self, when people were either selecting ideas whilst listening in real-time or arranging ideas after curation.

Between reflection-on-process and reflection-on-self, we tentatively suggest from our observations that there *might* be a potential trade-off relationship, which we visualise in Figure 4. When selecting AI-outputs and listening in real-time, participants tended to reflect on future directions to take their music, whilst reflecting on what their music means to them when combining AIGC. Future work is needed to support our limited evidence. However, this model supports definitions of reflection as moments where people sit back in contemplation [72, 93], and descriptions of its push-and-pull with moments of more instinctual reflection-in-the-moment [17]. Future work could explore whether this model is music specific, e.g., high reflection-on-process when selecting AIGC might occur in any ideation stage.

5.1.2 Confirmatory Trends. Reflection-on-process remained high across all sessions, and was higher in comparison to other RiCE scores for software in Table 5, implying that participants often reflected on alternative directions to take their music. The high reflection-on-process scores are supported by our findings from Theme 2 (Reflection on Direction and Surprises), suggesting that participants would leverage surprising outputs to change the direction of their music. This corroborates Caramiaux and Fdili Alaoui's [21] findings that AI-artists leveraged surprising outputs in their creative process. However, we cannot attribute reflection-on-surprise to the AI tools exclusively because P5 reflected on a surprising cowbell sample they found in their live-coding environment, without using their AI. We also find reflection on glitches in some non-AI composition practices [32, 50].

Reflection-through-experimentation (Figure 3c) generally decreased over time, with participants converging in Session 2 and diverging in Session 3. A possible explanation for the convergence comes from Theme 6 (Reflection on Technical Challenges). Perhaps, some participants needed to first reflect on technical issues (cf. Candy's [17] reflection-for-action) in Session 1 before they could experiment with AIGC. Furthermore, we suggest that reflection-through-experimentation decreased in Session 3 for some participants (in Figure 3c) because they had already curated and decided how to organise their AIGC, no longer needing to experiment. P6 and P7, with high reflection-through-experimentation in Session 3, showed more unique stories (see below).

5.2 Discussion on Individual Findings

The first-person accounts enabled us to explore the individual ways that participants reflected on their composition sessions, and the unique ways that their tools led to different reflective experiences. For example, P7's chosen AI, CFEP [96], was the only example of an AI which required musical material to be written beforehand. This might explain their high levels of reflection-through-experimentation in Session 3 (in Figure 3c), where they first introduced CFEP to their process and started experimenting with its output.

P3 & P4 also offer unique insights as an example of collaborative practice. We observe how they split tasks to effectively navigate the co-creative AI process [53, 73]. For example, P4 initially generated ideas whilst P3 prepared the music software cf. reflection-for-action [17]. Furthermore, we note that their curating of AIGC was completed together, possibly highlighting the importance of selecting AI-outputs which had significance to both band members, and more closely mimicking how rock bands make music without AI by jamming [7]. We also compare this to how Yurman and Reddy [97] assigned meaning to AIGC within their study.

P2 was also unique in that their AI tool, MoM, did not require any input and only output music to their software. This gives a potential explanation for the descriptions in their first-person account on how it was easier to compose with smaller chunks of AIGC; for P2, editing the recording to be able to curate different ideas was a necessity – likely, as focused on film music, to identify interesting motifs [68]. P2 also does not mention challenges presented by tools which required inputs as in Theme 3 (Reflection for AI). For example, where P7 avoided writing material for drums as CFEP was trained on a dataset of piano music and might not perform well. Whilst our examples above do not directly mirror how Lewis's [61] art style was informed by AI suggestions, we suggest our findings echo their recommendation to consider the origin of an AI's data and its influence on people's creative practice.

A notable finding from our Thematic Analysis was in Theme 4 (Reflection on Feelings), which related to two participants' unique emotional responses to the AI. P2 suggested that their AI helped them to overcome imposter syndrome by providing material for them to extend. A different emotional response was from P7 who notes that their AI played music at a higher standard than themselves, helping them to realise their music beyond their own abilities. This corroborates Sturm [92] who found their AI helped them to express ideas they could not yet realise.

5.3 Discussion on Method & Limitations

Our first-person accounts offer insights into using AIGC in music composition but includes many variations from participant background, to the tools used, to the compositional techniques, and different genres, preventing statistical analysis or generalisation from our findings. There is clearly an opportunity design more controlled A/B tests to untangle these factors in future work. For example, the brief comparison between different RiCE measures in Table 5 suggests that AI encourages more reflection-on-process, which a more controlled A/B study design could examine. Future work could also unpack our findings further by focusing on different groupings of tools e.g. examining only the timbre-focused tools used such as RAVE [15] and neural resonator [35].

The first-person accounts helped us in speculating on explanations for patterns we observed from the RiCE [44] questionnaire measures, and brought complementary insights to the findings from our interviews. For example, the citations to researcher's inspirations in the first-person accounts are not captured by the RiCE questionnaire [44], nor other questionnaire measures typically used in creative HCI research [25, 59]. Investigating the impact of these more artistic influences and how to capture nuance in creative HCI and AI contexts might be fruitful future work.

We further note that our interview data contained many similar insights to the first-person accounts. Without conducting the interviews, participants could have completed the study at any time without the researcher needing to be present. In this case, participants could then also choose which moments to pause and reflect back on whilst composing, which would be more conducive to researchers exploring aspects of creative user experiences where interruptions pose a confounding variable such as flow states [29]. On the other hand, our structured activity meant that we were able to collect data from the unique perspectives of artist-researchers who have limited time to be creative [91], with the interviews acting as a useful contingency in case participants could not complete their first-person accounts. Our method contrasts other methods such as diary studies [8, 31] or autoethnography [61, 65, 76, 87], where more commitment is typically required from participants. Diary studies also tend to capture immediate thoughts, whilst our method required participants to retrospectively synthesise their thinking into first-person accounts. Nonetheless, we found that collecting and comparing a range of first-person perspectives was helpful and propose that a method which can capture multiple personal insights in a consistent way could complement research using current HCI methods.

5.4 Key Takeaways

We offer the following key takeaways from our research. The first-person accounts offer rich descriptions of a plurality of AI-based music composition practices, which could inform others' AI music making. This was enabled by our reflection board method, which could be applied by CST researchers to capture personal, subjective accounts. The findings of our RiCE metrics could be used to make numerical comparison with other reflective CSTs, supporting further research on reflective CSTs. By collecting RiCE metrics over time, we also learnt that human-AI collaboration in music making presents different types of reflection at different stages. Our hypothesised model in Figure 4 based on RiCE could be tested by musicians to leverage either reflection-on-self or reflection-on-process (cf. P6) to spark inspiration or ideation at different times, or leveraged by CST designers to scaffold modes of a reflective CST.

6 CONCLUSION

This paper explored how composers reflect across a range of AI music tools and composition approaches. We recruited artist-researchers with music and AI skills and tasked them with composing a piece of music using an AI tool of their choice. We contribute six first-person accounts from their practice, gathered through a novel data collection approach using *reflection boards*, where participants were asked to pause and reflect back on screenshots of their composing after every hour, triangulated with interviews and questionnaire measures. We suggest that our insights contribute to furthering the current understandings of how people reflect in creative user experiences.

CONTRIBUTION STATEMENT

Corey Ford contributed the conceptualisation, methodology, investigation, formal analysis, data curation, visualisation, project administration, writing of the original draft, and editing. **Ashley Noel-Hirst**, **Sara Cardinale**, **Jackson Loth**, **Pedro Sarmento**, **Lewis Wolstanholme**, **Elizabeth Wilson** and **Kyle Worrall** each contributed their music composition and first-person accounts. **Nick-Bryan Kinns** contributed as project supervisor and editor.

ACKNOWLEDGMENTS

Corey Ford, Ashley Noel-Hirst, Sara Cardinale, Jackson Loth, Pedro Sarmento and Lewis Wolstanholme are supported by the EPSRC UKRI Centre for Doctoral Training in Artificial Intelligence and Music (AIM) [EP/S022694/1]. Elizabeth Wilson is supported by the EPSRC and AHRC Centre for Doctoral Training in Media and Arts Technology (MAT) [EP/L01632X/1]. Kyle Worrall is supported by the EPSRC Centre for Doctoral Training in Intelligent Games & Games Intelligence (IGGI) [EP/S022325/1]. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

REFERENCES

[1] Samuel Aaron and Alan F. Blackwell. 2013. From Sonic Pi to Overtone: Creative Musical Experiences with Domain-specific and Functional Languages. In *Proceedings of the First ACM SIGPLAN Workshop on Functional Art, Music, Modeling & Design* (Boston, Massachusetts, USA) *(FARM '13)*. Association for Computing Machinery, New York, NY, USA, 35–46. https://doi.org/10.1145/2505341.2505346

[2] Berker Banar and Simon Colton. 2021. Generating Music with Extreme Passages using GPT-2. *Evo* 2021 Late Breaking Abstracts* (2021), 31. https://arxiv.org/pdf/2106.11804.pdf#page=34

[3] Eric P.S. Baumer, Vera Khovanskaya, Mark Matthews, Lindsay Reynolds, Victoria Schwanda Sosik, and Geri Gay. 2014. Reviewing Reflection: On the Use of Reflection in Interactive System Design. In *Proceedings of the 2014 Conference on Designing Interactive Systems* (Vancouver, BC, Canada) *(DIS '14)*. Association for Computing Machinery, New York, NY, USA, 93–102. https://doi.org/10.1145/2598510.2598598

[4] Steve Benford, Chris Greenhalgh, Andy Crabtree, Martin Flintham, Brendan Walker, Joe Marshall, Boriana Koleva, Stefan Rennick Egglestone, Gabriella Giannachi, Matt Adams, Nick Tandavanitj, and Ju Row Farr. 2013. Performance-Led Research in the Wild. *ACM Transations in Computer-Human Interaction* 20, 3, Article 14 (jul 2013), 22 pages. https://doi.org/10.1145/2491500.2491502

[5] Jesse Josua Benjamin, Heidi Biggs, Arne Berger, Julija Rukanskaité, Michael B. Heidt, Nick Merrill, James Pierce, and Joseph Lindley. 2023. The Entoptic Field Camera as Metaphor-Driven Research-through-Design with AI Technologies. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 178, 19 pages. https://doi.org/10.1145/3544548.3581175

[6] Marit Bentvelzen, Paweł W. Woźniak, Pia S.F. Herbes, Evropi Stefanidi, and Jasmin Niess. 2022. Revisiting Reflection in HCI: Four Design Resources for Technologies That Support Reflection. *ACM Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (March 2022). https://doi.org/10.1145/3517233

[7] Michele Biasutti. 2012. Group Music Composing Strategies: A Case Study Within a Rock Band. *British Journal of Music Education* 29, 3 (2012), 343–357. https://doi.org/10.1017/S0265051712000289

[8] Marion Botella, Jessica Nelson, and François Zenasni. 2019. It Is Time to Observe the Creative Process: How to Use a Creative Process Report Diary (CRD). *Journal of Creative Behavior* 53, 2 (2019), 211–221. https://doi.org/10.1002/jocb.172

[9] V Braun and V Clarke. 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3, 2 (Jan. 2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[10] V. Braun and V. Clarke. 2013. *Successful Qualitative Research: A Practical Guide for Beginners*. SAGE Publications.

[11] Jean-Pierre Briot, Gaëtan Hadjeres, and François-David Pachet. 2017. Deep Learning Techniques for Music Generation – A Survey. https://doi.org/10.48550/arxiv.1709.01620

[12] Nick Bryan-Kinns, Berker Banar, Corey Ford, Courtney N. Reed, Yixiao Zhang, Simon Colton, and Jack Armitage. 2021. Exploring XAI for the Arts: Explaining Latent Space in Generative Music. In *1st Workshop on eXplainable AI Approaches for Debugging and Diagnosis (XAI4Debugging@NeurIPS2021)*. https://xai4debugging.github.io/files/papers/exploring_xai_for_the_arts_exp.pdf

[13] Nick Bryan-Kinns, Corey Ford, Alan Chamberlain, Steven David Benford, Helen Kennedy, Zijin Li, Wu Qiong, Gus G. Xia, and Jeba Rezwana. 2023. Explainable AI for the Arts: XAIxArts. In *Proceedings of the 15th Conference on Creativity and Cognition (C&C '23)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3591196.3593517

[14] Nick Bryan-Kinns and Courtney N. Reed. 2023. *A Guide to Evaluating the Experience of Media and Arts Technology*. Springer International Publishing, Cham, 267–300. https://doi.org/10.1007/978-3-031-31360-8_10

[15] Antoine Caillon and Philippe Esling. 2021. RAVE: A Variational Autoencoder for Fast and High-quality Neural Audio Synthesis. *arXiv preprint* (2021). https://arxiv.org/abs/2111.05011

[16] Linda Candy. 2011. Research and Creative Practice. In *Interacting: Art, Research and the Creative Practitioner* (1st ed.), Ernest A. Edmonds and Linda Candy (Eds.). Libri Publishing UK, 33–59.

[17] Linda Candy. 2019. *The Creative Reflective Practitioner: Research Through Making and Practice* (1st ed.). Routledge. https://doi.org/10.4324/9781315208060

[18] L. Candy, S. Amitani, and Z. Bilda. 2006. Practice-led Strategies for Interactive Art Research. *CoDesign* 2, 4 (2006), 209–223. https://doi.org/10.1080/15710880601007994

[19] Linda Candy and Ernest Edmonds. 2018. Practice-based Research in the Creative Arts: Foundations and Futures from the Front Line. *Leonardo* 51, 1 (2018), 63–69. https://doi.org/10.1162/LEON_a_01471

[20] Baptiste Caramiaux and Sarah Fdili Alaoui. 2022. "Explorers of Unknown Planets": Practices and Politics of Artificial Intelligence in Visual Arts. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2, Article 477 (Nov 2022), 24 pages. https://doi.org/10.1145/3555578

[21] Baptiste Caramiaux, Fabien Lotte, Joost Geurts, Giuseppe Amato, Malte Behrmann, Frédéric Bimbot, Fabrizio Falchi, Ander Garcia, Jaume Gibert, Guillaume Gravier, Hadmut Holken, Hartmut Koenitz, Sylvain Lefebvre, Antoine Liutkus, Andrew Perkis, Rafael Redondo, Enrico Turrin, Thierry Viéville, and Emmanuel Vincent. 2019. *AI in the Media and Creative Industries*. Research Report. New European Media (NEM). 1–35 pages. https://inria.hal.science/hal-02125504

[22] Filippo Carnovalini and Antonio Rodà. 2020. Computational Creativity and Music Generation Systems: An Introduction to the State of the Art. *Frontiers in Artificial Intelligence* 3 (2020). https://doi.org/10.3389/frai.2020.00014

[23] Alan Chamberlain, Andy Crabtree, Tom Rodden, Matt Jones, and Yvonne Rogers. 2012. Research in the Wild: Understanding 'in the wild' Approaches to Design and Development. In *Proceedings of the Designing Interactive Systems Conference* (Newcastle Upon Tyne, United Kingdom) *(DIS '12)*. Association for Computing Machinery, New York, NY, USA, 795–796. https://doi.org/10.1145/2317956.2318078

[24] Minsuk Chang, Stefania Druga, Alexander J. Fiannaca, Pedro Vergani, Chinmay Kulkarni, Carrie J Cai, and Michael Terry. 2023. The Prompt Artists. In *Proceedings of the 15th Conference on Creativity and Cognition* (Virtual Event, USA) *(C&C '23)*. Association for Computing Machinery, New York, NY, USA, 75–87. https://doi.org/10.1145/3591196.3593515

[25] Erin Cherry and Celine Latulipe. 2014. Quantifying the Creativity Support of Digital Tools through the Creativity Support Index. *ACM Transactions on Computer-Human Interaction* 21, 4, Article 21 (jun 2014), 25 pages. https://doi.org/10.1145/2617588

[26] John Chowning. 1977. Stria. Commissioned by IRCAM (Paris) for the Institute's first major concert series: Perspectives of the 20th Century.

[27] David Collins. 2007. Real-time Tracking of the Creative Music Composition Process. *Digital Creativity* 18, 4 (Dec. 2007), 239–256. https://doi.org/10.1080/14626260701743234

[28] John Croft. 2007. Theses On Liveness. *Organised Sound* 12, 1 (2007), 59–66. https://doi.org/10.1017/S1355771807001604

[29] Mihály Csíkszentmihályi. 1990. *Flow: The Psychology of Optimal Experience.* Harper Collins, New York, USA.

[30] Zihang Dai, Zhilin Yang, Yiming Yang, Jaime Carbonell, Quoc V. Le, and Ruslan Salakhutdinov. 2019. Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context. arXiv:1901.02860 [cs.LG]

[31] Peter Dalsgaard and Kim Halskov. 2012. Reflective Design Documentation. In *Proceedings of the Designing Interactive Systems Conference* (Newcastle Upon Tyne, United Kingdom) *(DIS '12)*. Association for Computing Machinery, New York, NY, USA, 428–437. https://doi.org/10.1145/2317956.2318020

[32] Teodoro Dannemann, Nick Bryan-Kinns, and Andrew McPherson. 2023. Self-Sabotage Workshop: A Starting Point to Unravel Sabotaging of Instruments as a Design Practice. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Miguel Ortiz and Adnan Marquez-Borbon (Eds.). Mexico City, Mexico, 70–78. http://nime.org/proceedings/2023/nime2023_9.pdf

[33] Nicholas Davis. 2021. Human-Computer Co-Creativity: Blending Human and Computational Creativity. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* 9, 6 (Jun. 2021), 9–12. https://doi.org/10.1609/aiide.v9i6.12603

[34] Sebastian Deterding, Jonathan Hook, Rebecca Fiebrink, Marco Gillies, Jeremy Gow, Memo Akten, Gillian Smith, Antonios Liapis, and Kate Compton. 2017. Mixed-Initiative Creative Interfaces. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 628–635. https://doi.org/10.1145/3027063.3027072

[35] Rodrigo Diaz. 2024. *Neural Resonator VST: Generate and Use Filters based on Arbitrary 2D Shapes and Materials.* https://github.com/rodrigodzf/NeuralResonatorVST (Online).

[36] Rodrigo Diaz, Ben Hayes, Charalampos Saitis, György Fazekas, and Mark Sandler. 2022. Rigid-Body Sound Synthesis with Differentiable Modal Resonators. *arXiv preprint* (2022). https://doi.org/10.48550/arXiv.2210.15306

[37] Elisabeth T. Kersten-van Dijk, Joyce H. D. M. Westerink, Femke Beute, and Wijnand A. IJsselsteijn. 2017. Personal Informatics, Self-Insight, and Behavior Change: A Critical Review of Current Literature. *Human–Computer Interaction* 32, 5-6 (2017), 268–296. https://doi.org/10.1080/07370024.2016.1276456

[38] Carolyn Ellis, Tony E. Adams, and Arthur P. Bochner. 2011. Autoethnography: An Overview. *Historical Social Research / Historische Sozialforschung* 36, 4 (138) (2011), 273–290. http://www.jstor.org/stable/23032294

[39] Sarah Fdili Alaoui. 2019. Making an Interactive Dance Piece: Tensions in Integrating Technology in Art. In *Proceedings of the 2019 Designing Interactive Systems Conference (DIS '19)*. Association for Computing Machinery, New York, NY, USA, 1195–1208. https://doi.org/10.1145/3322276.3322289

[40] Sara Fdili Alaoui. 2023. *Dance-Led Research.* Computer Science. Université Paris Saclay (COMUE). https://inria.hal.science/tel-04059520/

[41] Rowanne Fleck and Geraldine Fitzpatrick. 2010. Reflecting on Reflection: Framing a Design Landscape. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction* (Brisbane, Australia) *(OZCHI '10)*. Association for Computing Machinery, New York, NY, USA, 216–223. https://doi.org/10.1145/1952222.1952269

[42] Göran Folkestad, David J. Hargreaves, and Berner Lindström. 1998. Compositional Strategies in Computer-Based Music-Making. *British Journal of Music Education* 15, 1 (March 1998), 83–97. https://doi.org/10.1017/S0265051700003788 Publisher: Cambridge University Press.

[43] Corey Ford and Nick Bryan-Kinns. 2022. Speculating on Reflection and People's Music Co-Creation with AI. In *Generative AI and HCI Workshop at CHI 2022*. https://qmro.qmul.ac.uk/xmlui/handle/123456789/80144

[44] Corey Ford and Nick Bryan-Kinns. 2023. Towards a Reflection in Creative Experience Questionnaire. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 763, 16 pages. https://doi.org/10.1145/3544548.3581077

[45] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the Landscape of Creativity Support Tools in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3290605.3300619

[46] Mafalda Gamboa, Michael James Heron, Miriam Sturdee, and Pauline H Belford. 2023. Screenshots as Photography in Gamescapes: An Annotated Psychogeography of Imaginary Places. In *Proceedings of the 15th Conference on Creativity and Cognition* (Virtual Event, USA) *(C&C '23)*. Association for Computing Machinery, New York, NY, USA, 506–518. https://doi.org/10.1145/3591196.3593370

[47] 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* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 233–240. https://doi.org/10.1145/642611.642653

[48] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Nets. *Advances in Neural Information Processing Systems* 27 (2014).

[49] A. M. Grant, J. Franklin, and P. Langford. 2002. The Self-Reflection and Insight Scale: A New Measure of Private Self-consciousness. *Social Behavior and Personality: An International Journal* 30, 8 (2002), 821–836. https://doi.org/10.2224/sbp.2002.30.8.821

[50] Andy Hamilton. 2020. The Aesthetics of Imperfection Reconceived: Improvisations, Compositions, and Mistakes. *The Journal of Aesthetics and Art Criticism* 78, 3 (2020), 289–302. https://doi.org/10.1111/jaac.12749

[51] Dorien Herremans, Ching-Hua Chuan, and Elaine Chew. 2017. A Functional Taxonomy of Music Generation Systems. *Comput. Surveys* 50, 5, Article 69 (Sep 2017), 30 pages. https://doi.org/10.1145/3108242

[52] Kristina Höök, Baptiste Caramiaux, Cumhur Erkut, Jodi Forlizzi, Nassrin Hajinejad, Michael Haller, Caroline CM Hummels, Katherine Isbister, Martin Jonsson, George Khut, et al. 2018. Embracing First-person Perspectives in Soma-based Design. In *Informatics*, Vol. 5. MDPI, 8.

[53] Cheng-Zhi Anna Huang, Hendrik Vincent Koops, Ed Newton-Rex, Monica Dinculescu, and Carrie J. Cai. 2020. AI Song Contest: Human-AI Co-Creation in Songwriting. In *21st International Society for Music Information Retrieval Conference*. Montréal, Canada. https://arxiv.org/pdf/2010.05388.pdf

[54] Cheng-Zhi Anna Huang, Ashish Vaswani, Jakob Uszkoreit, Noam Shazeer, Curtis Hawthorne, Andrew M Dai, Matthew D Hoffman, and Douglas Eck. 2018. Music Transformer: Generating Music with Long-Term Structure. *arXiv preprint* (2018). https://doi.org/10.48550/arXiv.1809.04281

[55] Layne Jackson Hubbard, Yifan Chen, Eliana Colunga, Pilyoung Kim, and Tom Yeh. 2021. Child-Robot Interaction to Integrate Reflective Storytelling Into Creative Play. In *Proceedings of the 13th Conference on Creativity and Cognition* (Virtual Event, Italy) *(C&C '21)*. Association for Computing Machinery, New York, NY, USA, Article 16, 8 pages. https://doi.org/10.1145/3450741.3465254

[56] Alexander Refsum Jensenius and Michael J. Lyons (Eds.). 2017. *A NIME Reader*. Current Research in Systematic Musicology, Vol. 3. Springer International Publishing, Cham. http://link.springer.com/10.1007/978-3-319-47214-0

[57] Martin Jonsson and Jakob Tholander. 2022. Cracking the Code: Co-coding with AI in Creative Programming Education. In *Proceedings of the 14th Conference on Creativity and Cognition* (Venice, Italy) *(C&C '22)*. Association for Computing Machinery, New York, NY, USA, 5–14. https://doi.org/10.1145/3527927.3532801

[58] Théo Jourdan and Baptiste Caramiaux. 2023. Machine Learning for Musical Expression: A Systematic Literature Review. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Mexico City, Mexico. https://hal.science/hal-04075492

[59] Andruid Kerne, Andrew M. Webb, Celine Latulipe, Erin Carroll, Steven M. Drucker, Linda Candy, and Kristina Höök. 2013. Evaluation Methods for Creativity Support Environments. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 3295–3298. https://doi.org/10.1145/2468356.2479670

[60] Diederik P Kingma and Max Welling. 2013. Auto-Encoding Variational Bayes. https://doi.org/10.48550/ARXIV.1312.6114

[61] Makayla Lewis. 2023. AIxArtist: A First-person Tale of Interacting with Artificial Intelligence to Escape Creative Block. In *Proceedings of the 1st International Workshop on Explainable AI for the Arts (XAIxArts), ACM Creativity and Cognition (C&C) 2023*. https://arxiv.org/abs/2308.11424

[62] Makayla Lewis, Miriam Sturdee, Mafalda Gamboa, and Denise Lengyel. 2023. Doodle Away: An Autoethnographic Exploration of Doodling as a Strategy for Self-Control Strength in Online Spaces. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 414, 13 pages. https://doi.org/10.1145/3544549.3582747

[63] Jackson Loth, Pedro Sarmento, CJ Carr, Zack Zukowski, and Mathieu Barthet. 2023. ProgGP: From GuitarPro Tablature Neural Generation to Progressive Metal Production. In *Proceedings of the 16th International Symposium on Computer Music Multidisciplinary Research*. https:

//arxiv.org/abs/2307.05328

[64] Todd Lubart. 2005. How Can Computers Be Partners in the Creative Process: Classification and Commentary on the Special Issue. *International Journal of Human-Computer Studies* 63, 4 (Oct. 2005), 365–369. https://doi.org/10.1016/j.ijhcs.2005.04.002

[65] Andrés Lucero. 2018. Living Without a Mobile Phone: An Autoethnography. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 765–776. https://doi.org/10.1145/3196709.3196731

[66] Andrés Lucero, Audrey Desjardins, Carman Neustaedter, Kristina Höök, Marc Hassenzahl, and Marta E. Cecchinato. 2019. A Sample of One: First-Person Research Methods in HCI. In *Companion Publication of the Designing Interactive Systems Conference 2019* (San Diego, CA, USA) *(DIS '19 Companion)*. Association for Computing Machinery, New York, NY, USA, 385–388. https://doi.org/10.1145/3301019.3319996

[67] Thor Magnusson. 2019. *Sonic Writing: Technologies of Material, Symbolic, and Signal Inscriptions.* Bloomsbury Publishing USA.

[68] Stephen McAdams. 2004. Problem-Solving Strategies in Music Composition: A Case Study. *Music Perception* 21, 3 (2004), 391–429. https://doi.org/10.1525/mp.2004.21.3.391

[69] Glenn McGarry, Peter Tolmie, Steve Benford, Chris Greenhalgh, and Alan Chamberlain. 2017. "They're All Going out to Something Weird": Workflow, Legacy and Metadata in the Music Production Process. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 995–1008. https://doi.org/10.1145/2998181.2998325

[70] A. McLean and Geraint Wiggins. 2010. Tidal–Pattern Language for the Live Coding of Music. In *Proceedings of the 7th Sound and Music Computing Conference*. 331–334.

[71] David R. Millen. 2000. Rapid Ethnography: Time Deepening Strategies for HCI Field Research. In *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (New York City, New York, USA) *(DIS '00)*. Association for Computing Machinery, New York, NY, USA, 280–286. https://doi.org/10.1145/347642.347763

[72] Jennifer A Moon. 2013. *Reflection in Learning and Professional Development: Theory and Practice.* Routledge, New York, USA.

[73] Michael Muller, Justin D Weisz, and Werner Geyer. 2020. Mixed Initiative Generative AI Interfaces: An Analytic Framework for Generative AI Applications. In *Proceedings of the 'The Future of Co-Creative Systems-A Workshop on Human-ComputerCo-Creativity" at the 11th International Conference on Computational Creativity (ICCC 2020)*.

[74] Daniel Müllensiefen, Bruno Gingras, Jason Musil, and Lauren Stewart. 2014. The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLOS ONE* 9, 2 (Feb. 2014), 1–23. https://doi.org/10.1371/journal.pone.0089642

[75] Molly Jane Nicholas, Sarah Sterman, and Eric Paulos. 2022. Creative and Motivational Strategies Used by Expert Creative Practitioners. In *Proceedings of the 14th Conference on Creativity and Cognition* (Venice, Italy) *(C&C '22)*. Association for Computing Machinery, New York, NY, USA, 323–335. https://doi.org/10.1145/3527927.3532870

[76] Ashley Noel-Hirst and Nick Bryan-Kinns. 2023. An Autoethnographic Exploration of XAI in Algorithmic Composition. In *The 1st International Workshop on Explainable AI for the Arts (XAIxArts), ACM Creativity and Cognition (C&C) 2023.* https://arxiv.org/abs/2308.06089

[77] Pauline Oliveros. 2005. *Deep Listening: A Composer's Sound Practice.* IUniverse.

[78] OpenAI. 2022. *ChatGPT: Optimizing Language Models for Dialogue.* https://openai.com/blog/chatgpt/ (Online).

[79] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, and others. 2019. Language Models are Unsupervised Multitask Learners. *OpenAI Blog* 1, 8 (2019), 9.

[80] Jeba Rezwana and Mary Lou Maher. 2022. Designing Creative AI Partners with COFI: A Framework for Modeling Interaction in Human-AI Co-Creative Systems. *ACM Transactions on Computer-Human Interaction* (Feb. 2022). https://doi.org/10.1145/3519026

[81] Adam Roberts, Jesse Engel, Colin Raffel, Curtis Hawthorne, and Douglas Eck. 2018. A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music. In *Proceedings of the 35th International Conference on Machine Learning (ICML '18)*, Vol. 80. PMLR, Stockholm, Sweden. https://proceedings.mlr.press/v80/roberts18a/roberts18a.pdf

[82] P. Sarmento, A. Kumar, CJ Carr, Z. Zukowski, M. Barthet, and Y. Yang. 2021. DadaGP: A Dataset of Tokenized GuitarPro Songs for Sequence Models. In *Proceedings of the 22nd International Society for Music Information Retrieval Conference* (Online). https://arxiv.org/abs/2107.14653

[83] Pedro Sarmento, Adarsh Kumar, Yu-Hua Chen, CJ Carr, Zack Zukowski, and Mathieu Barthet. 2023. GTR-CTRL: Instrument and Genre Conditioning for Guitar-focused Music Generation with Transformers. In *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*. Springer Nature Switzerland Cham, 260–275.

[84] Donald A. Schön. 1983. *The Reflective Practitioner: How Professionals Think in Action.* Basic Books Inc., London.

[85] Ben Shneiderman. 2022. *Human-Centered AI.* Oxford University Press.

[86] Ben Shneiderman, Gerhard Fischer, Mary Czerwinski, Mitch Resnick, Brad Myers, Linda Candy, Ernest Edmonds, Mike Eisenberg, Elisa Giaccardi, Thomas T. Hewett, Pamela Jennings, Bill Kules, Kumiyo Nakakoji, Jay Nunamaker, Randy Pausch, Ted Selker, Elisabeth Sylvan, and Michael Terry. 2006. Creativity Support Tools: Report From a U.S. National Science Foundation Sponsored Workshop. *International Journal of Human-Computer Interaction* 20, 2 (2006), 61–77. https://doi.org/10.1207/s15327590ijhc2002_1

[87] Katta Spiel. 2021. "Why are they all obsessed with Gender?" — (Non)binary Navigations through Technological Infrastructures. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (Virtual Event, USA) *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 478–494. https://doi.org/10.1145/3461778.3462033

[88] Angie Spoto and Natalia Oleynik. 2018. Library of Mixed-Initiative Creative Interfaces. http://mici.codingconduct.cc/aboutmicis/

[89] Sarah Sterman, Molly Jane Nicholas, Janaki Vivrekar, Jessica R Mindel, and Eric Paulos. 2023. Kaleidoscope: A Reflective Documentation Tool for a User Interface Design Course. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 702, 19 pages. https://doi.org/10.1145/3544548.3581255

[90] Anna Ståhl, Vasiliki Tsaknaki, and Madeline Balaam. 2021. Validity and Rigour in Soma Design-Sketching with the Soma. *ACM Transactions on Computer-Human Interaction* 28, 6, Article 38 (Dec 2021), 36 pages. https://doi.org/10.1145/3470132

[91] Miriam Sturdee, Makayla Lewis, Angelika Strohmayer, Katta Spiel, Nantia Koulidou, Sarah Fdili Alaoui, and Josh Urban Davis. 2021. A Plurality of Practices: Artistic Narratives in HCI Research. In *Proceedings of the 13th Conference on Creativity and Cognition* (Virtual Event, Italy) *(C&C '21)*. Association for Computing Machinery, New York, NY, USA, Article 35, 14 pages. https://doi.org/10.1145/3450741.3466771

[92] Bob Sturm. 2022. Generative AI Helps One Express Things for Which They May Not Have Expressions (Yet). In *Workshop on Generative AI and HCI at the CHI Conference on Human Factors in Computing Systems 2022*. https://kth.diva-portal.org/smash/get/diva2:1757906/FULLTEXT01.pdf

[93] Elizabeth Wilson, György Fazekas, and Geraint Wiggins. 2023. On the Integration of Machine Agents into Live Coding. *Organised Sound* (2023), 1–10. https://doi.org/10.1017/S1355771823000420

[94] Elizabeth Wilson, Shawn Lawson, Alex McLean, and Jeremy Stewart. 2021. Autonomous Creation of Musical Pattern from Types and Models in Live Coding. In *xCoAx 2021 9th Conference on Computation, Communication, Aesthetics & X*. 76–93. https://qmro.qmul.ac.uk/xmlui/handle/123456789/73475

[95] Kyle Worrall and Tom Collins. 2023. Considerations and Concerns of Professional Game Composers Regarding Artificially Intelligent Music Technology. *IEEE Transactions on Games* (2023), 1–13. https://doi.org/10.1109/TG.2023.3319085

[96] K. Worrall, Z. Yin, and T. Collins. 2022. Comparative Evaluations in the Wild: Systems for the Expressive Rendering of Music. *IEEE Transactions on Artificial Intelligence* (2022). (Under review).

[97] Paulina Yurman and Anuradha Venugopal Reddy. 2022. Drawing Conversations Mediated by AI. In *Proceedings of the 14th Conference on Creativity and Cognition* (Venice, Italy) *(C&C '22)*. Association for Computing Machinery, New York, NY, USA, 56–70. https://doi.org/10.1145/3527927.3531448

APPENDIX

Appendix material can be found at:

- https://github.com/thecoreyford/Reflection-Across-AI-Music

Compositions can be heard at:

- https://codetta.codes/reflection-across-AI-music/

More information on RiCE [44] can be found at:

- https://ricequestionnaire.github.io/