Timbre Tools: Ethnographic Perspectives on Timbre and Sonic Cultures in Hackathon Designs

Charalampos Saitis*
Centre for Digital Music
Queen Mary University of London
London, United Kingdom
c.saitis@qmul.ac.uk

Bleiz Macsen Del Sette
Centre for Digital Music
Queen Mary University of London
London, United Kingdom
b.delsette@qmul.ac.uk

Jordie Shier
Centre for Digital Music
Queen Mary University of London
London, United Kingdom
j.m.shier@qmul.ac.uk

Haokun Tian
Centre for Digital Music
Queen Mary University of London
London, United Kingdom
haokun.tian@qmul.ac.uk

Shuoyang Zheng
Centre for Digital Music
Queen Mary University of London
London, United Kingdom
shuoyang.zheng@qmul.ac.uk

Sophie Skach
Centre for Advanced Robotics
Queen Mary University of London
London, United Kingdom
s.skach@qmul.ac.uk

Courtney N. Reed*
Institute for Digital Technologies
Loughborough University London
London, United Kingdom
c.n.reed@lboro.ac.uk

Corey Ford†
Creative Computing Institute
University of the Arts London
London, United Kingdom
c.ford@arts.ac.uk

ABSTRACT

Timbre is a nuanced yet abstractly defined concept. Its inherently subjective qualities make it challenging to design and work with. In this paper, we propose to explore the conceptualisation and negotiation of timbre within the creative practice of timbre tool makers. To this end, we hosted a hackathon event and performed an ethnographic study to explore how participants engaged with the notion of timbre and how their conception of timbre was shaped through social interactions and technological encounters. We present individual descriptions of each team's design process and reflect on our data to identify commonalities in the ways that timbre is understood and informed by sound technologies and their surrounding sonic cultures, e.g., by relating concepts of timbre to metaphors. We further current understanding by offering novel interdisciplinary and multimodal insights into understandings of timbre.

CCS CONCEPTS

• **Human-centered computing → HCI theory, concepts and models**; *Empirical studies in HCI*; • **Applied computing → Sound and music computing**; *Ethnography*.

*Corresponding author.
†The author was employed at Queen Mary University of London during this work.

AM '24, September 18–20, 2024, Milan, Italy

ACM ISBN 979-8-4007-0968-5/24/09
https://doi.org/10.1145/3678299.3678322

KEYWORDS

Timbre, sonic cultures, hackathon, musical interaction, collaborative design, design thinking, ethnography

ACM Reference Format:
Charalampos Saitis, Bleiz Macsen Del Sette, Jordie Shier, Haokun Tian, Shuoyang Zheng, Sophie Skach, Courtney N. Reed, and Corey Ford. 2024. Timbre Tools: Ethnographic Perspectives on Timbre and Sonic Cultures in Hackathon Designs. In *Audio Mostly 2024 - Explorations in Sonic Cultures (AM '24), September 18–20, 2024, Milan, Italy.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3678299.3678322

1 INTRODUCTION

Timbre is an important sonic concept that is difficult to work with. It has been understood as a 'wastebasket' attribute: a catch-all category for everything not pitch, loudness, or timing related in sound and music perception [72]. Yet, an abundance of musical and sonic expression lies within that wastebasket—timbre is many things to many people and has many functions [16, 18, 23, 82]. If timbre is a wastebasket category, it has the power to coalesce and contrast the most diverse types of sonic elements ('waste') in perpetually mutating listening practices [20, 75]. In this wastebasket, we keep discovering elements in sounds that we did not initially notice or anticipate [28, 80]. To work with timbre therefore requires to constantly *negotiate* and redefine it within our experience and relevant sonic practices and cultures.

Timbre has mostly been looked at in technical ways. Two prevailing approaches to its study have sought to construct spatial representations—"timbre spaces"—where a limited set of different sounds occupy different regions of a low-dimensional (typically three-dimensional) space based on how timbrally close or far they are from one another. These spatial configurations are constructed from multidimensional scaling of dissimilarity ratings for pairs of sounds [47], or factor analysis of ratings of individual sounds along semantic differential scales (e.g., "bright–dull," "soft–rough" [84]). While these paradigms are effective in exploring the perceptual representation of timbre [46, 66], they are not designed to observe its dynamic emergence and constant negotiation in human-human and human-technology encounters.

The "inexhaustibility of sounds" [80, p. 25] requires digging into the wastebasket of timbre to extract new ways in which we might reconceive what timbre is and what we can do with it. One approach would be to examine how it is used and understood in actual creative practice, for example, when designing tools and instruments for crafting musical and sonic expression. This approach turns attention to the site that lies between the creating subject (e.g., instrument maker, sound designer, live coder) and the receiving subject (e.g., performer, listener): technology. It raises interesting but unexplored questions about the role and practice of timbre (*timbre thinking*) in the development and adoption of sound technologies (*design thinking*) and their surrounding sonic cultures. Contemplating technologically and sonic-culturally situated facets of timbre can expand and diversify—and often confront—our understanding of how timbre is perceived, represented, and generated.

This paper presents a detailed account of the *Timbre Tools Hackathon*: a hackathon that invited designers, technologists, developers, researchers, and other makers to work with timbre through the design of tools for supporting timbre exploration in the craft practice of digital instrument makers—*timbre tools*. Here both "instrument" and "maker" are broadly construed, including composers and performers who build bespoke instruments as part of their artistic practice as well as live coders [34]. Pre-hackathon engagement involved a series of workshops on relevant hardware and software platforms to support learning, stimulate ideas, and facilitate team formation. Participants were also presented with open-ended and creative prompts to guide their projects. The main hackathon event followed a week later and lasted two days. All activities were available to attend in person or remotely via the Discord communication platform.

We performed an ethnographic study to observe how practitioners from diverse sonic backgrounds engaged with the topic of timbre in the design process, specifically how their practice and idea of timbre was dynamically (re)shaped through social-collaborative interactions (talking about timbre) and technological encounters (working with timbre). Hackathons are time-bounded, low-pressure collaborative events that present themselves as *observatories of design thinking* [24, 27]. Accordingly, to explore the use and understanding of timbre in the design of tools for instrument makers, we draw on the phases of the Double Diamond design process model[1] and the related notion of *problem space* and *solution space* [42] to develop our research questions. Exploration of problem space and solution space follow a pattern of *divergent thinking* (exploring choices) followed by *convergent thinking* (making choices), visualised as two diamonds as seen in Figure 1. Specifically, our research questions (RQs) are:

- *Exploring the problem space* - **RQ1:** How do participants think about the concept of timbre in the design of tools for making instruments?; **RQ2:** What (collaborative) strategies do they use to conceptualise their design?

[1]https://www.designcouncil.org.uk/our-resources/the-double-diamond/

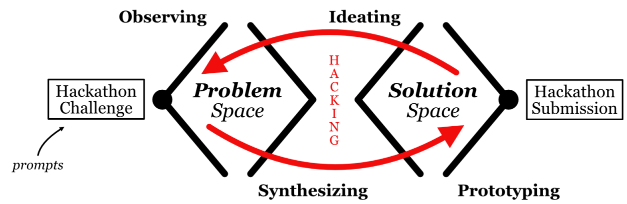

Figure 1: Hackathons as observatories of design thinking: Exploring the problem space and solution space follows a pattern of divergent thinking followed by convergent thinking, visualised as two diamonds (adapted from [27, 42]).

- *Exploring the solution space* - **RQ3:** How do participants use the tools currently available to them to develop their concepts? How do participants' choices of tools relate to how they think about timbre?

In summary, we offer the following contributions, drawn from ethnographic analysis of field notes and transcripts of both structured and unstructured interviews, as well as our own reflection on how we structured the hackathon activities:

- Descriptive accounts of the timbre tool designs and processes for 11 hackathon teams,
- Insights into how subjective experiences of timbre were negotiated by the hackathon teams through metaphor,
- Identification of these designs as embodiments of each teams' collective understanding and definition of timbre within their sonic culture(s), and
- Broader implications for design of similar subjective material for music research communities like Audio Mostly and generally for human-computer interaction (HCI) research.

Additionally, code and videos for all hackathon project submissions are available online at the hackathon's webpage.[2]

2 RELATED WORK

2.1 Timbre, Technology, and Sonic Cultures

Electronics, analog and digital, provide novel means to generate and shape sound. Their development has been paralleled by the development of sonic cultures—cultures of listening [20, 37], cultures of creative production and consumption [74, 79], and cultures of designing and making [59, 60], amongst many others. One of the first electronic instruments was the Theremin, which worked by means of electromagnetic fields (control) and heterodyning oscillators (sound). The sonic purity and eeriness of the latter has since associated electronic sound with "ethereal timbres" [15]. The Theremin was the accidental product of its inventor's research into proximity sensors during WWI [33]. Because they are both technologies and instruments, synthesizers have had a transformative effect on the sound of many musics [4] and also on sonic cultures of instrument making [60]: The distinctly "'80s sound" is tied to the electric piano preset of the Yamaha DX7 FM (frequency modulation) digital synthesizer [41]. Auto-Tune began as a pitch-correction plug-in for fine-tuning live-recorded vocals, but has since been used

[2]https://comma.eecs.qmul.ac.uk/timbre-tools-hackathon/

widely as a general and pronounced vocal effect that produces a signature timbre and sonic culture in popular music [6, 62].

Of course, pianos and violins are also technologies. Contemplating the timbral agency of piano pedaling in African-American jazz aesthetics, Dana Gooley articulates that instrumental timbres "are the products of an encounter between a person and a technology" [28, p. 121] (see also [64, 65]). The addition of new instrument-technologies to the orchestra in the eighteenth century foregrounded the formalisation of orchestration and a new form of attention to timbre [14].

Technology was also central in the formation of timbre perception as a research field. Computational methods for multidimensional scaling developed in the 1960s enabled researchers to consolidate responses from listening tests into spatial representations which they could also visualise [30]. Advances in digital signal processing and sound synthesis around the same time (e.g., FM) made it possible to link spectral, temporal, and energy features of recorded or synthesized audio signals with listeners' perceived timbre dissimilarity. Since then, the "timbre space" model and metaphor has dominated scientific discourse on timbre [73]. It has also been built into new technologies for music creation and performance [22, 70] as well as into audio classification schemes and formatting standards that support all kinds of digital music applications [52].

We may contrast the idea of timbre inscribed in monolithic timbre spaces with Dolan et al.'s argument that timbre emerges in a dynamic relay between technology, creation, and history [14, 15]. Ethnomusicological contributions suggest similarly [28, 29]. Could timbre exploration be understood as the primary guiding factor in the development of sound technologies? Reducing the understanding of the former to the operative functions of the latter (technological determinism), and conversely, viewing the history of music technology as a history of timbre exploration (technological constructivism) might potentially oversimplify the intricate and complex dynamics between sonic ideas and materiality, between history and the technological present [44]. Rather, perhaps timbre exploration can be understood as the (sometimes accidental) result of harnessing technical processes and materials for sonic purposes.

2.2 Hackathons, Ethnography, and Design

Hackathons have been used as a fast-paced design process since the early 1990s, with origins tied to Silicon Valley [43]. In their review of hackathons across HCI, Falk Olesen and Halskov [21] found that hackathon settings helped structure design thinking and created opportunities for designers and developers to collaborate on making activities. During hackathons, teams undergo a process with parallels to design processes [27]. For example, teams might start with "fuzzy" initial ideas which become more refined over time [68]. Teams also tend to go back and forth between phases of divergent thinking (generating several ideas) followed by convergent thinking (narrowing down ideas) to move from an initial problem space to a solution space (e.g., the double diamond model, see Footnote 1). In this way, hackathons serve as settings to observe how interdisciplinary teams conceptualise and create their designs [27], giving insight into teams' technical choices and negotiations around subjective concepts.

The subjective concepts to be negotiated within the design process of a hackathon project are often captured by its theme [54], which could be related to a diverse range sonic cultures. Indeed, hackathons have precedence outside of the walls of academia [17] and can serve as a method to engage a wider range of sonic cultures. Across HCI literature, there are several examples of sonic cultures which have been explored in a hackathon format. Zappi and McPherson [85] present how people appropriated instruments with limited affordances into their practice, unpacking their design process of digital musical instruments (DMI) where sounds could not be directly anticipated by the designers. Correia and Tanaka [11] followed a hackathon approach to explore a sonic culture of audiovisual performance tools, while Huang et al. [32] explored how developers and musicians used AI as a design material in music making.

Research into hackathons typically tends to draw upon observations made by non-participants, e.g., [27, 31, 36, 56, 58]. This parallels ethnographic approaches as used in HCI, capturing information on how activity unfolds "in-the-wild" [10]. Ethnographic approaches also give researchers the opportunity to examine collaborative aspects of how hackathon teams operate; for example, observing spatial formations which affect our social hierarchies, or how much or little participants engage in conversations [38]. We note that with recent shifts towards more digital and hybrid settings in hackathons [69] and musical collaboration [13], the rules of spatial formations change and social structures are often redesigned. In this paper, we draw particular influence from Millen's "rapid ethnography" [51], which suggests time-deepening strategies to ensure rich data collection within a limited time period: typical of hackathons. These strategies include focusing on key informants and having researchers follow up on key insights from their observations by interviewing participants in-the-moment. We take our hackathon setting as a window into exploring the how timbre is conceptualised in the sonic culture of timbre tool makers, following an ethnographically inspired method:

3 METHOD

The Timbre Tools Hackathon (hereby referred to as simply "the hackathon") was held in collaboration with The Audio Programmer (TAP),[3] an audio development company with a strong focus on fostering community and providing free educational resources. The hackathon ran for 48 hours and was accompanied by four workshops held in a one-day event one week before the hacking commenced. The hackathon was a hybrid event, held in person at Queen Mary University of London (QMUL) and supported online through the TAP Discord server. All workshops and presentations were streamed through the TAP YouTube channel.

Figure 2 illustrates our methodological approach and hypothesis. We expect timbre tools to be comprised through the interplay between timbre thinking and design thinking. Sonic cultures and experience with musical and creative practices are a resource the participants bring into the hackathon (top right of Figure 2, in pink). These influence timbre thinking (right circle of Venn diagram), which might also be inscribed in familiar or preferred sound technologies. In the same way, resources related to the structuring

[3] https://www.theaudioprogrammer.com/

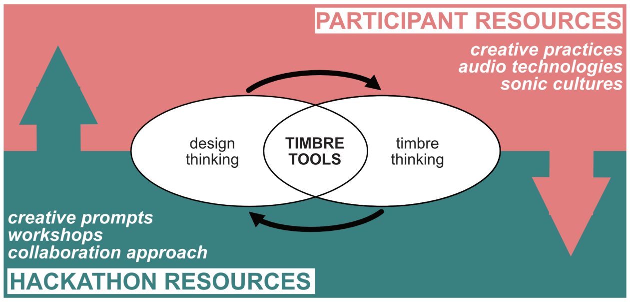

Figure 2: Methodological approach: Participant (top right, pink) and Hackathon resources (bottom left, teal) influence the iterative negotiation between timbre thinking and design thinking in the design of timbre tools (Venn diagram, centre).

of the hackathon (bottom left of Figure 2, in teal) provide a context that also influences the timbre tools designed. Other participants in collaboration, as well as the prompts and workshops offered, will also be entangled with the way timbre is negotiated at the team level. This will inform specific design thinking (right circle of Venn) choices to reconcile definitions of timbre and which of its aspects are incorporated into the timbre tools through the hackathon.

3.1 Participants

The hackathon aimed to bring together practitioners and researchers in the fields of music technology, audio programming, deep learning, interaction and design engineering. Our call for participation was directed at people from this broad range of disciplines, welcoming participants from all stages of practice and research. We shared the call using social media platforms and academic mailing lists. The call for participation was also communicated through TAP community channels, which allowed us to reach an audience of about 10,000 members (at the time of the call) interested in music technology, beyond our contacts in academia.

Of an initial 19 teams registered, 11 actively engaged during the hackathon activities (2 remote, 2 hybrid, 7 on site), totalling 30 participants. Participants were between 22 to 55 years old (average age 30 yrs). Of these, 5 participants identified as female, 1 as non-binary, and 22 as male. 15 participants were white, 6 Asian, 2 black, and 3 from mixed/other backgrounds. The majority of participants taking part in the hackathon resided in the UK at the time of the study (19), with the rest living in EU countries and one in India. In contrast, only 10 participants spent their formative years in the UK, while the rest attended to their instruction in China (3), India (3), Canada/USA (3), Uruguay (1), and EU countries (10). Finally, participants reported a moderate familiarity with the concept of timbre, scoring a mean average of 3.8/5 points on a 5 point Likert scale for the question "How familiar are you with the concept of timbre?". Six participants reported themselves as experts (5/5) and only two expressed a slight level of experience (2/5). A comprehensive breakdown of team compositions can be found in Appendix A.

The hackathon and ethnography study were approved by the QMUL Ethics of Research Committee (ref. number DSEECS23.137). Upon signing up, each participant was provided with an information and consent form describing the event and online platform. Participants also agreed to familiarise themselves with the Berlin Code of Conduct[4] and follow this conduct in all communications, in-person and online.

3.2 Pre-Hackathon Engagement

Before the main 48 hour hackathon event, we conducted a number of engagement activities with participants, which we describe below.

3.2.1 Warm-up Workshops. One week before the hacking event, a one-day event with four workshops was held in person at QMUL and live streamed on the TAP YouTube channel, with the aim of stimulating our participants' creativity. This aimed to provide opportunities for participants to deepen their knowledge or introduce themselves to some relevant resources for the hacking.

Four workshops were held on the following: (1) *SP Tools*,[5] a set of machine learning tools for low latency and real-time performance in Max/MSP; (2) *Bela*,[6] an open-source maker platform optimised for musical and audio performances [48]; (3) *Signal Flow*,[7] a Python sound synthesis framework for live coding and performance; and (4) *Neutone*,[8] a company dedicated to developing AI-powered plugins for musicians and artists. The goal of these workshops was two-fold: First, they served as an educational resource to introduce relevant technical concepts and platforms related to DMI development, audio signal analysis and synthesis, and machine learning. Second, the workshops encouraged potential participants to connect with each other and start forming collaborations. Each workshop lasted between 60 and 90 minutes. Workshop presenters for Signal Flow and SP Tools made themselves further available during the hacking period to support teams with their projects.

3.2.2 Team Formation. After an initial expression of interest collected in the previous three months leading to the hackathon, participants were encouraged to form teams of one to five members for their hacking project. Team registration opened a week before the event, at the end of the workshops. In addition to registering with us through our website, participants signed up to the TAP Discord server. Participants were invited to post hack project ideas onto Discord and to form collaborations over the subsequent week leading to the hackathon weekend.

3.2.3 Creative Prompts. To facilitate team formation and project ideas, we presented six open-ended and creative prompts. These emerged from semi-structured interviews with practising instrument makers and musicians, including live coding artists, conducted as part of our wider investigations into timbre and musical craft practice (but not reported in this paper). Interviewees were asked, among other questions: "*What broad ways could further support and enhance your timbre-based practice? What broad ways could help you to engage timbre in a more meaningful way?*" Using reflexive thematic analysis [5], we generated themes around using tools for

[4]https://berlincodeofconduct.org/
[5]https://rodrigoconstanzo.com/sp-tools/
[6]https://bela.io/
[7]https://signalflow.dev/
[8]https://neutone.ai/

the exploration and navigation of timbre space, for body and material based timbral mappings, and to analyse or predict timbral variations when altering spatial locations, recording hardware, and musicians' expressivity. In addition, participants also suggested a need for tools that help with timbre listening, tools to use timbre to improve the accessibility of music to people with disabilities, and tools to understand the cultural value of timbre.

Based on our reflections, these ideas were abstracted and extended to the following six prompts:

(1) **deep timbre** - *deep learning, deep listening, learning to listen*
(2) **to infinity and beyond** - *explore, navigate, generate, timbre space*
(3) **sensing timbre** - *sensors, materials, circuits, acoustics, embodiment*
(4) **timbrecheck** - *live performance, sound recording, timbre in space*
(5) **back to the future** - *analysis-by-synthesis, reflection, sound reproduction*
(6) **timbre for all** - *community, social interaction, accessibility*

3.3 Hacking and Submissions

3.3.1 Hacking Schedule and Spaces. The hackathon took place from 4 PM Friday 23 February to 9 PM Sunday 25 February, 2024. The first hour on Friday was dedicated to finalise teams and complete registrations. Hacking officially started Friday at 5 PM and finished Sunday at 5 PM. Teams were then granted an additional hour until 6 PM to finalise their submissions, record the video presentation, and prepare their repository. From 6 PM, the teams presented their submissions, which were live-streamed through the TAP YouTube channel. Winners were announced at 7:30 PM; afterwards the teams shared a final social moment with a few snacks and refreshments.

Each team was given both a private virtual space on the TAP Discord server and a shared physical space at QMUL, accessible from 10 AM to 10 PM each day, to hack and interact with each other. Participants could use both spaces at their discretion, being free to hack on site, online, or elsewhere, as they saw fit. We also encouraged everyone involved in the event to take breaks, have a regular sleep schedule, and generally act responsibly towards themselves and their team throughout the hacking. Light refreshments and snacks were provided on site. A group walk was also organised through a local park on the second day to encourage wellbeing and socialising.

3.3.2 Resources and Facilities. A living document of resources to work with and design timbre[9] was shared with all participants to support and inspire their hacking process. Participants could also freely access a dedicated Discord channel for each of the warm-up workshops (see Section 3.2.1), alongside the other resources present on the TAP Discord server.

In person teams could additionally ask for a kit comprising a Bela board, a breadboard, a Bela Trill touch sensor, a piezo contact mic, and a selection of jumper wires, alongside a limited stock of Arduino boards and additional sensors. We also provided access to a maker space with laser cutters, saws, a pillar drill, and soldering equipment. However, none of the in-person teams made use of these facilities.

3.3.3 Submission Criteria and Prizes. To be included in voting, each team was asked to submit: (1) A short two minute presentation of their project and ideas, and (2) A public repository of source code (or instructions on how to recreate your project) under an Apache-2.0 or compatible licence. Open sharing of ideas was a primary objective of the hackathon and permissive licenses help foster further work in the topic of timbre tools. Interactive demos, websites, and audio plug-ins were also welcomed but not required to be considered for the final selection.

The submissions were judged by a panel of experts (see Acknowledgments) who awarded prizes to all members of the first-place teams in four main categories: (1) *Best Idea/Research Direction*; (2) *Best Presentation/Demo*; (3) *Best Creative Hack* (oriented toward offbeat and artistic thinking); and (4) *Best Responsible Hack* (oriented toward responsible innovation). The provision of these prizes was to celebrate the most innovative ideas and teamwork, and to add a competitive element to help stimulate individual and team investment.

3.4 Data and Analysis

Below we detail the approach taken for data collection and analysis during the hackathon event.

3.4.1 Ethnography Team. A team of six ethnographers, including five of the authors, carried out on-site observations. The ethnography team collectively coordinated their efforts through an initial meeting at the start of the hackathon and subsequent daily updates. Each ethnographer contributed according to their schedule and were elected to observe teams on each day. Materials were collected using a shared Miro board[10], allowing each ethnographer to follow on from data collected by other ethnographers across their varied schedules. The ethnographers' timbre-related backgrounds and positionality are included in Appendix B.

3.4.2 Data Collection. All teams (on site, remote, hybrid) were asked to post self-reports via their private voice channels on Discord. Self-reports were meant to be bite-sized, taking no more than a few minutes to write, and possibly including media such as photo or screenshot, sketches, code snippets, or audio samples. Teams were asked to self-report their progress only when actively working, at an hourly rate, or otherwise as they saw fit, to avoid interrupting their work processes.

Inspired by Millen [51], the ethnography team employed both interviews (unstructured and semi-structured) and ethnographic observation in their data collection process. *Observations* were conducted mainly for on-site teams and were captured by researchers in the form of free-form field notes and drawings as a form of transcription [1]. *Unstructured interviews* were conducted impromptu to follow up on ethnographers' observations whilst being mindful to not interrupt hackathon participants' flow. This allowed the ethnographers to focus on "key informants" and take a more interactive approach to capturing rich information on key ideas within the hackathon's limited time period [51]. *Semi-structured interviews*

[9] https://github.com/comma-lab/timbre-resources

[10] https://miro.com/

were conducted up until one day after the end of the hackathon to investigate the teams' interactions with the hackathon theme, timbre, and technical tools (see Appendix C).

3.4.3 Analysis. After the hackathon, our range of collected data was organised within one shared document. Data was organised by team to be easily understandable for all ethnographers for a collaborative analysis [51] and to retain the individualistic qualities driving each team's design processes. Specifically, the document collated: all discord logs from teams and media, all Miro notes, sketches, observation notes and interview transcripts. This meant that the ethnographic notes, structured interviews, and unstructured interviews were consolidated together.

An iterative narrative analysis [76] approach was then taken to further reduce our broad data collection into descriptions of each team's design process "to help communicate the field learning to the rest" of the ethnographers [51]. This was inspired by related work on hackathons [31] and creativity-related HCI [25, 45, 77, 78] research showcasing a plurality of understandings across participants with diverse backgrounds and perspectives, similar to our hackathon teams. First, one author identified "anacdotes and repeated patterns"[61, p. 312] in the shared document to identify ways that: participants conceptualised timbre (RQ1), collaborated on their design (RQ2) and used tools to negotiate timbre (RQ3). They then brought these patterns together to create first-draft descriptions of each team's individual design process, presented in Section 4. Each narrative was accompanied by an overview of teams' backgrounds and a description of their hackathon submission to retain context on teams' sonic cultures.

Following the first-draft, all ethnographers met across several weekly meetings. In the meetings, all ethnographers reflected on both the collaborative document and descriptive accounts to both corroborate and enrich each narrative description with their own interpretations and perspectives, as informed by their unique backgrounds. For example, one ethnographer with particular expertise in human-human interaction brought forth additional details on each team's use of collaborative space where applicable (cf. RQ2). We emphasise that narrative analysis has no prescriptive method but is closer to a craft practice, where narratives are informed by the author's critical readings of their collected data [26].

The similarities and differences across these narratives were further collectively reviewed and analysed by all authors through weekly meetings, considering our research questions outlined in Section 1. We identified several similarities and differences in how (1) participants created mappings, (2) listened to audio in their design process, (3) negotiated timbre with their choice of technology, and (4) manipulated generated audio. We describe these reflections in the discussion.

4 FINDINGS

In this section, we report on the design process, collaborative approach, and the resulting timbre tools submitted by each team (10/11 teams submitted). Some examples of the submission can be seen in Figure 5. Where applicable, we give details on a team's use of collaborative space; it was not possible to observe fully remote and one-person teams in terms of spatial formations and nonverbal behaviour. Some examples of spatial formations and constellations mentioned in this section are sketched out in Figure 3 and Figure 4, which will also be reflected and discussed more in-depth in 5.2. Table 1 summarises the background of each team, their submissions, the tools and platforms they used, and the ways they negotiated timbre. We give information on each team individually to showcase the diversity of tools and techniques we observed. Direct quotes are taken from interviews with teams.

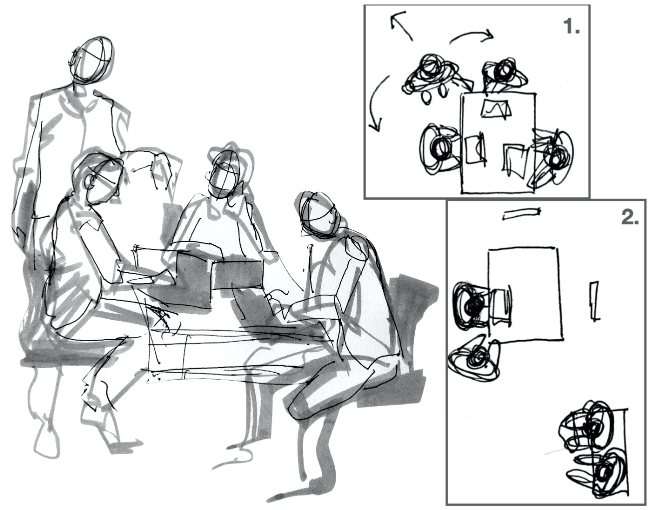

Figure 3: Drawings of hackathon participants working in teams in different spatial constellations.

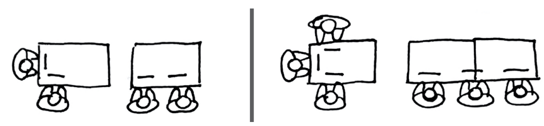

Figure 4: Examples of different spatial formations in L-shape, and lined up next to each other of teams with two (left) and three (right) members.

Team 1. The team had one developer, two researchers specialising in audio, AI, and music creativity, and a digital luthier who designs interactive musical systems. One member has a background in live performance with audio software, and one in jazz guitar and electronic/electroacoustic music improvisation.

Submission: A Max MSP[11] patch that visualises timbre features of an audio input in real-time using a continuous plotted graph. The timbre features were extracted with a combination of semantic descriptors and Mel Frequency Cepstral Coefficients (MFCC) [7]. Selected features can be mapped into a Realtime Audio Variational autoEncoder (RAVE) [8] audio synthesis model's latent space to guide the synthesis process in real-time.

Design Process: Initially, the team brainstormed around why timbre is absent from the DMI makers' toolbox. They converged onto three main issues: (1) timbre is difficult to visualise, (2) timbre descriptors are inaccessible, and (3) timbre is easy to generate but difficult to guide. Discussing these issues, they further converged to

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

Table 1: A summary of hackathon teams' backgrounds, the tools they used in the hackathon, and their ways of negotiating timbre in the design process.

	Size	Background	Tools they used	How they negotiated timbre in the design process
T1	4	Audio developer (1) Researcher (2) Digital music instrument maker (1)	Max MSP	visualisation of audio metrics such as spectral centroid, spread, skewness, and MFCC; streaming audio snippets;
T2	5	Music and audio researcher (3) Audio developer (1) Audio engineer (1)	Tölvera Python Library, SignalFlow, PureData	"organic"; latent space parameters
T3	1	Digital music instrument maker	Arduino, Teensy Audio Library	audio playbacks; "random" and "predictable" sounds; "capability of an instrument and what one can make out of it"
T4	5	Sound and music computing student (4) Lecturer (1)	Python scripts for analysis, Max/MSP, JUCE for development	metaphors such as "bubbly", "trippy", "crunchiness"; "predictable aspect" and "complex aspect" of features
T5	4	Music and AI researcher (4)	Python notebooks for analysis and prototype, Flask for development	text prompts; parameters in a joint embedding between text and timbre
T6	2	Audio researcher (1) Sound designer (1)	PureData	audio impulse responses of the algorithm
T7	3	Music and AI researcher (3)	Python notebook	audio metrics including spectral centroid, spectral flux, spectral flatness, CQT, and a few MFCC
T8	2	Music researcher (1) Software engineer (1)	Javascript, GPT3.5	text prompts; audio playbacks
T9	2	Music and AI researcher (2)	Python notebooks	audio metrics including spatial flatness, inharmonicity, danceability scores; "smooth transition", "abrupt transitions"
T10	1	Lecturer in digital music	PureData	latent space parameters; "non-nature sound" and "nature sound"
T11	1	Game developer	JavaScript, OpenFrameworks	sample parameters such as pitch, volume, playback speed

their aim: to create a tool that uses specific timbral characteristics of audio signals to guide the parameters of an audio generation tool. Next, they explored different audio feature extraction approaches such as MFCC, spectral centroid, and skewness. They were fascinated by the large number of timbral features that can be extracted, noting that "re-mapping [extracted timbral features] to other synth parameters can lead to interesting results". To focus their thinking, they considered which features are meaningful as synth parameters and drafted an initial patch.

Collaborative Space: Regarding the created interactional spaces of this team, working in a hybrid constellation determined a screen focused approach also for in-person and face-to-face members. The screen dominantly served as a shared interface and when communicating through it, explanatory gestures were used [13], although pen and paper was also used for some idea sharing between the in-person members. Throughout the hackathon, the used work space as well as their postural movement was restricted to a table. However, over time, the in-person members exhibited an increasingly relaxed and "open" [49] nonverbal behaviour.

***Team 2**.* Three members of the team described themselves as researchers in Music Information Retrieval (MIR) and Interaction Design, whilst the other two members are a developer and an engineer working in the field of audio analysis. Two members have music performance/production backgrounds; one member has a degree in classic piano.

Submission: A sound synthesis program that maps data from an artificial life simulation to the parameters of a RAVE model for audiovisual performance.

Design Process: The team was interested in "organic" timbral qualities and aimed to "incorporate timbre into an artificial life simulation in such a way that timbre serves as a tool to explore the simulation and vice versa". They started by brainstorming how to create parallels between artificial life simulation data and sound, and operated fully online through individual independent lines of work, keeping timbre-related thinking at a meta-level. The team decided to use software libraries which were new to most members, including the Tölvera library [2], so a few design decisions are made based on technical affordances. This led to later design decisions being technocentric, for example, they controlled a latent space vector using the absolute positions of particles instead of a more complex measure. Lots of effort was made to integrate multiple existing tools: "the approach was to incorporate first all the tools we wanted

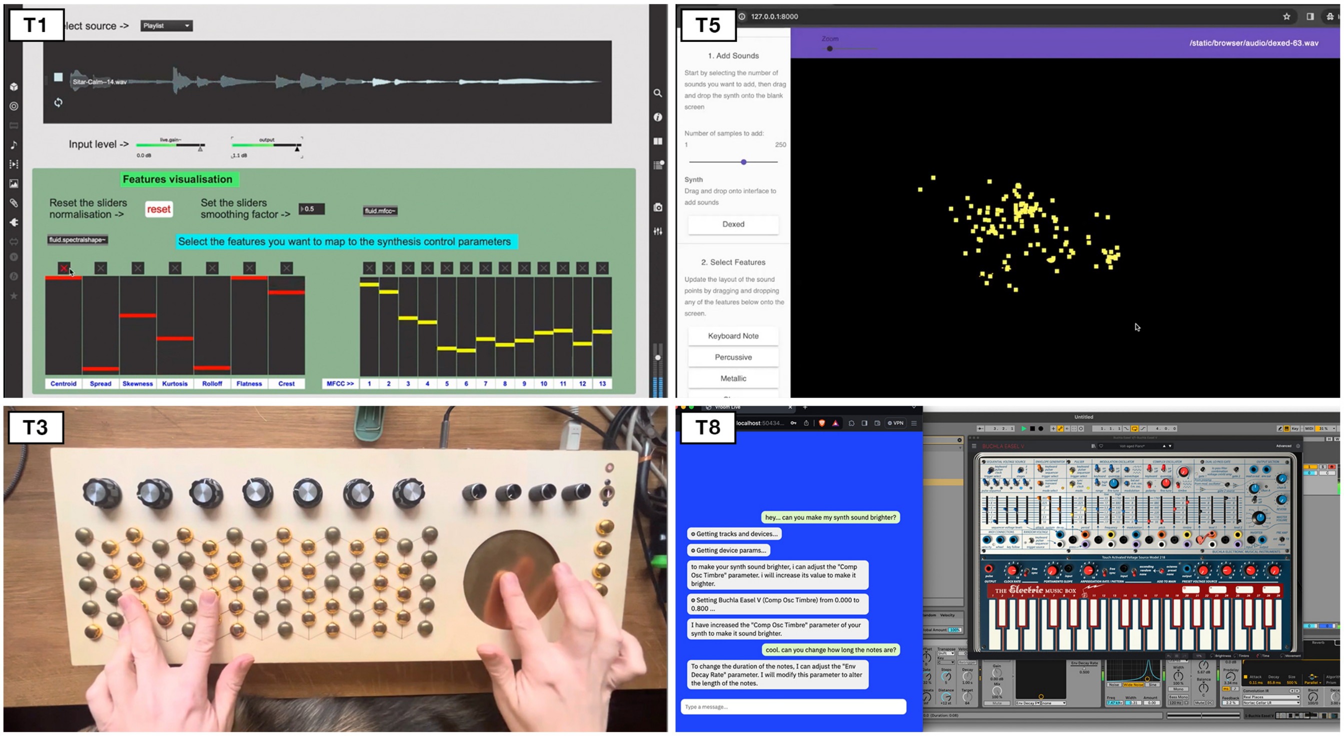

Figure 5: Screenshots of the Max MSP patcher submitted by Team 1, the web interface submitted by Team 5, a video demonstration of the prototyped instrument submitted by Team 3 using their audio library, and the web interface submitted by Team 8.

to have available[...] and only then consider the timbral possibilities they offer". The members see the outcome of the hackathon as both a chance to learn new technical skills and a new way of thinking about the dimensional aspects of timbre.

Team 3. A one-member team, self-described as a DMI maker and music technology enthusiast.

Submission: An audio library for microcontrollers that musically randomises synth patches and generates real-time expression macros to assign to sensors.

Design Process: The team started the hackathon with a clear submission idea motivated by their current musical practice. Their objective was to find an approach for sound design that avoids changing loads of parameters manually, drawing inspiration from early electronic instruments, such as the theremin, which balanced musical expression (having continuous control over parameters real-time) and complexity. They chose Teensy Audio as the development platform given their previous knowledge and preference for making physical instruments. During hacking, they actively thought about how to dial in the right amount of randomness to open up chances of creating deployable musical sounds without being too predictable. The team conceptualised timbre by considering the capabilities of instruments and what one can make out of it, connecting these to the textures an instrument can have: "I think of timbre kind of as oscillators and what you do to oscillators, not as snippets of audio".

Team 4. The team is composed by four students in sound and music computing and led by a researcher/lecturer who focuses on real-time machine learning in the same field.

Submission: A Max patch, which extracts features of an image and maps its output to musical attributes in a JUCE plugin. A slinky spring can be used in front of a webcam to interact with the plugin's sound controls.

Design Process: The team started from the assumption that objects in the physical world can create a more natural timbre interaction compared to clicking on a digital interface. This motivated them to explore different ways of modifying sonic characteristics through physical metaphors related to the body. Their skill set guided tool choices. The team started by integrating techniques in computer vision to capture the features of a physical object (a slinky spring), and used these features to control timbre or produce sound. They established several features of image and sound and attempted to categorise them into "predictable aspects", such as parametric features including the amount of distortion, MIDI, image brightness, toy presence, RGB values, and "complex aspects", such as features in the frequency domain like transforming 2-D images into 1-D frequencies using Fast Fourier Transform (FFT). After defining the set of features, the team drew upon the concept of cross-modal mapping to connect them to sound, such as mapping the gradient of the image (crunchiness) to distortion. They also came up with several metaphor-like languages to describe the characteristics of sounds, for instance the amount of red to "bubbly", the amount of green to "trippy".

Team 5. The team is comprised of researchers in the field of music and AI with a range of backgrounds in music composition, performance, and production.

Submission: A web application that uses a text prompt to identify a set of parameters that produce a timbre best matching the prompt. The group sees this process as "creating a guide to navigate the almost infinity parameter space of a synth" and "an exploration into the corresponding between language and timbre via a pre-trained model on paired audio and text descriptors".

Design Process: The team main objective was to create a tool that uses text prompts to access a point in a large parameter space. At the beginning of the project, they were open about the design space of the tools at hand and actively explored their possibilities. The team discussed relevant techniques and sketched out the overall structure of their software system. A wide range of techniques were considered including the Contrastive Language-Audio Pretraining (CLAP) embedding model [19], an evolutionary algorithm and a software frameworks such as Flask and Django. They assigned roles for each member, from front-end design to back-end algorithm design, and started working individually, following a typical software development-styled approach. One branch of the team looked at a large language model trained on paired audio-text and experimented with how it could be used to guide the parameter search. The other branch looked at integrating related issues, such as front-to-back-end communication and loading VST plug-ins into their development environments. Members converged near the end of the project to discuss integration and code-related issues. During this process, they also discussed which parameters in a synth are important and meaningful and can lead to timbral changes.

Collaborative Space: All team members were physically present in the work space provided. Over time, they formed different subgroups and spatial formations, exhibiting a wide range nonverbal social signals and orientations [38], like L shapes, circular formations, or lined up positions, seen in Figure 3. In addition to screens as shared interfaces, they made use of white boards as analog tools for idea sharing. Ethnographers observed that increased bodily movement correlated with a decreased dominance relating to a work task (such as a more administrative or assistive role), and vice versa. Further, the team divided into pairs occupying different physical spaces (sofas or desks) for working on specific tasks of their project. There seemed to be a very clear social structure and organisation assigning team roles and tasks in this group.

Team 6. A researcher in spatial audio and a sound designer – both working with emphasis on sonic interaction.

Submission: A mobile application that can record audio samples and apply them as convolution reverb on sound.

Design Process: The team was interested in convolution reverb and hoped to experiment with its ability to turn any recording into an audio effect. Therefore, when designing the interaction, they focused on its convenient and portable aspects. This motivated them to choose the MobMuPlat platform, which is a mobile music platform that can use phone sensors as input devices for PureData. During hacking, they ran a lot of experiments with the impulse responses of the algorithm. They focused particularly on the feature of discovering the timbral perspective of sounds in the real world, and applying them as a sound effect.

Collaborative Space: As a pair, they worked alongside each other throughout most of the duration of the hackathon, and, despite attending in person, preferred to use a work space off-site the offered hackspace, which limited the ethnographic observations.

Team 7. The team is made of researchers in the field of digital music and AI. One has a professional background in classical music performance.

Submission: A prototype that loops through a cluster of orchestral chord samples according to mouse position on a 2D plane.

Design Process: The team started initially from an existing orchestral music audio dataset, provided by one of the team members. They set out with an plan to cluster the audio and use the clusters to drive a granular synth. They also planned on using a 3D interface and interacting with the clusters through mouse movements. Most of their experiments were run in Python notebooks. The experiments initially started by analysing the relationship between chord and audio features such as spectral centroid, spectral flux, spectral flatness, and MFCC. A few figures were plotted to verify the features. After the analysis, they worked on clustering samples and experimented with ideas.

Collaborative Space: This team's shared space remained a lined up formation for most of the hackathon, as illustrated in Figure 4. Slight variations in body posture and orientation occured during verbal exchanges without mutual gaze (eye contact). Members seemingly taking a more dominant role within the group performed fewer movements and were more screen focused than others. Less dominant members showed a more active bodily behaviour, turning to others and looking off screen more regularly.

Team 8. One member of the team is a researcher in music and machine learning. The other is a founder/engineer/musician working in the broader digital music field.

Submission: A web application that chats with a GPT agent about a desired sound output. The AI agent can then change the parameters of a synth loaded in Ableton based on the discussion.

Design Process: The team's idea was to build an educational tool for Ableton that allows users to change a VST's parameters by describing the desired timbre to an AI agent. They already have experience working with text-to-audio models, and they wanted to extend the interaction into a "shared negotiation and explanations" approach. This choice is motivated by the will of using "expression in language [to get closer to the desired output of the sound] rather than moving knobs". The team started by connecting a GPT3.5 agent and Ableton using OSC, then took a UX-oriented approach to iterate over the exploration and prototype. A few design decisions were limited by the availability of technical tools. A core mapping-related issue was identified during their hacking as "the transformation from language to [synth] parameters relies on a black box model (GPT)", which can be affected by the user's ability to describe a timbre.

Collaborative Space: This team's two members created a balanced interaction switching between face to face conversations with eye contact, as well as screen focused while sitting next to each other. They changed their posture accordingly, turning torsos to each other and away from the screen, showcasing a more conversational behaviour than other teams where the screen played a dominant role as an interaction interface, including changing their spatial formation from a line to an L-shape. This also led to mirroring each other's postures, such as crossing legs or leaning forward [67].

Team 9. Two PhD researchers in the field of music informatics and machine learning/AI. One often performs as a DJ and the other has a background in music production.

Submission: A demonstration of how computed audio metrics, including metrics for danceability [50], change when a DJ transits from one track to another.

Design Process: The team started the hackathon with the initial aim to make a tool for DJs that recommends songs based on timbre, connecting with one of the member's experience as a DJ. After deciding on datasets and environments, they experimented with audio metrics such as MFCC, spatial flatness, inharmonicity, and danceability scores using Python notebooks. These features were used to visualise timbre similarities between each song in a DJ set. They also used these timbral features to create an algorithm that estimates DJ transition lengths based on timbral distances, and observed two types of transitions: "Smooth transitions tend to interpolate between timbre clusters; abrupt transitions tend to exhibit distinct timbre clusters."

Collaborative Space: As other teams working in a pair mostly interacting through a screen interface, their spatial formation mostly remained lined up next to each other with little eye contact, like in Figure 4 (left). However, for self organised check-ins, they turned away from their screen or shared one screen amongst them to discuss next steps and tasks for their submission. For these occasions, they also adapted their posture and turned towards each other.

Team 10. A one person team comprised of an academic researcher in MIR and deep learning.

Submission: A PureData patcher that controls four latent space axes using the spatial orientation of a phone.

Design Process: The team started by experimenting with a deep variational auto-encoder for audio synthesis. Decisions on system designs and technology choices were mostly based on intuition and availability of tools. For instance, they decided to explore the latent space using a phone application since it was a simple way to send sensor data via OSC. The team explored different ways to wire the phone and AI model together, and became interested in the morph between drone sounds and birds twittering and described the sonic characteristics as either a "non-nature sound" or "nature sound."

Team 11. A one-person team who describes themselves as a game developer and creative technologist.

Submission: No submission was made due to technical issues, but they still presented their idea of creating a gestural audio controller.

Design Process: The team's initial idea was to create an audio-visual piece using gestural control. Their design process prioritised visual over audio. They expressed doubts about their definition of timbre, so they hoped to start by manipulating audio with different parameters such as pitch, volume, playback speed, or layering different samples. The team chose to play with samples rather than synthesis algorithms because they felt it was easier. They ended up with a hand-tracking program as a starting point for a gestural controller but had not thought about the mapping between audio and gestures. As part of their submission, they presented also their further plans of connecting audio parameters with dance performance.

5 DISCUSSION

In this paper, we investigated how groups of music technology developers and musicians considered the role of timbre in our hackathon. Our ethnographic approach collated interviews, participants' notes and accounts throughout the hackathon, which we synthesised into 11 descriptive accounts of the hackathon teams' design processes. Below, we offer our own reflections on the hackathon projects, revisiting each of our three research questions in Sections 5.1–5.3, respectively, and offer key takeaways across our discussion in Section 5.4. This is followed by a discussion of some limitations in Section 5.5, with highlighted opportunities for addressing them in future work.

Going forward, we refer to the teams using their team number as listed above (T1 - T11) in Section 4 and Table 1.

5.1 Reflections on RQ1: Negotiating Timbre

There were several ways that the participants thought about the concept of timbre whilst designing tools for instrument making. Overall, we observed that all the teams, in some shape or form, used metaphors as a way to reach a mutual understanding of timbre [63, 66]. These metaphors were sometimes abstract, using, for instance, multimodal shapes and colours to communicate about timbre by referencing other sensory experiences [9, 39]. For example, T5 were interested in identifying dimensions of timbre that could lead to different visualisations that were meaningful to understand. Multimodal metaphors also transitioned to the mapping process, as we observed for example with T4, who mapped the amount of colour gradient from their webcam feed to the amount of audio distortion, based on their underlying, mutual understanding that images with more grey should sound "crunchier" and more distorted.

As a tool to understand the world, abstract metaphors are derived from our embodied understanding and individual interpretation [40]. Several teams explored embodied and personal perspectives in their understanding of timbre, as seen for example in T3, who linked timbre to their personal physical experience of interacting with oscillators, or T11 who tried to directly map timbre generation with hand movements.

Mappings of timbre to meaningful metaphors were also focused around semiotics or semantic meanings. For example, T8 was interested in using ways people talk about sound to approach desired timbre outputs. They noted how "experienced musicians don't have knowledge about production but have the vocabulary about it" (cf. [63, 64]) which motivated them to leverage a ChatGPT-based model [57] – viewing the use of text-based prompts as a more natural way to describe and therefore interact with timbre, especially for non-expert software users. We also saw a text-based approach from T5: text prompts guided a parameter search for the CLAP embedding model [19], a neural network model connecting text descriptions with audio, affording text-to-audio generation. Here, T5 also had a mutual understanding of a conceptual link between timbre and language, seeing this more as a way to limit exploration of the wider timbre space.

The notion that teams use metaphors to communicate timbre corroborates studies in contemporary metaphor theory on signifier and signified [39], that is, using multimodal representations to map understanding from a known concept (signifier) to another unknown one (signified), and it is suggested that metaphors, including colour metaphors, are bound up in other domains through mediums such as language and categorisations [9]. As the way we describe things comes from pre-existing distinctions we learned from experience, the tools we use have a significant impact on how they are designed and communicated [74]. In the case of timbre, these distinctions can be learned within other modalities such as colour and visual references, and how timbre is communicated would be a factor of/thought about within the context of the tool being used to construct it. Given this, it is particularly important to note that metaphor is dynamic and flexible [63, 83]. As a tool for aligning respective subjective experience into something that is mutually understandable, metaphors are non-domain specific but come from places that co-opts different people’s knowledge to reach a working solution, this can be updated over time as knowledge changes.

Given the reliance on metaphor, we suggest that teams negotiated timbre by reducing its concept down to experiences that they were easily able to communicate given the team’s respective background. The understandings of timbre we observed might arise from attempting to reconcile the subjectivity of our individual perceptions of audio. As timbre is a broad category of experience and has many dimensions, to understand this nuance and reach mutual understandings, teams *had* to reduce the dimensionality to convey specific, targeted features of mutual interest. Given this, perhaps future investigations unpacking aspects of timbre from the “wastebasket” should emphasise the variety of metaphor approaches and what they might represent.

5.2 Reflections on RQ2: Collaborative Strategies

Amongst various understandings of timbre and metaphor used across each team’s design approaches, we noted links between teams’ understandings of timbre and their goals in the problem spaces they explored. To approach the concept of timbre, some teams tended to form concrete design objectives: T1, T3 and T9 all started the hackathon with a distinct idea of what they wanted to address in their design. This mirrors design literature on how teams start with unclear initial ideas and then reach a solution by iteratively converging and diverging over time [68][12]. For example, T3 had a preconceived motivation to focus on balancing randomness and expressivisity, later converging on more precise goals based on their choices of tools, given their availability and how the tools afforded manipulating timbre. Others started with a more interaction-led approach, exploring ways to generate or manipulate audio with regard to timbre: T4 and T11 for instance explored embodied approaches to interacting with timbre.

This embodiment was also reflected in the way the teams collaborated in the given physical space. With a community work space available, and the screen as a key shared interface, we observed different social dynamics and hierarchies unfolding throughout the hackathon. For example, how much bodily movement was performed, and how much eye contact was established relied on the initial spatial formations the teams took, as well as on how dominant the screen functioned as a communication interface. We observed that team members who moved around more (T5, T7), either changing position entirely or turning towards other members, seemed to take a less dominant role within the team. They adjusted accordingly, while the dominant members remained in more static postures. This became particularly obvious in groups with more than 2 members. The postural behaviour of the ones working in pairs was overall more balanced (T9, T8, T6, T1), with either mutually open or relaxed postures, or mutually little general movement and eye contact. Teams with more defined roles, such as T5 who followed a typical software development approach, even broke away from one another once tasks were delineated mirroring their dominance.

The teams working hybrid or with the screen as the central focus of interactions expectedly had less eye contact despite face to face interactions. Spatial formations determining collaborative settings changed according to the dominance of the screen, too. Rather than circular or L-shaped orientations, team members were lined up next to each other and communicated via the screen interface. This formation has previously been uncommon in face-to-face interaction [38, 49] and has been established in hybrid sessions as found in this hackathon.

5.3 Reflections on RQ3: Tools and Concepts of Timbre

In examining how hackathon teams conceptualised timbre through choice of metaphors, and negotiated team dynamics through their use of space whilst collaborating, we note that many negotiated their conceptualisation of timbre further given their choice of tools [74]. Most teams used frequency domain methods to both analyse and manipulate audio for timbre, corroborating a rich history of research on computing timbre features on magnitude spectra or spectrograms [7]. Even T4, who focused their project around interaction with a slinky, needed to work with technical representations of timbre, filtering audio spectrums in the frequency domain to manipulate audio with cross-modal data – despite their project focus taking a more embodied perspective on timbre. Other spectral audio features, such as MFCC, were also used by several teams (e.g. T1, T7, T9) and acted as a way for them to understand and describe the timbre of audio. This represents a particular techno-scientific understanding of timbre (see also Section 2.1) which might be a consequence of our recruitment strategy – many participants are researchers in MIR or closely related fields. This resonates with the interdisciplinary challenges raised by Aucouturier and Bigand [3] who highlight how objective conceptualisations of timbre in MIR are at tension with the more subjective understandings of timbre from music perception researchers, creating barriers between the data which MIR timbre tools provide and their meaningful interpretation in other domains (see also [71]). Or, it might be representative of the types of timbre tools available and the need for teams of researcher-developers to be able to work with tools they both understand.

Across the variety of tools used by hackathon participants, many do not afford listening to audio in real-time, which we suggest informed how participants thought about the concept of timbre.

[12]See Footnote 1.

For example, T5, T7, T8, and T9 all used Python notebooks where playback could only occur on compiling. In general, we found that teams using notebooks tended to focus more on inspecting visual elements of timbre. For example, T7 experimented in a Python notebook and largely analysed visual representations of timbre rather than assessing outputs through listening. We suggest that this influenced how people conceptualised timbre. Perhaps, using high-level concepts of timbre in lieu of listening, e.g., using MFCC or spectral centroids, was more helpful for certain teams (T1, T7, T9) because they could more easily reach common understandings. It is also possible that teams were relying on the features/constraints of their chosen tools because of how the perception of audio tends to vary on a more subjective and individual level [3]. Some teams used playback more frequently to aid their experimenting whilst hacking together their project, e.g., T4 experimenting with different representations of their webcam input and listening to its output in real-time. We suggest this is notable as live playback is fundamental in other music practices [25, 35, 53] including, we suggest, with many of the AI tools explored by our hackathon teams. Future work could explore the role of audio playback in the timbre tool designers toolkit and explore more deeply how this changes their understandings of timbre.

5.4 Summary

Overall, we suggest that participants' choice of tools seemed to inform how they conceptualised timbre in their submissions. Through the hackathon, participants have been relatively free to shape their vision of timbre, and this variety was captured by the range of tools and negotiations around timbre in our descriptions of each team's design process. Yet, whilst every team has their own way of approaching their design, we suggest that there is a shared information infrastructure for modelling timbre [52]. For example, we observed that teams had to find a common ground in terms of software, technical representations of timbre, as well as understandings of timbres as metaphors. There were also common patterns of negotiation between the choice of software and more practical conceptualisations of timbre e.g. using understandings of timbre embedded in software choices, which consequently influenced the types of sound people would create and how they would then work with these sounds.

We see a shared understanding of timbre as a space. However, there is a lack of agreement on what this space is, for instance, regarding its dimensions and shape. While there is a need for a definition of using space as a metaphor to approach timbre, any set of descriptors or representations used by the participants will have a specific cultural history surrounding it, and finding a unified solution might not be possible. Therefore, the entanglement between the initial set of dimensions and metaphors that people rely upon when discussing timbre, and the root tools that people rely upon when approaching the creation of a Timbre Tool, *becomes* the reflections of the participants' sonic cultures (see [52]).

Returning to Figure 2 and our approach to the hackathon, we can see that the tools themselves are a reification of this metaphorical communication. The designs are a representation (metaphor) of each group's understanding of timbre, manifested in the context of the hackathon. Reed et al.'s model of metaphor-based communication [63] demonstrates how metaphors negotiate subjective lived experiences between individuals in a cyclical way to reach a mutual understanding. Similarly, we note that our methodological approach involves reconciling realms of design and timbre thinking. The end result – the tool devised by each team – is a product of this mutual understanding. Extending beyond the timbre tools designed in this study, we argue that others dealing with subjectively experienced parameters might find a methodological approach similar to ours useful. The resulting technological implementations are coalescence and negotiation of the designers' lived experience come-to-life. Revisiting other technologies introduced in Section 2.1, we can further suggest that designs come to exist as products of the specific enculturing (in both sonic and other cultures) and interests of the designer(s).

5.5 Limitations and Future Work

Our current study has several limitations that could be addressed in future work. Foremost, our insights and findings are exploratory, and limited to the micro-culture within each team. It is clear that the method we used does not create generalisable insights for a broad range of cultural contexts. We see the work more as generative, noting several potential connections between the unique participant teams and how this informed their design processes, acting as a starting point for studying how members within a team form common grounds for concepts and collaborate with regard to timbre. We observed many complicated overlaps between teams' background, goals, motivations and technical expertise, which informed their understandings of timbre and hackathon project designs. Future work could apply theoretical frameworks or analytical techniques which embrace the entanglement in our collected data [52]. For example, a diffractive analysis approach [55, 81] might be well suited to capture the nuance and complex entanglement between participants and technologies in our hackathon context.

With our hackathon method, we followed an open-ended strategy to ask teams to self-report their design process. Teams are allowed to decide what they want to report and how to frame their report. It is possible that teams wanted to only give positive feedback, especially where some of the researchers are academics teaching hackathon participants. On the other hand, teams' self reports were useful when considered alongside our ethnographic data collection, affording a degree of freedom that helped us to collect a variety resources instead of concentrating solely on one level of detail. Our approach was also unobtrusive, not being too disruptive to hackathon teams' creative flow.

We also acknowledge that our recruitment and advertisement strategies for the hackathon directed our sample towards the sonic culture of programmers and developer-oriented individuals. The hackathon event was held in collaboration with the TAP community, and our promotions for the event targeted mostly research-related communities e.g., many researchers are in the Centre for Digital Music at QMUL. We acknowledge that this likely directed the sonic cultures involved in the event to programmer and developer-oriented design approaches. The warm-up workshops and living document of timbre resources also likely influenced the choices of tools selected; for example, we might have examined novelty effects where participants' were excited about using tools that they had just learnt. Additionally, we did not deeply examine into how our participants' demographics (Table A1), cultural backgrounds, or general musical tastes would have also influenced their perception and conceptualisation of timbre. Future editions of this hackathon could expand our findings by recruiting participants and designating teams, with the aim to drive conversation between participants and explore what timbre means as a facet of one's larger musical experience.

6 CONCLUSION

What can ethnographic approaches teach us about a social and collaborative understanding of timbre? Through the provided structure of a Hackathon we were able to extract different approaches towards integrating timbre as an active part in creating tools and technologies in music. Rather than being a limited technoscientific idea rooted in the psychoacoustical "timbre space" model [52], our study presents an ethnographic and multimodal understanding of timbre from the technologist's perspective, looking outside of the "wastebasket" box of timbre and learning about its social and collaborative qualities, informing future development of tools to assist timbre exploration in musical craft practice that can be adapted by a large community. With this exemplary work, we aim to establish hackathons and ethnographic approaches as a valid method to evaluate nuanced, hazily defined phenomena in the musical domain and beyond. Rather than providing definitive answers, we seek to—as one reviewer put it—"start interesting and important conversations about how we study effects that are socially and culturally situated."

ACKNOWLEDGMENTS

The Timbre Tools Hackathon was organised by Saitis, Del Sette, Shier, and Tian, in partnership with Joshua Hodge and The Audio Programmer community. Prizes were sponsored by Bela, Cycling '74, Juce, and the Audio Developer Conference. Submitted projects were judged by Joshua Hodge, Jean-Baptiste Thiebaut, Daniel Jones, Rodrigo Constanzo, Christian Steinmetz, Christopher Mitcheltree, and Saitis. We thank all hackathon participants and student helpers for their enthusiasm and contributions.

The ethnographic study was undertaken by Del Sette, Zheng, Skach, Reed, and Ford under the direction of Saitis and with the advice of Reed and Ford. We thank Landon Morrison for also contributing to the ethnography team and for fruitful discussions. Del Sette, Shier, Tian, and Zheng are, and Ford was when the work was done, research students at the UKRI Centre for Doctoral Training in Artificial Intelligence and Music, supported jointly by UK Research and Innovation (grant number EP/S022694/1) and Queen Mary University of London.

We thank three anonymous reviewers for insightful feedback on an earlier version of this paper.

REFERENCES

[1] Saul Albert, Claude Heath, Sophie Skach, Matthew Tobias Harris, Madeline Miller, and Patrick GT Healey. 2019. Drawing as transcription: how do graphical techniques inform interaction analysis? *Social Interaction Video-Based Studies of Human Sociality* (2019).

[2] Jack Armitage, Victor Shepardson, and Thor Magnusson. 2024. Tölvera: Composing With Basal Agencies. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. Utrecht, NL.

[3] Jean-Julien Aucouturier and Emmanuel Bigand. 2012. Mel Cepstrum & Ann Ova: The Difficult Dialog Between MIR and Music Cognition. In *Proc. International Society for Music Information Retrieval (ISMIR) Conference*. 397–402.

[4] Eliot Bates. 2021. The interface and instrumentality of Eurorack modular synthesis. In *Rethinking Music through Science and Technology Studies*, Antoine Hennion and Christophe Levaux (Eds.). Routledge, 170–188.

[5] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[6] Ragnhild Brøvig and Anne Danielsen. 2016. Autotuned Voices: Alienation and "Brokenhearted Androids". In *Digital Signatures: The Impact of Digitization on Popular Music Sound*, Ragnhild Brøvig and Anne Danielsen (Eds.). MIT Press, Chapter 7, 117–132.

[7] Marcelo Caetano, Charalampos Saitis, and Kai Siedenburg. 2019. *Audio Content Descriptors of Timbre.* Springer International Publishing, Cham, 297–333. https://doi.org/10.1007/978-3-030-14832-4_11

[8] Antoine Caillon and Philippe Esling. 2021. *RAVE: A Variational Autoencoder for Fast and High-Quality Neural Audio Synthesis.* arXiv:2111.05011 [cs, eess] http://arxiv.org/abs/2111.05011

[9] Clive Cazeaux. 2002. Metaphor and the Categorization of the Senses. *Metaphor and Symbol* 17, 1 (Jan. 2002), 3–26. https://doi.org/10.1207/s15327868ms1701_2

[10] 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

[11] Nuno N. Correia and Atau Tanaka. 2015. Prototyping Audiovisual Performance Tools: A Hackathon Approach. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, Edgar Berdahl and Jesse Allison (Eds.). Louisiana State University, Baton Rouge, Louisiana, USA, 319–321. https://doi.org/10.5281/zenodo.1179044

[12] Cycling '74. 2024. *Max [software].* https://cycling74.com/products/max

[13] Thomas Deacon, Nick Bryan-Kinns, Patrick G.T. Healey, and Mathieu Barthet. 2019. Shaping Sounds: The Role of Gesture in Collaborative Spatial Music Composition. In *Proceedings of the 2019 Conference on Creativity and Cognition* (San Diego, CA, USA) *(C&C '19)*. Association for Computing Machinery, New York, NY, USA, 121–132. https://doi.org/10.1145/3325480.3325493

[14] Emily I. Dolan. 2013. *The Orchestral Revolution: Haydn and the Technologies of Timbre.* Cambridge University Press.

[15] Emily I. Dolan and Thomas Patteson. 2018. Ethereal Timbres. In *The Oxford Handbook of Timbre*, Emily I. Dolan and Alexander Rehding (Eds.). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190637224.013.27

[16] Emily I. Dolan and Alexander Rehding (Eds.). 2021. *The Oxford Handbook of Timbre.* Oxford University Press.

[17] Andrew Dubber. 2016. You Have 24 Hours to Invent the Future of Music: Music Hacks, Playful Research and Creative Innovation. In *Business Innovation and Disruption in the Music Industry.* Edward Elgar Publishing, 211–228.

[18] Nina Sun Eidsheim. 2019. *The race of sound: Listening, timbre, and vocality in African American music.* Duke University Press.

[19] Benjamin Elizalde, Soham Deshmukh, Mahmoud Al Ismail, and Huaming Wang. 2023. CLAP Learning Audio Concepts from Natural Language Supervision. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* 1–5. https://doi.org/10.1109/ICASSP49357.2023.10095889

[20] Cornelia Fales. 2018. Hearing Timbre: Perceptual Learning among Early Bay Area Ravers. In *The Relentless Pursuit of Tone: Timbre in Popular Music.* Oxford University Press. https://doi.org/10.1093/oso/9780199985227.003.0002

[21] Jeanette Falk Olesen and Kim Halskov. 2020. 10 Years of Research With and On Hackathons. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 1073–1088. https://doi.org/10.1145/3357236.3395543

[22] Stefano Fasciani. 2020. Interactive Computation of Timbre Spaces for Sound Synthesis Control. (2020), 69–78. https://doi.org/10.25370/array.v20152528

[23] Robert Fink, Melinda Latour, and Zachary Wallmark (Eds.). 2018. *The relentless pursuit of tone: Timbre in popular music.* Oxford University Press.

[24] Meagan Flus and Ada Hurst. 2021. Design at hackathons: new opportunities for design research. *Design Science* 7 (2021), e4.

[25] 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 *Proceedings of the 2024 ACM Conference on Creativity and Cognition (C&C '24)* (Chicago, IL, USA). ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3635636.3656185

[26] Arthur W Frank. 2010. *Letting stories breathe: A socio-narratology.* University of Chicago Press.

[27] Kiev Gama, George Valença, Pedro Alessio, Rafael Formiga, André Neves, and Nycolas Lacerda. 2023. The developers' design thinking toolbox in hackathons: a study on the recurring design methods in software development marathons. *International Journal of Human–Computer Interaction* 39, 12 (2023), 2269–2291.

[28] Dana Gooley. 2013. Jazz piano pedaling and the production of timbral difference. *Keyboard Perspectives* 6 (2013), 101––126.

[29] Paul D Greene and Thomas Porcello (Eds.). 2005. *Wired for sound: Engineering and technologies in sonic cultures.* Wesleyan University Press.
[30] John Michael Grey. 1975. *An Exploration of Musical Timbre.* Ph. D. Dissertation. Stanford University.
[31] James Hodge, Sarah Foley, Dan Lambton-Howard, Laura Booi, Kyle Montague, Sandra Coulter, David Kirk, and Kellie Morrissey. 2023. Exploring Participants' Representations and Shifting Sensitivities in a Hackathon for Dementia. *ACM Trans. Comput.-Hum. Interact.* 30, 3, Article 46 (jun 2023), 35 pages. https://doi.org/10.1145/3571814
[32] 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 *Proc. International Society for Music Information Retrieval (ISMIR) Conference.* Montréal, Canada. https://arxiv.org/pdf/2010.05388.pdf
[33] Don Ihde. 2021. A Finnish turn: digital and synthesiser musical instruments. *Journal of New Music Research* 50, 2 (2021), 165–174.
[34] Alexander Refsum Jensenius. 2022. *Sound actions: Conceptualizing musical instruments.* MIT Press.
[35] Pirkko Juntunen, Inkeri Ruokonen, and Heikki Ruismäki. 2014. The impact of playback in learning musicianship skills and musical communication. *The changing face of music and art education : Education Interdisciplinary Journal for Music and Art Pedagogy - CFMAE journal* 6, 1 (2014), 49–68.
[36] Joakim Karlsen and Anders Sundnes Løvlie. 2017. 'You can dance your prototype if you like': independent filmmakers adapting the hackathon. *Digital Creativity* 28, 3 (2017), 224–239. https://doi.org/10.1080/14626268.2017.1351992
[37] Mark Katz. 2010. *Capturing Sound: How Technology has Changed Music.* University of California Press, Berkeley and Los Angeles, CA.
[38] Adam Kendon. 1990. *Conducting interaction: Patterns of behavior in focused encounters.* Vol. 7. CUP Archive.
[39] George Lakoff and Mark Johnson. 2008. *Metaphors we live by.* University of Chicago Press.
[40] George Lakoff, Mark Johnson, and John F Sowa. 1999. Review of Philosophy in the Flesh: The embodied mind and its challenge to Western thought. *Computational Linguistics* 25, 4 (1999), 631–634.
[41] Megan Lavengood. 2019. What makes it sound'80s? The Yamaha DX7 electric piano sound. *Journal of Popular Music Studies* 31, 3 (2019), 73–94.
[42] Tilmann Lindberg, Christoph Meinel, and Ralf Wagner. 2011. Design thinking: A fruitful concept for IT development? *Design thinking: Understand–improve–apply* (2011), 3–18.
[43] Thomas J Lodato and Carl DiSalvo. 2016. Issue-oriented hackathons as material participation. *New Media & Society* 18, 4 (2016), 539–557. https://doi.org/10.1177/1461444816629467
[44] Thor Magnusson. 2019. *Sonic Writing Technologies of Material, Symbolic, and Signal Inscriptions.* Bloomsbury Academic, New York, NY.
[45] Juan P Martinez Avila, Vasiliki Tsaknaki, Pavel Karpashevich, Charles Windlin, Niklas Valenti, Kristina Höök, Andrew McPherson, and Steve Benford. 2020. Soma Design for NIME. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, Romain Michon and Franziska Schroeder (Eds.). Birmingham City University, Birmingham, UK, 489–494. https://doi.org/10.5281/zenodo.4813491
[46] Stephen McAdams. 2019. The Perceptual Representation of Timbre. In *Timbre: Acoustics, Perception, and Cognition.* Springer International Publishing, 23–57.
[47] S. McAdams, Suzanne Winsberg, Sophie Donnadieu, Geert De Soete, and Jochen Krimphoff. 1995. Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research* 58, 3 (1995), 177–192.
[48] Andrew McPherson and Victor Zappi. 2015. An environment for submillisecond-latency audio and sensor processing on BeagleBone Black. In *Audio Engineering Society Convention 138.* Audio Engineering Society.
[49] Albert Mehrabian. 1969. Significance of posture and position in the communication of attitude and status relationships. *Psychological Bulletin* 71, 5 (1969), 359.
[50] Alessandro B. Melchiorre and Markus Schedl. 2020. Personality Correlates of Music Audio Preferences for Modelling Music Listeners. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (Genoa, Italy) *(UMAP '20).* Association for Computing Machinery, New York, NY, USA, 313–317. https://doi.org/10.1145/3340631.3394874
[51] 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
[52] Landon Morrison. 2022. Hearing Timbre at the Crossroads. *Kalfou* 9, 2 (Fall 2022), 328–336. https://www.proquest.com/scholarly-journals/hearing-timbre-at-crossroads/docview/2763578225/se-2
[53] Chris Nash and Alan Blackwell. 2012. Liveness and Flow in Notation Use. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME).* University of Michigan, Ann Arbor, Michigan. https://doi.org/10.5281/zenodo.1180547
[54] Alexander Nolte, Ei Pa Pa Pe-Than, Abasi-amefon Obot Affia, Chalalai Chaihirunkarn, Anna Filippova, Arun Kalyanasundaram, Maria Angelica Medina Angarita, Erik Trainer, and James D Herbsleb. 2020. How to organize a hackathon–A planning kit. *arXiv preprint arXiv:2008.08025* (2020).
[55] Charlotte Nordmoen and Andrew P. McPherson. 2022. Making space for material entanglements: A diffractive analysis of woodwork and the practice of making an interactive system. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22).* Association for Computing Machinery, New York, NY, USA, 415–423. https://doi.org/10.1145/3532106.3533572
[56] Jeanette Falk Olesen, Nicolai Brodersen Hansen, and Kim Halskov. 2018. Four factors informing design judgement at a hackathon. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction* (Melbourne, Australia) *(OzCHI '18).* Association for Computing Machinery, New York, NY, USA, 473–483. https://doi.org/10.1145/3292147.3292155
[57] OpenAI. 2022. *ChatGPT: Optimizing Language Models for Dialogue.* https://openai.com/blog/chatgpt/ (Online).
[58] Ei Pa Pa Pe-Than, Alexander Nolte, Anna Filippova, Christian Bird, Steve Scallen, and James Herbsleb. 2022. Corporate hackathons, how and why? A multiple case study of motivation, projects proposal and selection, goal setting, coordination, and outcomes. *Human–Computer Interaction* 37, 4 (2022), 281–313.
[59] Jon Pigott and Aidan Taylor. 2024. Sonic cultures of making: DIY sound and electronics since 1981. *Sound Studies* (2024), 1–19. https://doi.org/10.1080/20551940.2024.2304439
[60] Trevor Pinch and Frank Trocco. 2004. *Analog Days: The Invention and Impact of the Moog Synthesizer.* Harvard University Press.
[61] Jenny Preece, Yvonne Rogers, Helen Sharp, David Benyon, Simon Holland, and Tom Carey. 1994. *Human-computer interaction.* Addison-Wesley Longman Ltd.
[62] Catherine Provenzano. 2018. Auto-Tune, Labor, and the Pop-Music Voice. In *The relentless pursuit of tone: Timbre in popular music*, Robert Fink, Melinda Latour, and Zachary Wallmark (Eds.). Oxford University Press, Chapter 7, 159–181.
[63] Courtney N. Reed, Paul Strohmeier, and Andrew P. McPherson. 2023. Negotiating Experience and Communicating Information Through Abstract Metaphor. 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 185, 16 pages. https://doi.org/10.1145/3544548.3580700
[64] Charalampos Saitis, Claudia Fritz, Gary P. Scavone, Catherine Guastavino, and Danièle Dubois. 2017. Perceptual evaluation of violins: A psycholinguistic analysis of preference verbal descriptions by experienced musicians. *Journal of the Acoustic Society of America* 141, 4 (2017), 2746–2757.
[65] Charalampos Saitis, Hanna Järveläinen, and Claudia Fritz. 2018. The Role of Haptic Cues in Musical Instrument Quality Perception. In *Musical Haptics*, Stefano Papetti and Charalampos Saitis (Eds.). Springer, Cham, 73–93.
[66] Charalampos Saitis and Stefan Weinzierl. 2019. The Semantics of Timbre. In *Timbre: Acoustics, Perception, and Cognition*, Kai Siedenburg, Charalampos Saitis, Stephen McAdams, Arthur N. Popper, and Richard R. Fay (Eds.). Springer, Cham, 119–149.
[67] Jeffrey Sanchez-Burks, Sally Blount, and Caroline A. Bartel. 2006. Fluidity and Performance in Intercultural Workplace Interactions: The Role of Behavioral Mirroring and Relational Attunement. *Ross School of Business Paper* 1039 (Sept. 2006), 56. https://doi.org/10.2139/ssrn.928584
[68] Elizabeth B.-N. Sanders and Pieter Jan Stappers. 2008. Co-creation and the new landscapes of design. *CoDesign* 4, 1 (2008), 5–18.
[69] Cleo Schulten, Alexander Nolte, Daniel Spikol, and Irene-Angelica Chounta. 2022. How do participants collaborate during an online hackathon? An empirical, quantitative study of communication traces. *Frontiers in Computer Science* 4 (2022). https://doi.org/10.3389/fcomp.2022.983164
[70] Jordie Shier, Charalampos Saitis, Andrew Robertson, and Andrew McPherson. 2024. Real-time timbre remapping with differentiable DSP. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME).* Utrecht, NL.
[71] Kai Siedenburg, Ichiro Fujinaga, and Stephen McAdams. 2016. A comparison of approaches to timbre descriptors in music information retrieval and music psychology. *Journal of New Music Research* 45, 1 (2016), 27–41.
[72] Kai Siedenburg and Stephen McAdams. 2017. Four distinctions for the auditory "wastebasket" of timbre. *Frontiers in Psychology* 8 (2017), 272043.
[73] Kai Siedenburg, Charalampos Saitis, Stephen McAdams, Arthur N. Popper, and Richard R. Fay (Eds.). 2019. *Timbre: Acoustics, Perception, and Cognition.* Springer, Cham.
[74] Joe Snape and Georgina Born. 2022. Max, music software and the mutual mediation of aesthetics and digital technologies. In *Music and Digital Media: A Planetary Anthropology*, Georgina Born (Ed.). UCL Press, London, 220–266. http://www.jstor.org/stable/j.ctv2pzbkcg.11
[75] Makis Solomos. 2020. *From Music to Sound: The Emergence of Sound in 20th-and 21st-Century Music.* Routledge, Abingdon.
[76] Andrew Sparkes and Brett Smith. 2008. Narrative Constructionist Inquiry. In *Handbook of Constructionist Research*, James Holstein and Jaber Gubrium (Eds.). Guilford Publications, London, 295–314.

[77] 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
[78] Bob L. Sturm, Oded Ben-Tal, Úna Monaghan, Nick Collins, Dorien Herremans, Elaine Chew, Gaëtan Hadjeres, Emmanuel Deruty, and François Pachet. 2019. Machine Learning Research That Matters for Music Creation: A Case Study. *Journal of New Music Research* 48, 1 (2019), 36–55. https://doi.org/10.1080/09298215.2018.1515233
[79] Paul Théberge. 1997. *Any Sound You Can Imagine: Making Music/Consuming Technology.* Wesleyan University Press, Hanover, NH.
[80] Daniel Villegas Vélez. 2018. The Matter of Timbre: Listening, Genealogy, Sound. In *The Oxford Handbook of Timbre*, Emily I. Dolan and Alexander Rehding (Eds.). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190637224.013.20
[81] Ron Wakkary, Doenja Oogjes, Nazmus Sakib, and Armi Behzad. 2023. Turner Boxes and Bees: From Ambivalence to Diffraction. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 790–807. https://doi.org/10.1145/3563657.3596081
[82] Zachary Wallmark. 2022. *Nothing but Noise: Timbre and Musical Meaning at the Edge.* Oxford University Press.
[83] Bodo Winter. 2019. *Sensory Linguistics: Language, perception and metaphor.* John Benjamins.
[84] Asterios Zacharakis, Konstantinos Pastiadis, and Joshua D. Reiss. 2014. An Interlanguage Study of Musical Timbre Semantic Dimensions and Their Acoustic Correlates. *Music Perception* 31 (2014), 339–358.
[85] Victor Zappi and Andrew McPherson. 2018. Hackable Instruments: Supporting Appropriation and Modification in Digital Musical Interaction. *Frontiers in ICT* 5 (2018). https://doi.org/10.3389/fict.2018.00026

A PARTICIPANT BACKGROUNDS

Self-reported demographic information and timbre familiarity of participants are summarised in Table A1.

B ETHNOGRAPHER POSITIONALITY

Del Sette has a Western classical music background focused on electronic and computer music, both through artistic experience and academic studies. His timbre perspectives are rooted especially in concrete and acousmatic music. He currently works in HCI research for music listening healthcare interventions with a participatory approach.

Zheng is a PhD student in the field of explainable AI for musical audio synthesis. His long-term research goal focuses on developing interactive and understandable AI systems that facilitate musical creation and expression, and on understanding how these technological advances impact artistic practices. They also have a background in electronic music composition and production.

Skach is a researcher and designer in human-centric computing, smart textiles and soft robotics. Her research experience comprises social behavioural studies, using ethnographic methods, wearable sensing, and conversation analysis tools to study embodied interaction. During the hackathon, Sophie focused on observing nonverbal communication, including spatial orientation, posture, gaze, and other group dynamics.

Reed grew up in the Northeast USA with Western pop-rock and classical vocal music, later studying Western classical and American jazz traditions at university. She is experienced as a semi-professional singer in Western vocal styles, including choral music and Baroque and Romantic opera, and electronic music; her timbre perspectives are deeply rooted in vocal pedagogy from these experiences. She currently works as an HCI researcher in DMI design, focusing on designing for vocalists and vocal bodies.

Ford is a HCI researcher from England, usually working in the Creativity Support Tool sub-field. They typically use mixed-methods with standard HCI data collection approaches (e.g. questionnaires, interviews) and recently used first-person and ethnographic methods. They have explored how novices and experts use AI to make music but hasn't focused directly on timbre. Their perspective on timbre is influenced by their Western degree training in music technology.

Morisson is a music theorist who studies the role of technological mediation in 20th- and 21st-century sonic practices, focusing on electroacoustic sound, timbre, microtonal systems, and popular music culture. His current research focuses on how the idea of timbre is constructed in audio technologies and software, and what this means for a politics of machines listening.

Saitis has a background in both audio technology and psychoacoustics, with experience in the experimental investigation of timbre primarily through the dissimilarity and semantic differential paradigms as well as audio content analysis; his timbre perspectives are deeply rooted in Western orchestral instrument tones from these experiences.

C SEMI-STRUCTURED INTERVIEW PROMPTS

- Introduce yourselves, the team, why are you doing the hackathon this weekend, what brought you together?
- What are you working on?
- What kind of problem/s are you trying to solve?
- How does it/ do they relate to the creative prompts?
- How do you (as a team) think about timbre in your project?
- Did anyone had to think differently about timbre than they usually do in their own Musical/craft/coding/other practice?
- What tools are you using/ do you plan to use? Why?
- The project (the "tool") you are developing, who is it for? What kind of maker do you think might use it? For what kind of instrument/s might it be useful?

D LIST OF ABBREVIATIONS

CQT Constant-Q Transform (CQT) is an algorithm that converts an audio signal into spectral components suitable for musical representation.
CLAP Contrastive Language-Audio Pretraining (CLAP) [19] is an audio-language model that can extract latent representations of audio and text.
FFT Fast Fourier Transform (FFT) is an algorithm that converts an audio signal into spectral components that provide frequency information about the signal.
MSP Max Signal Processing (MSP) is a visual programming language for interactive music and audio programs [12].
MFCC Mel Frequency Cepstral Coefficients (MFCC) is a compressed audio feature obtained from the Log-Mel spectrogram.

Table A1: Self-reported demographic information and timbre familiarity of participants grouped per hackathon team.

	Age	Gender identity	Ethnic group	Formative years	Current residence	Timbre familiarity
T1	24-38	Male (4)	White (2) Mixed (1) Asian (1)	Italy (1) India (1) Romania (1) Netherlands (1)	UK (2) Netherlands (1) India (1)	average moderate
T2	24-35	Male (3) Non-Binary (1) Prefer not to say (1)	White (2) Black (1) Other (1) Prefer not to say (1)	UK (2) Portugal (1) Spain (1) Uruguay (1)	Iceland (1) Portugal (1) UK (1) Germany (1) France (1)	moderate/extensive
T3	39	Male	White	UK	UK	somewhat
T4	23-55	Female (2) Male (3)	White (3) Prefer not to say (2)	Poland (1) Italy (1) German (1) USA (1) Finland (1)	Denmark (5)	moderate/extensive
T5	26-33	Female (1) Male (3)	White (2) Mixed (1) Asian (1)	UK (2) India (1) Canada (1)	UK (4)	average moderate
T6	24	Female (2)	Asian (2)	China (2)	UK (2)	slight/somewhat
T7	28-55	Male (2) Prefer not to say (1)	White (1) Black (1) Prefer not to say (1)	UK (2) USA (1)	UK (3)	average moderate
T8	22	Male (2)	White (2)	UK (2)	UK (2)	somewhat/moderate
T9	25-26	Male (2)	White (1) Asian (1)	Wales (1) China (1)	UK (2)	somewhat/moderate
T10	40	Male	White	Belgium	UK	moderate
T11	35	Male	Asian	India	UK	somewhat

RAVE Realtime Audio Variational autoEncoder (RAVE) [8] is a variational autoencoder for fast and high-quality audio synthesis. Pretrained RAVE models can be run in Max in realtime.

OSC Open Sound Control (OSC) is a data transmission protocol for realtime messaging between devices.

VST Visual Studio Technology (VST) is an audio-plug-in interface for integrating audio effects and synthesisers into digital audio workstations.

JUCE JUCE is an open-source cross-platform C++ framework that supports the development of audio plug-ins in formats like VST3.