

Rescuing Performance from the Demo: Co-Designing Drum Gesture Mappings with a Percussionist

Jordie Shier *

Centre for Digital Music
Queen Mary University of London
London, UK

Teresa Pelinski

Centre for Digital Music
Queen Mary University of London
London, UK

Charalampos Saitis

Centre for Digital Music
Queen Mary University of London
London, UK

Andrew Robertson

Ableton AG
Berlin, Germany

Andrew McPherson

Dyson School of Design Engineering
Imperial College London
London, UK

Abstract

Augmenting instruments with sensors and neural network mappings is a well-explored digital musical instrument design approach. While augmentations can create new expressive opportunities, they also exert aesthetic influence and can constrain musicians’ gestural language, which, if left unchecked, can lead to technological capture. To examine this, we conducted a study with a professional percussionist, co-developing a gesture mapping toolkit and recording a ten-track album. Drawing on the concept of productive dissonance, our study aimed to hold the musician’s aesthetic in tension with technological constraints. This, along with a practice-based reflective approach, supported the development of a continuous gesture recognition method for percussive mapping and surfaced insights into the design process. We identify knowing-when as a form of tacit knowledge that supported productive dissonance, and raise an open question: absent a musician’s broader social context, how do we know whether a technology’s influence is genuinely supporting their practice?

1 INTRODUCTION

Designing digital musical instruments (DMIs) often involves creating mappings between sensors and audio synthesis algorithms (Hunt et al., 2003), allowing for gestural manipulation of diverse sound sources. Machine learning and artificial intelligence are increasingly being used to create these mappings, supporting complex relationships between input and sound generation (Jourdan and Caramiaux, 2023). When applied to existing instruments, such mappings can allow musicians to “recycle” their existing musical skills within a new timbral sound world (Tremblay and Schwarz, 2010).

However, recent critical discourse has raised the risk of reification of mapping, where the metaphor of mapping comes to define musical performance itself (McPherson et al., 2024). In this paper, we examine this risk and identify a concrete consequence, which we call *technological capture*, where the constraints of a mapping system overwhelmingly shape musical practice and performance, ultimately preventing musicians from “recycling” their musical skills.

To engage with this tension, we propose a bespoke co-design approach strongly influenced by technical practice research (TPR) (Pelinski et al., 2025). More specifically, a researcher (the first author) worked closely with a professional percussionist over the course of four months to co-develop a drum gesture mapping toolkit and incorporate it into the percussionist’s practice. Creating and recording a musical performance with our co-designed mappings was a key goal, providing a musical anchor to technical practices.

The contributions of this paper are twofold. First, as outcomes of the co-design, we share the co-developed drum gesture mapping toolkit built in Max4Live and a ten-track album presented as an annotated portfolio (Gaver and Bowers, 2012) on an accompanying website.¹ The toolkit contributes a novel method for real-time mapping of continuous percussive gestures (e.g., brushing) using neural networks, extending the trigger-based approach to drum augmentation. Second, we share insights from attempting to put the critical discourse on mapping into practice: our collaboration was designed to work against the risk of technological capture by holding a professional musician’s aesthetic in tension with the constraints of our system and by anchoring the work in the making of an album rather than a technical demonstration. We found that sustaining that tension required a form of judgement we characterise as *knowing-when* (Ryan, 2009): tacit knowledge of when to attend to the technology and when to the music. We argue this is not a protocol that can be specified in advance, but a judgement cultivated through sustained collaborative practice. Beyond these contributions, our study surfaced a limitation of the laboratory setting itself. We reflect on this in our discussion and present it as an open question to the community: absent a musician’s broader social context, how do we know whether a technology’s influence is genuinely supporting their practice?

2 “RESCUING THE PERFORMANCE FROM THE DEMO”

The provocation “rescue performance from the demo” offered by Norman et al. (2025) highlights a key tension in designing and performing with new technology for music,

*j.m.shier@qmul.ac.uk

¹<https://jordieshier.com/projects/aimc2026/>

including DMIs. Is the performance about the music or about demonstrating the affordances of the new technology? These goals are often at odds with each other in an unresolvable way. We argue that this tension is actually a positive, echoing what McPherson et al. (2024) describe as “productive dissonance” and the notion of *dissensus* (Norman et al., 2025). It is when this tension resolves in the direction of technology that the driving aesthetic force behind a musical performance is subordinated by the goals of technology development. We borrow the term *capture* from regulatory theory (Dal Bó, 2006), where it describes the process by which a regulatory body comes to serve the interests of the industry it was designed to regulate. By analogy, *technological capture* describes the process by which musical practice comes to serve the interests of technology development rather than the other way around.

Gestural control in musical interaction design offers a useful reflection on when and how “technological capture” manifests. Tuuri et al. (2017) describe the *choreographing effects* of technology on musical gesture, querying whether it is the musician who controls the technical system or the other way around. They propose two modes of technological choreographing: *push and pull effects*. An example was demonstrated by Jack et al. (2017) in a performance study with a digital percussive instrument, where musicians discarded gestures that produced no sonic result on the instrument (push effects) and converged on the most ergonomically convenient gestures that achieved sonic results (pull effects). Jack et al. use the term *bottleneck* to characterise attributes of a DMI that act to constrain a gestural language and steer a musician to play in certain ways.

Bottlenecks in DMI mapping are largely unavoidable due to the often necessary process of representing a musical performance in a computationally efficient form. Representations are useful for analysing phenomena (McPherson et al., 2024; Saitis et al., 2025) and as input to machine learning algorithms, but they can become problematic when a reduced symbolic representation of music is inverted to become its generative basis. For example, if a DMI applies an onset detector to an audio stream, then it presupposes that “onset” is a musically meaningful concept, even though onsets might not be a stable, pre-existing entity in the absence of an onset detector (Reed et al., 2024). A musical worldview containing onsets can be aesthetically desirable in certain contexts: for instance, a kick drum trigger is often used in metal performance to achieve timbral consistency and clarity during rhythmically fast or complex sections (Weatherhead, 2025). But to assume it is universal to all music would be problematic. This underscores the importance of contextualising technical systems within a “well-defined musical aesthetic” (McPherson et al., 2024), where unsuitable representations can generate productive dissonances to guide design.

To unpack these tensions, our study centres on directly collaborating and co-designing with a professional musician to provide a well-defined musical aesthetic to work against.

3 METHODOLOGY

In this section, we summarise how the study-collaboration was structured around co-design and practice research, and provide details on documentation and reflection.

3.1 Study Co-Design

The study was conducted over four months at Queen Mary University of London qMedia Studios, structured in three phases. The first phase consisted of a semi-structured background interview and initial co-design session, during which the researcher learned about the musician’s broader practice and aesthetic background, introduced the technical starting point, and conducted an initial brainstorming session to identify musical directions. The second phase consisted of five co-design sessions (2.5 hours each), spaced approximately two weeks apart, interleaved with periods of technical development by the researcher. Each session followed a consistent structure: a brief semi-structured reflection on the previous session, a co-design working block focused on developing and refining a mapping for a specific musical context, recording a short musical performance, and a reflection on the performance to set goals for the next sprint. The third phase consisted of two full recording days structured to replicate a real studio session, with making music, rather than technical development, as the explicit driving goal to anchor the co-design in a compositional aim. Thoroughly documenting the study was an important consideration to support reflexive practices, and is detailed in Section 3.3. The musician was paid at the local union hourly and daily rates throughout.

3.2 The Participants-Collaborators

The study was conducted by the first author (referred to as *the researcher*) and Jem Doulton (referred to as *the musician*), who was hired as a musical collaborator. The study originated after the researcher was introduced to the musician at a concert; the musician expressed interest in the researcher’s work on percussive gesture mapping, and the study design (detailed in Section 3.1) was planned collaboratively over subsequent conversations. Background information on the musician (collected from the background interview) and researcher is provided here to situate their collaboration and individual aesthetic background.

3.2.1 The Musician

Jem Doulton is a professional percussionist working in London, UK, playing in multiple musical projects ranging from folk country to free improv. He describes his aesthetic as occupying a space between rock and jazz, both sonically and in its energy and attitude, in his own words, music that has a “fuck you” quality.² The musician considered tacit knowledge a key element of his practice, for instance in the selection of cymbals, drums, sticks, and other “trinkets” to achieve a desired sonic result during a performance.

Jem put special value on the dexterity and “hands-on” character of his practice, an aspect which became a touchstone in the co-design study, which he described as “marrying the tech and the human.” In his past experience, incorporating digital technology into his practice had sometimes worked against this dexterity. In particular, he described an experience of playing in a band where over the course of several years his full acoustic drum kit was slowly replaced by electronic drum pads to the point where he felt he could have been replaced by a backing track: “I felt trapped ultimately, and that is where I think you can take the tech a little bit too far.” During the co-design study we repeatedly returned to this anecdote as an example of what we wanted

²Henceforth, all literal quotes are from the collected study documentation unless indicated otherwise.

to avoid—the “technological capture” we have described in Section 2.

3.2.2 The Researcher

The researcher, who is the first author of this paper, has a background in computer science and received formal training in audio engineering, and music theory/composition at an undergraduate level in a contemporary classical context. Prior to academia, he was a professional musician and performed and produced electronic music in several collaborative projects. In his primary musical project, Napoleon Skywalker, he produced and performed electronic dance music using Ableton Live and analogue synthesizers, and performed alongside a live drummer. Sound design and synthesizer programming are central to his musical practice, influenced by experimental electronic dance music and hip-hop. This practice has directly informed his academic research on real-time percussion-to-synthesizer timbre mapping and DMI design more generally. His technical research in this area includes the application of deep learning and evolutionary algorithms to estimate synthesizer parameters and to facilitate real-time, audio-driven interaction.

3.3 Documenting and Reflecting

In this study-collaboration we followed a practice research approach; this means that our principal research method is the collaborative design *practice* itself (Bulley and Şahin, 2021). Practice research is not uncommon in instrument design (Bowers and Green, 2018; Johnston and Ferguson, 2016; Pelinski et al., 2026) or generally in design (Bang et al., 2023; Gaver, 2012; Devendorf et al., 2020) or the arts (Borgdorff, 2010; Spatz, 2015). *Practice research* is however a broad term; here we draw primarily from the technical practice research (TPR) that Pelinski et al. (2025) propose, also inspired by reflective design (Sengers et al., 2005) and feminist HCI (Rode, 2011). TPR emphasises documenting and reflecting on the practice as closely as possible to the moment of it; seeking to capture the mutual mediations between technical development, aesthetics, and social dimensions; and, importantly, their evolution as the process unfolds. Pelinski et al. also point out the relevance of documenting and reflecting on the technical development as it unfolds; as the more technical phases of instrument design are often disregarded in practice research approaches to DMI design.

The practice was documented in a variety of forms. All studio sessions were audio and video recorded, including the interactions between researcher, musician, and studio visitors, which were transcribed and later thematically coded (Braun and Clarke, 2019); this was also the case for all interviews conducted outside of the studio sessions. Technical development was documented through version control (git) and through the researcher’s design journal.

We followed several parallel strategies to foster reflection: The researcher’s design journal served not only to document the technical process but also for the researcher to reflect on it as it unfolded. git commit messages were used in a similar way. Between sessions, the researcher went through the previous sessions’ documentation, incorporating the musician’s references and aesthetic dimensions into the technical development loops. The researcher also encouraged the musician’s reflection through questions asked at the beginning of each session and after periods of performance to query their felt experiences, inspired by

Petitmengin (2006). Additionally, the second author conducted five semi-structured interviews with the researcher periodically throughout the study, asking questions related to the progress of technical development and prompting reflection about the evolution of the co-design itself. These interviews served as a form of dialogic reflection and also captured “snapshots” of the process through the viewpoint of the researcher. The insights presented in the later sections are the product of revisiting the collected documentation and further articulating these reflections.

4 TOOLKIT AND PERFORMANCE

Augmenting percussion instruments with sensors and mappings to synthesizers has been explored in both commercial tools such as Sensory Percussion³ and open-source systems such as DataKnot.⁴ Both tools support percussive gesture recognition, including onset detection, classification, and mapping of drum strokes to samples or synthesis parameters. The technical starting point for this study was a set of gesture mapping tools built on DataKnot and developed by the researcher (Shier et al., 2024, 2025).

4.1 Gesture Mapping Toolkit

The toolkit can be broadly divided into two groups: objects that perform gesture recognition on streaming percussive audio and map the resulting features to synthesizer parameters, and objects that support musical structure within a live performance context such as audio looping and MIDI sequencing. Figure 1 provides a visual overview of the stages of development throughout the co-design timeline and Figure 2 overviews the neural network mapping systems described in the following subsections. Full technical details are provided in an appendix on the accompanying website.

4.1.1 Discrete Gesture Recognition

The discrete gesture recognition component uses onset detection to determine when a drum has been struck, followed by audio feature extraction and classification with a lightweight MLP. This is the standard approach in existing percussive mapping tools—onset detection provides a reliable, low-latency event trigger, and classification between playing zones such as drum head and rim was sufficient for the performances we developed. However, as we discuss in Section 5.1.1, this approach carries a representational assumption: it renders all percussion as a stream of discrete events, which is not universal to all playing techniques or musical contexts.

4.1.2 Continuous Gesture Recognition

A significant portion of the study was devoted to extending the toolkit to support continuous gesture recognition, which involves mapping gestures that sustain over time rather than triggering discrete events. To our knowledge, recognising a sustained gesture as a category distinct from discrete strikes and using it to drive control-rate parameters of a parametric synthesizer has not been implemented in a real-time percussive performance context.⁵ We focused on two gesture types. The first were sustained gestures such as

³<https://sunhou.se/sensorypercussion>

⁴<https://github.com/rconstanzo/data-knot>

⁵Sensory Percussion has a continuous interpolation mode; however, it is still based on a sequence of discrete onset triggers.

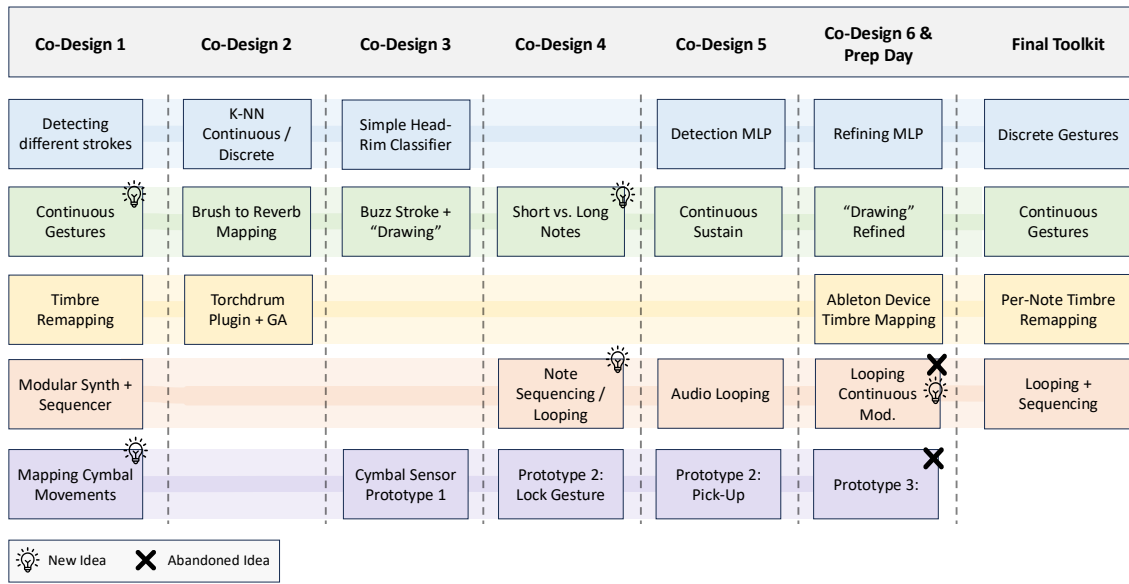


Figure 1: Key developments and insights throughout the various co-design sessions, and what elements of the toolkit ended up being included in the final kit. A cymbal sensor was explored, but didn't make it into the final kit. Section 5.1.2 provides additional information on the cymbal sensor development and subsequent abandonment.

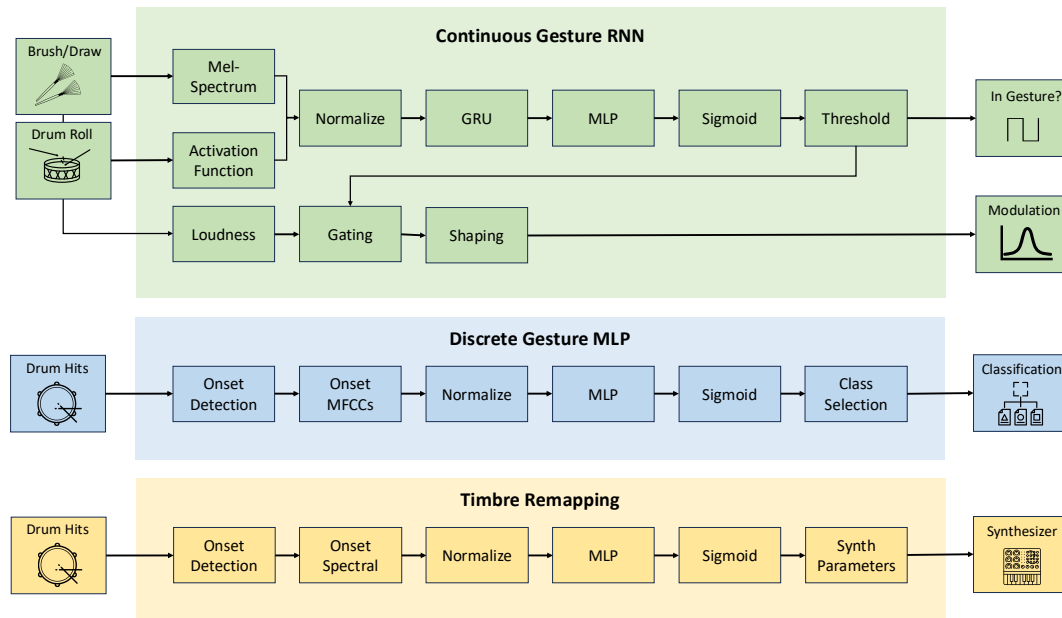


Figure 2: Overview of the three gesture recognition systems: the continuous RNN, discrete MLP, and timbre remapping. The continuous gesture RNN was trained on either brushing/drawing gestures or to detect buzz rolls—these are separate networks, but share the same architecture. The output of the continuous RNN is a binary signal indicating whether a continuous gesture has been detected and a continuous modulation signal. The discrete MLP is trained as a multiclass classifier based on different types of drum hits (i.e., head vs. rim). Timbre remapping uses a neural network trained as a regressor on pairs of spectral audio features and synthesizer parameters for a specific synthesizer. Each recognition algorithm supports a different style of playing and were combined in various ways across the album recording.

wire brush strokes on a drum membrane. An initial NMF-based approach (Lee and Seung, 1999; Smaragdis and Brown, 2003) was explored to isolate brush-like spectral components, but was unable to reliably exclude discrete strikes. We subsequently trained a recurrent neural network (RNN) — specifically a gated recurrent unit (GRU) (Cho et al., 2014) following the architecture used by Engel et al. (2020) — on manually labelled recordings from the co-design sessions, using a Mel-scaled spectral input representation. The second gesture type was buzz stroke rolls. Here the same RNN approach was applied using amplitude envelope features as input, trained to distinguish buzz strokes from regular double strokes. Both models were implemented for real-time inference using the RTNeural plugin for Max/MSP,⁶ built on RTNeural (Chowdhury, 2021), with frame-based features mappable to parameters in Ableton Live.

4.1.3 Timbre Remapping and Performance Devices

The toolkit also includes objects for timbre remapping (Stowell and Plumbly, 2010; Shier et al., 2025), from percussive input to synthesizer parameters, and performance devices including audio looping and melodic MIDI sequencing. These are described in more detail in Appendix A.

4.2 Musical Performance

The aesthetic direction for the album emerged early in the collaboration. During the first session the musician mentioned a previous solo project, a drum and electronics album he had made as a soundtrack to a play,⁷ and noted that he had “always wondered if I’d ever do another [one].” This reference lay dormant until much later in the study, when the format of the recording crystallised: short, dense grooves—“a beat tape”—with heavy synthesizer sounds that “could potentially be danced to.” The resulting album, recorded over two full days structured to replicate a real studio session, consists of ten tracks and constitutes the musical output of the study. It is presented as an annotated portfolio (Gaver and Bowers, 2012) with edited video performances available on the accompanying website. A public release with professional mix and mastering is planned. Figure 3 shows a photo taken from the recording day of the musician performing.



Figure 3: Performance photo of the musician playing the augmented drum kit designed as a part of this study.

⁶https://github.com/spluta/RTNeural_Plugin

⁷Short Stories by Jem Doulton: <https://jemdoulton.bandcamp.com/album/short-stories>

5 PRODUCTIVE DISSONANCE IN PRACTICE

In this section we present design arcs from the study-collaboration that illuminate where productive dissonance was sustained and where it surfaced limitations.

5.1 Representational Dissonances

Representational dissonances arise when the computational representation of gesture imposed by a mapping system fails to capture the range of a musician’s embodied practice. It describes the gap between the computational representation of a gesture (what the system can “see”) and the gesture itself (what the musician is actually doing). The following arcs trace how this gap surfaced through contact with practice, and how we designed in response to this dissonance.

5.1.1 Continuous Gestures

The initial mapping system introduced to the musician used onset detection to trigger sounds in response to drum hits. A representational dissonance arose when the musician pulled out a wire drum brush and asked: “how does a brush sound?” The resulting sound was not reflective of the gesture, unsurprisingly, as continuous brushing gestures are outside the representational scope of an onset detector. The musician suggested that these types of continuous gestures could be mapped to synthesizer parameters, an avenue that occupied a significant portion of the study. The continuous gesture recognition started with brushing and was expanded to include a *drawing* gesture with the tip of a drumstick, visually demonstrated in Figure 4, and buzz stroke rolls, which the musician described as being distinguishable from controlled rolls like double strokes.

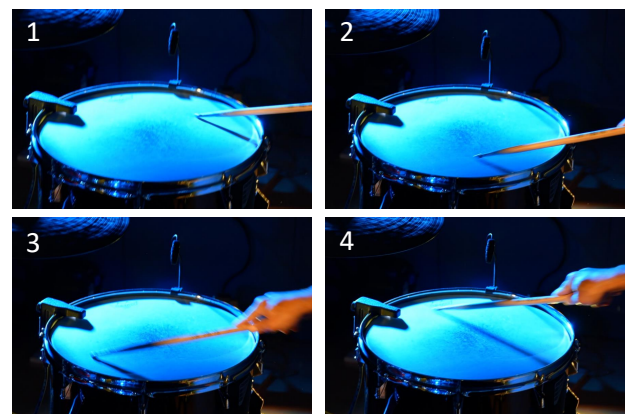


Figure 4: Drawing gesture used for continuous modulation.

Extending the system with a neural network for continuous gesture detection expanded the representational boundaries of the onset detector, but it also introduced an artificial delineation between “discrete” and “continuous” gestures. This model was still “incorrect”—the boundary is an artefact of the system, not a feature of the practice—but it was a re-negotiated model of the musician’s practice that proved useful during periods of music composition.

However, more subtle representational dissonances surfaced as we developed this approach. For instance, even with this richer representation, the temporal envelope of the extracted feature sequences didn’t always match the gestural intention of the musician. During the final co-design session the musician suggested that a circular brushing

motion, suggestive of a sinusoidal low-frequency oscillator, could be looped to provide a backing track for him to play over. The feature envelope wasn't precise enough to support a stable loop. Waveshaping was introduced to shape the envelopes, but alone couldn't make up for the temporal discrepancies between gestural intention and the sonic result. We abandoned the looping idea (Figure 1), but kept the real-time continuous feature mapping for live performance. Importantly, music composition was part of the design process: the gestural friction surfaced the representational limit, but it was the goal of making music that supplied a reason to do something about it. Knowing when to push against representational limitations using technical practices and when to accept and switch into a compositional mindset emerged as an important form of judgement in our collaboration.

The productive dissonance created by the musician's practice brought representational limits into sharper relief. Without this, the original representational framework would likely have remained as is, not because its limitations weren't known, but because there was no compelling reason to push against them from within a purely technical perspective. It was the friction of the musician's gestural language that created the impetus to develop continuous gesture recognition.

5.1.2 Cymbal Sensor

A contrasting arc illustrates what happens when this balance tips too far toward technical aims. An IMU sensor attached to the cymbal, motivated by the performative possibilities of a visible, physical gesture, led to a spiral of interventions: each technical solution surfaced a new problem, pulling focus progressively further from musical practice. A visit from one of the researcher's supervisors, who had not been present for the preceding weeks of development, offered the outside perspective needed to see the spiral clearly. We decided to scrap the sensor ahead of the recording session. The core idea remained compelling, but the musical rationale had receded. There was no musical anchor to return to because the idea had originated technically rather than emerging from the musician's practice.

5.2 Navigating Performance-Demo Ambiguity

A persistent ambiguity ran through the collaboration about whether the primary goal was developing technology or making music. Despite the study's explicit motivation to support the musician's practice, the context of being a study in the first place had an effect of recentring the role of technology:

Researcher: My PhD is about introducing tech to drums. [But], I'm trying to be principled and [make it] about the music.

Musician: Nah nah, it's about the tech, but we can make it also about the music. Then that means the tech has been successful.

The tension between the dual objectives of the collaboration was present throughout, and was directly negotiated during a performance on the final day when a classification algorithm was producing incorrect results:

Musician: I guess it's a question of whether we, or you, rather, accepting that there's a couple of errant notes here and there and sometimes it doesn't

trigger. But the vibe is still there, isn't it? So on a musical level, the vibe is still working. It's just sort of more on the tech level.

Researcher: If it's musically there, then that's perfect.

Musician: Yeah, sweet. It's not going to be played any better, is it? Let's move on.

Musical merit was identified by the musician, but attention was pulled towards a concern for demonstrating the effectiveness of the technology. If this was indeed the only aim, then this performance would have been deemed unacceptable, but switching the frame of reference to one focused on musical merit led us to the decision to accept the performance. This is a form of knowing-when: knowing when to evaluate using a technical lens or an aesthetic one.

However, a more subtle ambiguity arose when the researcher further questioned the musician about their experience during the performance in light of the technical issues:

Researcher: I think it is interesting if you were distracted when you were playing. If you're thinking "I'm trying to get this work and it's not really doing what I want it to."

Musician: No, I mean, I'm sort of one of those not-perfectionist-perfectionists. I like the slight glitchiness anyway. So it's not like "this thing's not working, and it's not doing what I wanted to." It's basically doing what I want it to do. And then every so often, it doesn't quite catch. And then it throws in a note when it hasn't been asked to throw in a note. But I kind of like that too. It goes in with the sort of the world we live in, a glitchy ADHD world.

The musician's response might be read as expressing a genuine aesthetic preference for the "glitchiness" of the mapping. However, "glitchy" is an aesthetic that he hadn't mentioned during our study prior to this point—it emerged as an accommodation to the classifier failing, and was retroactively accepted as a positive. It is difficult to assess whether this was truly a novel aesthetic discovery, or whether it was a rationalisation of a limitation that was absorbed because of the context of the study and focus on technology.

Answering this is complex. The musician himself had framed the collaboration partly as a search for "a new impetus, a new way to approach the next bit of writing," which means some degree of aesthetic influence from the technology was not just acceptable but desired. Determining whether the aesthetic influence of the technology acts as a desirable creative catalyst or as a form of capture is central to the goals of our work. However, we find ourselves ill-equipped to answer this question in isolation from the musician's broader social and professional context. We elaborate on this in the discussion.

6 DISCUSSION

The findings in Section 5 surface three questions. First, what forms of tacit knowledge supported productive dissonance in our collaboration? Second, how do idiomatic representational assumptions become invisible without contact with practice—and what does productive dissonance

reveal about them? Third, and hardest to answer: when technology exerts aesthetic influence on a musician's practice, how do we know whether that influence is genuinely supportive or simply going uncontested? We address each in turn.

6.1 Knowing-When

Ryan (2009) described a knowing-when in electroacoustic improvisation as a tacit judgement about when to intervene with an algorithm, which is similar to a type of tacit knowledge the musician described in achieving sonic specificity during live performance. We identify an analogous knowing-when in DMI co-design, one that supported productive dissonance in multiple ways during our study-collaboration. Section 5.1.1 described knowing when to push against representational limits and when to accept them and compose within their constraints. Section 5.2 involved knowing when to evaluate using an aesthetic lens rather than a technical one. Switching that frame was what allowed the music to take precedence over concern for demonstrating the technology's effectiveness.

A third type of knowing-when concerned the researcher's own role within the collaboration. The researcher started the project with the intention of remaining as aesthetically neutral as possible, acting as a translator between the musician's artistic goals and the technical development. As the study progressed, the researcher noted that taking more aesthetic initiative had a positive effect in the collaboration, even if this meant resigning from that neutrality: "[the process is more effective] when I am also applying my musical judgements ... I'm starting a band with the musician."⁸ The musician reflected on this dynamic in kind: "I think things sort of come naturally once you start hearing the different sound worlds." This process, however, raises a deeper question: did the researcher's aesthetic influence expand to fill space that the musician's broader social context might otherwise have occupied? We return to this question below.

Knowing-when is not a design method that can be specified in advance; it is a form of judgement that develops through sustained collaborative practice. Central to this was recognising representational assumptions that had been normalised through repeated use.

6.2 Productive Dissonance

The development of continuous gestures in Section 5.1.1 sheds light on a type of representational normalisation present in our technical practice with augmented percussion instruments. Classification algorithms, including onset detection and neural network-based detectors, are a representational foundation of the entire toolkit. The use of these techniques as a starting point was pragmatic as much as anything else. Onset detection solves a real problem and was used in earlier work by the researcher to facilitate mapping to MIDI-based instruments. Similarly, classifying between head and rim hits, or continuous and discrete gestures, is a solution to a mapping problem where the output comprises discrete classes like MIDI pitch for creating melodies. These techniques are useful, and have become somewhat idiomatic in augmented percussion and DMIs more generally. However, their usage imposes an ideological commitment on a musical practice; the act of classification enacts the concept of separable classes upon a musical practice where none may have been present before.

⁸Taken from the researcher's study journal

A designer may be well aware of the false precision of these representations, but chooses to use them anyway out of pragmatism or necessity, particularly when interfacing with tools that assert their own representational commitments (e.g., MIDI). The risk is that through repeated use, representational convenience becomes idiomatic and eventually reified.

In our work, contact with practice, specifically a musician pulling out a wire brush and trying a mapping based on onset detection, surfaced the representational tension. Without this productive dissonance, there was insufficient force to destabilise the representational commitments of the researcher's original design. The musician helped push out of this. Of course, the new continuous gesture method itself is just another classifier that introduces an artificial boundary. Lossy representations may be unavoidable, but productive dissonance at least makes their assumptions visible—which is a precondition for designing within them consciously rather than perpetuating them.

In our study-collaboration, productive dissonance surfaced representational tensions, which led to the development of new representations that supported the musician's gestural language and composition of an album. However, during reflection on the musical results, a more subtle question arose related to the context of the laboratory.

6.3 The Hermetic Laboratory

The difficulty in answering whether the glitchiness described in Section 5.2 was a positive aesthetic influence or an example of capture points to an ecological risk when working with musicians in an academic setting. Waters (2021) describes musical instruments as not merely "things" to be studied, but also a constitutive part of a larger ecosystem including performers, instruments, and environments. With regard to musical aesthetics, Born (2010) argues that treating an aesthetic as being embedded in a social context is required to accurately portray a musical practice. Therefore, isolating a musician in a laboratory study, even with the goal of composing original music, risks severing the social dimension of a practice, potentially undermining a friction needed to distinguish genuine aesthetic development from accommodation to technological constraint. The musician wanted the technology to provide a new creative impetus, which means some degree of aesthetic influence was desired; but without his broader practice present to evaluate that influence, would the glitchiness have been accepted?

The cymbal sensor example in Section 5.1.2 exemplifies a related dynamic. A technical intervention spiral built up over the course of several sessions. The idea had clear musical applicability, but originated from a technical possibility and became increasingly driven by a technical aim instead of a clear musical need. Only once the supervisor visited and observed the situation were we able to break from the cycle of iterative prototyping. In both cases the study setting worked against our aims of creating a productive dissonance with technical aims. Technical aesthetics and institutional aims naturally filled the void left by weakened social dimensions. This suggests that working with a musician is a necessary but insufficient condition for sustaining a productive dissonance.

This is not to say that the social dimensions were completely absent: the musician posted clips of our sessions on his social media; shared conversations he had about

our work with collaborators; and we discussed public performances and a release of the album we created through traditional publishing avenues. Additionally, the researcher himself attended musical concerts that the musician was playing at, which led to impromptu discussions related to their practice that were situated in a broader context. All these are suggestive of a porousness of the laboratory. Nonetheless, sustained engagement with a broader network was limited, and mostly restricted to the conclusion of the study. Our findings suggest deeper and sustained engagement with these dimensions is required to answer the complex, but important, questions related to technical impact on musical practice. Situating design within an active project with real social commitments, and ones that exist outside the regular streams of academic art production, is one approach that would take this engagement further—and one that reflective practice-based methods are well-positioned to investigate further.

7 CONCLUSIONS

In this work, we took up a provocation from Norman et al. (2025): “rescue performance from the demo.” This highlights a central tension in designing musical performances where technology development is a significant concern—is the performance about the music, or about demonstrating the affordances of the technology? When the latter occurs under the guise of composing new music, what does this reflect about the impact of a DMI on the practice it is incorporated into?

We propose that this may be indicative of technological capture, where the constraints of a DMI overwhelmingly influence a musical practice, trapping it in a prescriptive aesthetic range. Following recent critical work on mapping, we conducted a long-term collaboration with a professional percussionist, engaging his aesthetic as a method for creating productive dissonance with the constraints of a percussion gesture mapping toolkit—a guiding force for technical practice. This was partially successful, resulting in a toolkit with a novel method for mapping continuous percussive gestures using RNNs, and a ten-track album recorded with the musician, presented as an annotated portfolio.

Through a reflective practice-based approach, we surfaced a form of tacit knowledge we relate to knowing-when (Ryan, 2009), and identified a limitation of the academic co-design setting: isolating the social dimensions of a musician’s practice risks undermining the aesthetic weight they bring to the work. This surfaces an irresolvable tension absent a broader social context: how do we know whether the aesthetic introduced by the technology is truly supporting the musician’s aims, or whether the conditions of the study lead to its unquestioned acceptance? We put this forward as an open question, and suggest that a deep engagement with the social dimensions of a practice is a promising way to approach an answer.

AUTHOR DECLARATIONS

The study was deemed low-risk and received ethical approval from the Queen Mary University of London EECS Devolved School Research Ethics Committee. All AI models used during this study were trained using data recorded by the authors within the study. Claude Sonnet 4.6/Opus 5 was used to support the writing of this paper as a tool to edit text originally written by the authors. All ideas

and technical aspects were contributed by the authors and collaborators.

ACKNOWLEDGEMENTS

The authors would like to extend their deep gratitude to Jem Doulton for his commitment to the co-design process and for contributing his knowledge and experience. Thank you to Carson Gant for editing the performance videos and for mixing the audio, and to Alex Bonney for mastering the tracks. Thank you to Rodrigo Constanzo, whose DataKnot toolkit provided the foundation for this work, for the ongoing collaboration and continued conversations on drum performance and technology.

REFERENCES

- Bang, T. G., Fdili Alaoui, S., and Schwartz, E. (2023). Designing in Conversation With Dance Practice. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI ’23, pages 1–16.
- Borgdorff, H. (2010). The Production of Knowledge in Artistic Research. In Biggs, M. and Karlsson, H., editors, *The Routledge companion to research in the arts*. Routledge, New York.
- Born, G. (2010). The Social and the Aesthetic: For a Post-Bourdieuian Theory of Cultural Production. *Cultural Sociology*, 4(2):171–208.
- Bowers, J. and Green, O. (2018). All the Noises: Hijacking Listening Machines for Performative Research. In *Proceedings of the International Conference on New Interfaces for Musical Expression*.
- Braun, V. and Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4):589–597.
- Bulley, J. and Şahin, O. (2021). Practice Research - Report 1: What is practice research? Report, PRAG-UK, London, UK.
- Cho, K., van Merriënboer, B., Bahdanau, D., and Bengio, Y. (2014). On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. In Wu, D., Carpuat, M., Carreras, X., and Vecchi, E. M., editors, *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*, pages 103–111.
- Chowdhury, J. (2021). RTNeural: Fast Neural Inference for Real-Time Systems. arXiv:2106.03037 [eess].
- Dal Bó, E. (2006). Regulatory Capture: A Review. *Oxford Review of Economic Policy*, 22(2):203–225.
- Devendorf, L., Andersen, K., and Kelliher, A. (2020). Making Design Memoirs: Understanding and Honoring Difficult Experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI’20, pages 1–12.
- Engel, J., Hantrakul, L. H., Gu, C., and Roberts, A. (2020). DDSP: Differentiable digital signal processing. In *International Conference on Learning Representations*.
- Gaver, B. and Bowers, J. (2012). Annotated portfolios. *Interactions*, 19(4):40–49.

- Gaver, W. (2012). What Should We Expect From Research Through Design? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 937–946, New York, USA. Association for Computing Machinery.
- Hunt, A., Wanderley, M. M., and Paradis, M. (2003). The Importance of Parameter Mapping in Electronic Instrument Design. *Journal of New Music Research*, 32(4):429–440.
- Jack, R. H., Stockman, T., and McPherson, A. (2017). Rich gesture, reduced control: the influence of constrained mappings on performance technique. In *Proceedings of the 4th International Conference on Movement Computing*.
- Johnston, A. and Ferguson, S. (2016). Practice-Based Research and New Interfaces for Musical Expression. *Leonardo*, 49(1):71–71.
- Jourdan, T. and Caramiaux, B. (2023). Machine Learning for Musical Expression: A Systematic Literature Review. In *New Interfaces for Musical Expression*.
- Lee, D. D. and Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791.
- McPherson, A., Morrison, L., Davison, M., and Wanderley, M. M. (2024). On mapping as a technoscientific practice in digital musical instruments. *Journal of New Music Research*, 53(1-2):110–125.
- Norman, S. J., Stapleton, P., and Bowers, J. (2025). NIME: A Mis-User’s Manual. In *New Interfaces for Musical Expression*.
- Pelinski, T., McPherson, A., and Fiebrink, R. (2025). Ways of knowing, ways of writing: technical practice research in new musical instrument design. *Journal of New Music Research*, pages 1–14.
- Pelinski, T., Melbye, A. P., and McPherson, A. (2026). When AI doesn’t sound like AI: Negotiating aesthetic expectations in technology-mediated musical practice. *Organised Sound*, pages 1–11.
- Petitmengin, C. (2006). Describing one’s subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, 5(3-4):229–269.
- Reed, C. N., Benito, A. L., Caspe, F., and McPherson, A. P. (2024). Shifting Ambiguity, Collapsing Indeterminacy: Designing with Data as Baradian Apparatus. *ACM Trans. Comput.-Hum. Interact.*, 31(6):73:1–73:41.
- Rode, J. A. (2011). A theoretical agenda for feminist HCI. *Interacting with Computers*, 23(5):393–400.
- Ryan, J. (2009). Knowing When, Music and Machines IX, seminar convened by B. Hogg, S. J. Norman, Culture Lab, Newcastle University. <https://www.youtube.com/watch?v=pO3h9LMj9mY> (Accessed 2026-04-20).
- Saitis, C., Reed, C. N., Noel-Hirst, A., Lepri, G., and McPherson, A. (2025). (De)Constructing Timbre at NIME: Reflecting on Technology and Aesthetic Entanglements in Instrument Design. In *New Interfaces for Musical Expression*.
- Sengers, P., Boehner, K., David, S., and Kaye, J. J. (2005). Reflective design. In *Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility*.
- Shier, J., Constanzo, R., Saitis, C., Robertson, A., and McPherson, A. (2025). Designing Percussive Timbre Remappings: Negotiating Audio Representations and Evolving Parameter Spaces. In *International Conference on New Interfaces for Musical Expression*.
- Shier, J., Saitis, C., Robertson, A., and McPherson, A. (2024). Real-time Timbre Remapping with Differentiable DSP. In *International Conference on New Interfaces for Musical Expression*.
- Smaragdis, P. and Brown, J. (2003). Non-negative matrix factorization for polyphonic music transcription. In *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (IEEE Cat. No.03TH8684)*, pages 177–180, New Paltz, NY, USA. IEEE.
- Spatz, B. (2015). *What a Body Can Do: Technique as Knowledge, Practice as Research*. Routledge, New York, USA.
- Stowell, D. and Plumbley, M. D. (2010). Timbre remapping through a regression-tree technique. In *Sound and Music Computing (SMC)*.
- Tremblay, P. A. and Schwarz, D. (2010). Surfing the Waves: Live Audio Mosaicing of an Electric Bass Performance as a Corpus Browsing Interface. In *Proceedings of the International Conference on New Interfaces for Musical Expression*.
- Tuuri, K., Parviainen, J., and Pirhonen, A. (2017). Who Controls Who? Embodied Control Within Human–Technology Choreographies. *Interacting with Computers*, 29(4):494–511.
- Waters, S. (2021). The entanglements which make instruments musical: Rediscovering sociality. *Journal of New Music Research*, 50(2):133–146.
- Weatherhead, A. (2025). Drum Triggers: Discussing Pros and Cons. *Drummerworld*. <https://www.drummerworld.com/articles/news/drum-triggers-pros-and-cons-live-performance/> (Accessed 2026-04-21).