



PDF Download
3771594.3771600.pdf
16 February 2026
Total Citations: 0
Total Downloads: 145

Latest updates: <https://dl.acm.org/doi/10.1145/3771594.3771600>

RESEARCH-ARTICLE

Meaningful Interactions in Human-AI Musicking

OLIVER MILES, University of Nottingham, Nottingham, Nottinghamshire, U.K.

ADRIAN HAZZARD, University of Nottingham, Nottingham, Nottinghamshire, U.K.

SOLOMIYA MOROZ, University of Nottingham, Nottingham, Nottinghamshire, U.K.

LAURA BISHOP, University of Oslo, Oslo, Oslo, Norway

CRAIG VEAR, University of Nottingham, Nottingham, Nottinghamshire, U.K.

Open Access Support provided by:

University of Nottingham

University of Oslo

Published: 30 June 2025

[Citation in BibTeX format](#)

AM '25: 20th International Audio Mostly Conference

June 30 - July 4, 2025

Coimbra, Portugal

Meaningful Interactions in Human-AI Musicking

Oliver Miles
School of Computer Science
University of Nottingham
Nottingham, United Kingdom
Oliver.Miles1@nottingham.ac.uk

Adrian Hazzard
School of Computer Science
University of Nottingham
Nottingham, United Kingdom
Adrian.Hazzard@nottingham.ac.uk

Solomiya Moroz
Department of Music
University of Nottingham
Nottingham, United Kingdom
solomiya.moroz@nottingham.ac.uk

Laura Bishop
RITMO Centre for Interdisciplinary
Studies in Rhythm, Time and Motion
University of Oslo
Oslo, Norway
laura.bishop@imv.uio.no

Craig Vear
School of Computer Science and the
Department of Music
University of Nottingham
Nottingham, United Kingdom
Craig.Vear@nottingham.ac.uk

Abstract

In this paper we discuss our latest iteration of *RAMI* (the Robotic Arm for Musicking Inclusively) as prototype digital score. We deploy *RAMI* as interactive technology for the purpose of improvisation ensemble work, documenting the narrative journey of a single elite musician—*M*—as they engage with its nascent qualities over three extended workshop sessions. This study builds in part on the musical togetherness model of human-human interaction [4], and necessarily draws adjacent concepts in robotics, human computer interaction (HCI), and digital score. Recording musicking interactions with *RAMI*, we elicit and report on *M*'s thematic perceptions of instances of: familiarization; engagement strategy; decision making; connection, agency, and ensemble; assertions of schemas of *RAMI*; and moments of enjoying *RAMI*. Beyond the rich, narrative insight that these six themes offer, we critically assess these as variably (un)certain in nature; as particularly certain when asserting two schemas—*RAMI* as a dancer, and *RAMI* as a score, and as indicative of an underlying continuum of musician perception that recognises *RAMI* as an AI tool for musicking, on the one hand, and on the other hand, as an AI agent for embodied collaboration.

CCS Concepts

• **Human-centered computing** → **Interaction design theory, concepts and paradigms; Empirical studies in collaborative and social computing.**

Keywords

Embodied AI, musicking, meaning making, creativity, digital score

ACM Reference Format:

Oliver Miles, Adrian Hazzard, Solomiya Moroz, Laura Bishop, and Craig Vear. 2025. Meaningful Interactions in Human-AI Musicking. In *20th International Audio Mostly Conference (AM '25), June 30–July 04, 2025, Coimbra, Portugal*. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3771594.3771600>



This work is licensed under a Creative Commons Attribution 4.0 International License. *AM '25, Coimbra, Portugal*

© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2065-9/25/06
<https://doi.org/10.1145/3771594.3771600>

1 Introduction

The creative practice of musical improvisation is a challenging but potentially rewarding context in which to apply artificial intelligence (AI) and robotics. In this work we report on a thematic analysis of musician *M*'s reflections on a human-AI collaboration involving embodied musicking (the act of doing music [32]). We focus on their journey of interaction with a digital score [35] called *RAMI* ("Robotic Arm for Musicking Inclusively") and their nascent perception of *RAMI*'s provocative, responsive, and collaborative affordances during human-AI improvisation. We draw motivation from prior work in this area, in particular *Jess+* [39], which charted how a musical trio improvised with a digital score embodied as an AI robotic arm and developed meaningful creative relationships with it that suggested a shared experience of musical togetherness.

Our work is motivated by broad questions about how processes of creative meaning-making and musical togetherness, which have been well-studied in the context of human-human collaborative musicking, unfold during a musician's interaction with *RAMI*. Our findings suggest that *M* evaluates *RAMI* through the schema of "digital score", suggesting its function as largely a tool for collaborative musicking. However, we also find that *M* uses the schema of "dancer", suggesting that *RAMI* offers embodied affordances akin to human interaction. Both schemas have potential to be viewed as instantiations of the musical togetherness paradigm [4].

Informed largely by the extent to which perceived togetherness factored in *M*'s experience, we discuss our findings in terms of expectations of embodied AI, the creation of performative variety through unpredictability as a resource for musicking, and the acceptable bounds or parameters of an AI improvising agent contained within a digital score concept. Moreover, we suggest how these affordances of *RAMI* might feed back into the AI loop as an ambition for future work.

2 Related Work

Our related work covers meaning-making in music ensemble playing, robotics, and human machine collaboration, embodied AI, and concepts of musical togetherness. From these domains we draw out implications for our work and establish the core elements of *RAMI*'s system design.

2.1 Meaning Making in Music Ensembles

Music-making is inherently social; the majority of musical performances occur within ensembles, and even solo performances are often shaped by the presence, real or imagined, of an audience. Music-making is also an act of creativity. Across diverse musical cultures, improvisation serves as a key mechanism for generating new material. This can be observed in the jam sessions that contribute to pop songwriting, the spontaneous solos in jazz, the improvisatory practices in folk music, and the extemporization that precedes orchestral composition in classical music. Creativity emerges when play evolves into the invention of novel and meaningful musical expressions [13]. A central aim underlying this project involves understanding how *RAMI* facilitates meaning-making among musicians during the act of music-making.

For the purposes of this research, we conceptualize musical creativity and meaning-making from the perspective of the musicians themselves. This perspective aligns with Christopher Small's notion of musicking, which he defines as "to music is to take part." [32]. Small elaborates that participation in musicking can occur in various capacities, including performing, composing, listening, and even dancing. Crucially, he argues that meaning is formed through the relationships established within the act of musicking. In the context of digital scores, this entails forging connections not only with fellow musicians but also with agents, sounds, spaces, and presences encountered within the musicking environment [32].

RAMI offers a novel platform for musicking. Musicians who play with it for the first time must undergo a process of sense-making in which they figure out what it is, how it works, and how it might be integrated into their creative practice. Sense-making is a theoretical concept that describes the active, dynamic, and iterative process through which people acquire understanding about the world [12]. Part of the process involves accommodating incoming perceptions into schemas (cognitive frameworks) that are already in memory [41]. In this work, we report on a musician's experiences with *RAMI* over the course of several sessions and uncover some of the strategies that they used to "make sense" of *RAMI*'s functions and behaviour. We use the concept of sense-making as a framework to understand the musician's reflections.

2.2 Robotics and Human-Machine Collaboration

We are increasingly witnessing deeper intersections between human musicians and autonomous machines. This trend is reflected in the broader discourse within the HCI community, where topics of human-machine collaboration are central. Research in this field examines various concerns, including the negotiation of roles, autonomy, and control [1, 3, 17], co-creativity [11, 21, 36] sense-making [18, 27], and transparency, accountability, and trust [2, 30]. The relationship between humans and autonomous systems is evolving, challenging traditional notions of interaction in HCI.

This project advances research at the intersection of music co-creativity and robotics, drawing on prior work in computational creativity [22], AI and music [25], and robotic musicianship [7, 19, 40]. As this field continues to evolve, numerous approaches are being developed to enable human-robot music interaction, particularly focusing on real-time, in-the-loop collaboration. For instance, the

cooperative AI system at the heart of *In A Silent Way* [23] utilizes performance data to generate real-time sound and emoticons, fostering a sense of trust between human musicians and the AI. Furthermore, McCormack [24] proposes a framework for optimizing human-AI creative collaboration, which has implications for extending human-robotic musicking. Our research builds upon these foundations by focusing on the embodied relationships between musicians, robotic agents, and the co-creative flow of musical performance. The long-term goal is to explore how such interactions can transform human musical creativity.

2.3 Embodied AI and Human-Computer Integration

A core aspect of our investigation lies in understanding embodied AI, which directly informs the design innovations presented in this research. Embodied AI encompasses robotics, where AI operates through a physical body, but in the realm of musicking, embodiment extends beyond physical form. It encompasses the relational and communicative dynamics between musicians and their instruments, primarily mediated through sound and performative presence.

Vear [36] highlights that embodiment in music refers to the process by which musicians internalize their sound, integrating it into their bodily sense of being. Rather than merely producing sound externally, musicians experience musicking as a deeply immersive state in which they become the sound they create. This process is reciprocal; musicians not only shape their own sonic expressions but also attune themselves to the sounds of others, forging a shared presence through sound. This dynamic interaction resembles a dance—one of sensing, responding, playing, and engaging with others in an intricate and fluid exchange.

To incorporate AI and robotics into this "dance", these technologies must be designed from an embodied perspective. Our project adopts a definition of embodied AI as an "intelligent agent whose operational behaviour is determined by percepts interacting to the dynamic situation within which it operates" [26]. This definition rests on two key principles: (a) AI is not confined to a purely cognitive model, and (b) meaning-making is best understood through the lens of embodied cognition. This perspective builds upon Rodney Brooks's work on behavioral AI [8], particularly his concept of "coping machines". However, rather than designing AI to navigate generic dynamic environments, we focus on an environment where the principal mode of interaction and existence is music.

2.4 Contingencies of Musical Togetherness

Experiencing musical togetherness, defined as a sense of musical connection linked to feelings of social connection and pleasure, is part of what makes collaborative music-making rewarding for musicians [4]. When people play music together, they draw on shared knowledge and the affordances of real-time interaction to co-create musical gestures [29]. This gives them the sense of sharing and realizing musical ideas together.

Interactive computer musical systems, such as those utilising algorithmic and AI programming, are often designed with a goal of capturing liveness or the presence of other musicians [5]. George Lewis, a pioneer musician in computer-music interactions and creator of *Voyager* system, characterised it as designing a system that

he will not only “play” but will also “play with him” [20]. Previous works that support this approach also highlight the importance of feeling the presence of an interacting partner, even in a duet performance in virtual reality, as a satisfying aspect of a musical collaboration [34].

Most research on musical togetherness concerns musicians’ live interaction with each other. People take more creative risks when they can communicate multimodally (i.e., both hear and see each other) and when they have prior experience improvising together [14], suggesting that these factors ease coordination of musical ideas. Sometimes strengthening togetherness in the long term requires perturbing coordination in the moment, in a way that scaffolds the emergence of new creative ideas [28]. Feelings of togetherness reduce when musicians must interact remotely [16] or when they play with an artificial partner that does not replicate the breadth of human creativity [10, 34]. The study of *Jess+* [39] suggested that human musicians can integrate an interactive digital score into the social dynamic of their group. In the current study, we asked whether a solo musician might similarly establish a social connection to *RAMI* and experience a sense of musical togetherness with it.

When it comes to harnessing perceptions of meaningful moments in heavily networked interactions, Skarth-Hayley and Green-smith develop a novel sonification system that speaks to a contextual challenge of integrating both “live and retrospective” meaning making [31]. This challenge is similarly evident in our work: while *RAMI* does not intentionally use sound to communicate its own musical ideas, it does provoke questions of latency between the live, discrete, and sensed moments of musician response, and the comparatively retrospective, conversational qualification of interaction which forms the analytic focus of this piece.

2.5 *RAMI* as Digital Score: System Design

Although this study did not aim to validate *RAMI* as a digital score, it is important to define what a digital score means in this context. Vear defines a digital score as “a communications interface of musical ideas between musicians utilizing the creative potential of digital technology” [35]. In this sense, *RAMI* is not an open system for improvised movement but is instead embedded with specific behaviors and beliefs that shape its interactivity. A “musical idea” here is encapsulated in an interactive AI, whose responsive behaviors are expressed through the gestural movements of the arm. The musician’s role in realizing this digital score is to respond to these arm gestures by creating sounds on their instrument.

While this abstraction of a score as a platform to communicate a musical idea challenges traditional score norms, a closer examination of Vear’s definition of a digital score reveals that *RAMI* fulfills both its purpose and function:

- **Purpose:** The digital score serves as a technically mediated communication interface that connects the creativity of one musician, another musician, and the listener’s creative mind.
- **Function:** The digital score represents the ideas within the musician’s mind using digital technology, allowing those ideas to be translated into sound during performance through the technique and creative interpretation of another musician (human or machine).

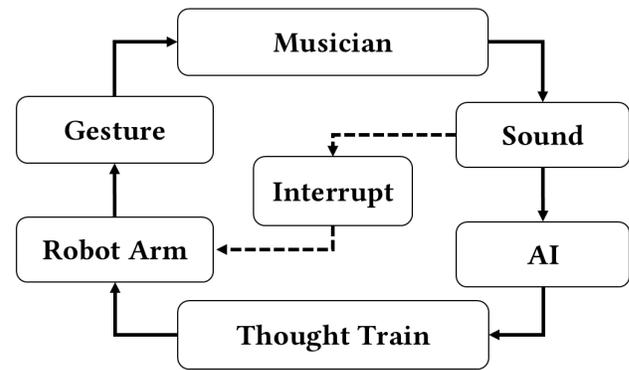


Figure 1: *RAMI* System Interaction

In the case of *RAMI*, the musical idea extends from the previous *Jess+* experiment [38], which was embedded with gestures derived from two 1960s graphic scores: Cardew’s *Treatise* and Wolff’s *For 1, 2, or 3 players*. These gestures were slightly modified to allow for 360-degree arm motion. How *RAMI* moves, how it reacts, and even why it reacts in certain ways are all shaped by the belief systems embedded within it, as its primary purpose is to function as a digital score.

RAMI is designed as an embodied AI robotic system for real-time shared musical creativity. This is an in-the-loop system, where a musician realises the work embedded in *RAMI* through improvised response. *RAMI* is not a sonic contributor, but rather the gestures made by the robotic arm are a manifestation of the embedded musical idea encapsulated by the digital score’s AI design. The musician’s response is left open to the interpretation by the individual, but nonetheless *RAMI* was conceived as an integral actor in the ensemble’s collaboration and in the realisation of the digital score.

The system interaction design in Figure 1 illustrates how *RAMI* is listening to the audio input of the musician, and then responds with a physical gesture, which the musician then takes as abstract music material to make a response. The system design and functionality of *RAMI* is based on *Jess+*, and is discussed in detail in [38]. In short, starting at the top right side of the diagram, direct [Sound] from M is processed by the AI Factory [AI] (a collection of 5 neural networks, predicting in real-time), which outputs a constant stream of predicted data. A secondary module selects one of these streams for a short time (between 2-8 seconds), thereby producing a [Thought train] value. This is then used to determine a robotic gesture selected from a pre-defined/ composed library of gestures inspired by the Wolff and Cardew pieces (mentioned above), which the robot arm uses to move, or interrupt an ongoing move. Those gestures are then interpreted by the musician who makes sound as part of their ongoing improvisation which is picked up by the microphone in [Sound] thereby closing the loop. That main loop can also be bypassed when the sound is too loud, interrupting the robot’s gestures and moving on to a new main cycle (called the [Interrupt] response).

The AI is a combination of symbolic algorithms and machine learning neural networks trained on embodied musicking data,

discussed in detail in [37]. As such, it is designed to sense the outside [musicking] world through the amplitude of the sound waves, not the frequency of the pitches, and to interpret this through the interpretive qualities embedded in the neural network training (discussed in more detail in [38]). The data that stream from the AI factory represents a busy mind with an abundance of trains of thought. This solution was chosen as it had been proven to stimulate novel musical experiences in the previous experiments such as *Jess+* (these are discussed in detail in [39]).

3 Methodology

The work charted here is a case study of a professional improviser, referred to as *M*, collaborating with *RAMI* over three study sessions. In this paper, we report and distill insights from the subjective experiences of *M* that unfolded during these sessions. This study was part of a larger project that aimed to develop a methodology for exploring musicians' meaning-making during interaction with *RAMI* based on three sources of data: subjective interviews with the musicians, their real-time physiological responses, and *RAMI*'s data logs. Although this paper is concerned only with *M*'s subjective experiences, it is nonetheless important to detail the broader study procedure and setup to illustrate the conditions that will have shaped our participant's experience.

The research team consisted of five researchers, who all brought a range of expertise to the study. Among them, four members are experienced musicians in their own right.

3.1 Study Set-up

We conducted three study sessions over a 5 month period with a professional improviser (*M*) (See 4.1 for a description of *M*'s background.) Each workshop followed a similar protocol, although adjustments were made as part of our methodological development. These will be highlighted where relevant to the focus of this paper. The workshops were held at the host institution, in a flexible black-box lab area. The lab area was configured for this study as illustrated in Figure 2. The study setup consisted of a central performance area (where *M* sat facing *RAMI*'s robotic arm), tables for system computers and members of the research team, video cameras to capture the improvised performances, and a large screen to conduct post-performance video protocol. Some adjustments were made to the placement of the equipment in each workshop, without affecting the protocol.

3.2 Study Procedure

In the following, we outline the study procedure for the *RAMI* workshops (as illustrated in Figure 3). The duration of each workshop was approximately 4-5 hours and consisted of four stages.

- **Stage 1 - Orientation:** Workshops 1 and 2 involved an initial framing discussion centered on the question, "What makes for a meaningful ensemble experience from your perspective?" This allowed *M* to reflect on their practice and identify key topics for later discussions on improvisation. In Workshop 1, key points were noted on flip chart paper. In Workshop 2, the discussion revisited these points, exploring changes and additional thoughts from *M*. This discussion

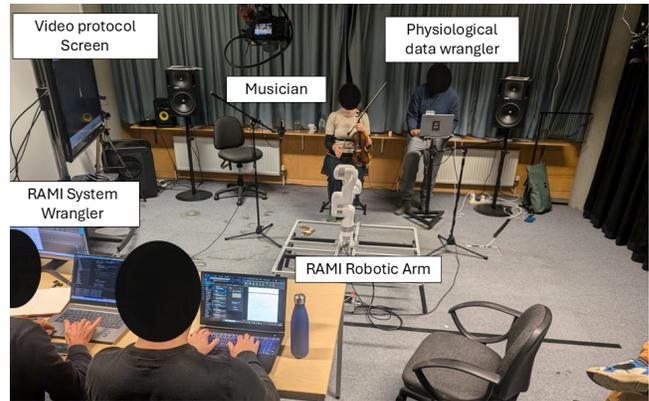


Figure 2: Setup of study area for *RAMI* workshops

was recorded and further annotations were made. The framing was considered complete by Workshop 3, where the flip charts were displayed for reference but no new discussion was conducted.

- **Stage 2 - Performance:** Stage 2 focused on *M*'s improvisations, starting with solo improvisations where physiological data was captured at varying intensities to establish a baseline for qualitative measuring. These solo performances were followed by collaborative improvisations with *RAMI*, which took place twice in Workshop 1, three times in Workshop 2, and twice in Workshop 3.
- **Stage 3 - Video Recall Protocol:** This involved *M* reflecting on each improvisation immediately after its completion. *M* watched a recording of the improvisation with a researcher, identifying "moments of interest" for discussion. The researcher occasionally paused the video to prompt discussion, sometimes relating to Stage 1 topics or asking about *M*'s thoughts and feelings. In the first workshop, a two-stage process was tested, where *M* marked points of interest during the first viewing and discussed them during the second. This marking step was abandoned after the first workshop as it did not add value and prolonged the process unnecessarily.
- **Stage 4 - Summative reflection:** At the end of each workshop, a summative semi-structured interview was conducted. These interviews began with the question, "Tell us about today, how has it been?" They focused on *M*'s thoughts and feelings about improvising with *RAMI*, revisited topics from the video recall protocols, compared the *RAMI* experience to previous ensemble practice, and explored changes in *M*'s approach over the course of the three workshops.
- **Stage 5 - Final Thoughts:** In addition to the four-stage workshop structure, a final semi-structured interview was conducted with *M* three months after the last workshop. The delay allowed *M* to reflect with a more detached perspective. This interview also provided an opportunity for the research team to present a work-in-progress analysis of the workshops and gather *M*'s thoughts and clarifications.

3.2.1 *Data Collection Summary.* The data were collected in 4 stages comprising framing discussions for each session; timestamped

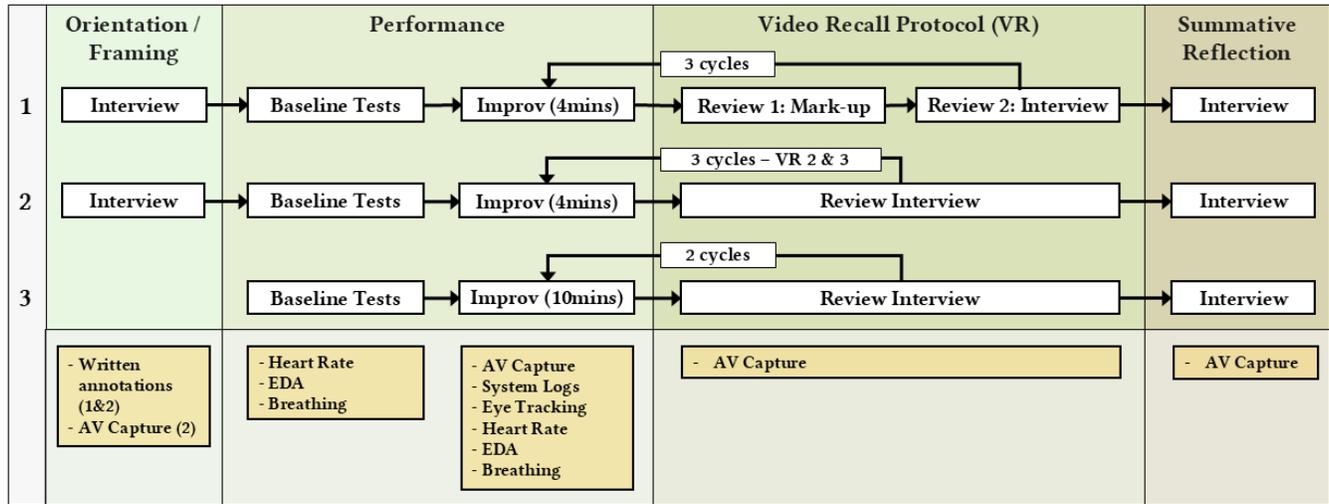


Figure 3: Overview of RAMI Methodology

video protocol discussions in which *M* provided rich, elaborative commentary on their musicking; post-performance summary discussions, and a semi-structured follow-up interview conducted three months after the final performance.

3.2.2 *Data Analysis.* In this paper we focus on the presentation of a reflexive thematic analysis of *M*'s transcribed reflections on their interactions with *RAMI*, predominantly inspired by the well regarded 6-step approach [6]. In an attempt to provide a worked example as others have done [9], we offer the following:

- (1) An initial *data familiarization* stage, beginning at the point of organizing transcripts as belonging to one of three contexts of production; the framing discussion, the video protocol (VP) semi-structured interviews, and the summative reflections (see Figure 3).
- (2) We engaged in a collective but independent *initial coding* of data. This involved formalizing verbatim extracts of (*M*'s) reflections (particularly in the VP datasets) as well as clear instances of nascent themes that were immediately apparent from the outset.
- (3) We then engaged in a *theme search*, individually proposing comprehensive sets of themes that described in detail the nature of interactions with *RAMI*.
- (4) At the point of *reviewing themes*, we collaborated as a team to compare and contrast themes, assessing their apparent significance in terms of their universality, as well as their linkage to the same sessions and extracts.
- (5) We *mapped and named* 6 top-level themes, which, through a process of discussion, build toward the following narrative of human-AI interaction that, for *M*, was descriptive of perceptions of meaningful interaction with *RAMI*.

4 Findings

In the following, we present our findings. This begins with a background description of our musician (*M*), illustrating their practice

and approach to ensemble improvisation. The background description draws on material from the framing, summative and final interview discussions. This is followed by a description of the 6 themes identified below from our thematic analysis, which illustrate *M*'s journey with *RAMI*. These are elaborated on in sequence in Fig. 5 and attributed to specific sessions, improvisations, and video recall protocol extracts "(Sn.In.VPn)".

(1) *Familiarization* - the process of getting to know *RAMI* as an improvising partner; (2) *Engagement strategies* - specific approaches that drew on musical expertise, processes and artifacts of preparation, embodied practices of mirroring, mimicking, and matching *RAMI*'s movement, and responding to *RAMI*'s sound; (3) *Decision making and challenging* - conscious, effortful moments of taking responsibility, communicating musical ideas, getting (un)stuck in loops, motivation, (in)security, and developing musical themes, form, and structure; (4) *Connection and Agency in Ensemble Interaction* - moments of feeling (dis)connected, leading/following *RAMI*, shifting senses of agency, and the role of chance and coincidence; (5) *Schemas of RAMI* - musician led characterizations of *RAMI* as, variably, a composer, a performer, a dancer, a teacher, a provocateur, an accompanying artist, a score, and a restriction; and (6) *Enjoying RAMI* - moments of explicit satisfaction, contingent on relationships between changes to dynamics (speed and perceived sound), moments of synchronicity, and precision. Importantly, these should be read as human perspectives on the nature of interaction with *RAMI* that are, on the one hand, descriptive of *M*'s authentic, subjective experience of musicking, while at the same time emerging affordances of *RAMI* as a prototype.

4.1 Musician's (*M*) Background and Practice

M is a classically trained violinist and university graduate who works as a freelance musician and teacher. While they still perform traditional classical repertoire in various ensembles, their primary focus is free improvisation in both live performances and recordings.

This practice includes solo and ensemble settings, often incorporating diverse stimuli like dancers, DJs, text, graphic scores, artwork, and museum objects. In their human-human improvisation they employ a number of strategies that help them to be "fully present in the moment" and make decisions as they arise. This approach allows *M* to "always strive to evolve" and foster collective creativity. While *M* values spontaneity, they occasionally plan ahead by deliberately limiting material or predetermining temporal structures. In the course of their improvisation, they are open to responding to any and every unexpected stimuli that might arise, for example, when working with dancers, the sounds of their feet on the ground or breathing are as valid as their physical gestures. On this point, they gave a further example: "I've played [with] a DJ. [...] Sometimes I respond to her energy, [...] like the sound she makes with her nails while doing the decks." *M* also embraces disruption as a core element of their practice, such as using "preparations" on the violin, which alters its sound and response: "Putting preparations on the string disrupts what I'm used to as a classical musician. [...] It invites me to discover something new, [...] rather than relying on what I know." Accidental events, like hitting the bow against the wall, can also become key elements in *M*'s improvisation, as they embrace the unexpected to shape their performance.

The framing discussion conducted as the first step in study session 1, which was also reflected on at the start of study session 2, led with the question, what makes for a meaningful ensemble performance from your perspective? *M* noted that a successful and meaningful ensemble improvisation is one where you are collectively "taken out" on a journey through a flow state. For this to occur, there needs to be a "shared sound and aesthetics [...] and contextual understanding" which typically emerges organically through trial and error and "heightened listening". All players are responsible for the sensitive development of this shared understanding, which can help establish a bond of trust. Trust, in turn, bestows "permission" to take risks, to "push and pull", to "converse and interpret", which promotes creativity. Furthermore, *M* stated the need to "have a frame" in place, to "image the improvisation already composed". By way of an example, *M* noted that in previous settings the ensemble would—when possible—take time to explore what a combined sound world would be (i.e., develop and establish a frame) prior to the performance or recording.

4.2 Familiarization

Familiarization was an important feature of interaction for *M*, who from the outset described a process of non-verbal conversation: As a first encounter it was important to both "get to know" *RAMI* in terms of the kind of output it is capable of, simply by "trying different things" (S1.I1.2VP). As they progressed through performances, *M* found that they were increasingly "pre-empting" the actions and outputs of *RAMI*, and that this in turn felt "more communicative" (S1.I2.VP). In the final session, we start to see what this familiarity is contingent on. *M* is specifically concerned with the quality of speed and matching the speed of *RAMI*:

"I am thinking ahead [...] I'm anticipating that I've kept the same, more or less speed, and I'm noticing there's quite a lot of [...] quick movements" (S3.I1.VP)

The sense of being familiar with *RAMI* in the moment was also contingent on a growing personal inventory of first encounter[s]. In one sense, familiarization was driven by our musician from the start through experimental practice, but increasingly, it is suggested as an affordance of *RAMI* when it is perceived as sensitive to these early experiments. In this sense, *RAMI* is perceptive and learning, and to some extent an increasingly "comfortable" improvisation partner:

"Well compared to my very first encounter [...] there is some kind of connection because I've felt that [...] it was learning a little bit of you know what I was doing, but yeah I'm a bit more comfortable with my decision to [...] present this material" (S3.I1.VP).

Nevertheless, *M* reiterated the centrality of sound as a stimulus, concluding the final session by emphasizing the importance of "getting to know these kind of sounds that it's making" (S3.I1.VP). Familiarization then, appeared contingent on not only a growing sense of *RAMI* as an improvisation agent, but also a growing familiarity with its output.

4.3 Engagement Strategies

We note several formalized engagement strategies that helped to maintain this state of familiarity between *M* and *RAMI*, which draw heavily on *M*'s status as an expert musician and improviser. Because of this expertise, successful engagement could be seen as contingent on specific strategies, such as: the use of instrumental preparation; the act of mirroring, mimicking, and matching physical movements of *RAMI*; and the utilization of *RAMI*'s natural sounds.

In terms of preparation, which in this case refers to the process of preparing material that is placed or attached to the violin to disrupt or modify its timbre or response, this was tangible through the use of "a little 3D printed disc that a fellow composer's given me". This afforded prior engagement based on instability, uncertainty, and ostensibly matched the perception of *RAMI* as itself unstable and uncertain. As well as a reaction to and matching of *RAMI*'s uncertain nature, preparation was an active attempt to try to prompt *RAMI* and provoke it to respond.

Another strategy was the use of mirroring, mimicking, and matching *RAMI*'s movement. *M* focused on complementing *RAMI*'s physical gestures with their own, and there was often a sense that this shaped the resulting sound. This strategy progressed throughout the performances to some quite specific moments, where *M* was able to articulate precisely the "small, jerky movements" that most inspired their response.

"I suppose this point here when I'm trying not to respond as I have done in the past [...] I'm just narrowing my left hand [...] Not particularly focused in any kind of pitch. It's just this sort of [...] movement I'm trying to find. Not a direct copy, but I'm still kind of searching for this" (S2.I3.VP)

Intrinsically tied to the physiology or *RAMI*'s robotic arm, entrainment of action was anecdotally correlated by *M* between the tip of *RAMI*'s arm and our musician's left-hand; between the extremity of its "tall" gestures and higher pitches and conversely, between its lower extremities and lower pitches.

A final strategy employed was the use of *RAMI*'s natural sound. While we note that *RAMI* was not designed to produce sound, its pneumatic action, swishing, and clicking noises provided incidental material that *M* chose to connect to by "...listen[ing] to...mechanical sounds that I can pick up to respond to".

"[...] I'm enjoying this step because I caught the bottom string by accident, but it seemed to sort of fit with the sort of small jerky movements. So then I continued that and started exploring that" (S2.I2.VP)

We note that the incidental and accidental are closely connected in this sense, in that our musician's positive embrace of their own musical "mistakes" often proves fundamental to perceptions of successful improvisation.

4.4 Decision Making

Improvising with *RAMI* placed greater emphasis on *M*'s sense of decision making. This was manifest through *M*'s perception of taking responsibility for successful interactions with *RAMI* and resulting material, along with the commitment to the performance of specific musical ideas. These types of decision making stemmed from *M*'s perception of leading, rather than being led by *RAMI*. However, active decision making was also present in uncomfortable moments of "getting stuck in the loop" with *RAMI*, and so becoming (un)stuck was a significant point of decision making that seemed to suggest it was *RAMI* that was in fact leading.

In terms of taking responsibility for material, *M* remained hopeful for a feeling of connection, but nevertheless felt inclined to take the lead: "Well I'm looking and listening to go off and do my own thing regardless [but] I'm hoping [and] looking for a connection". Avoiding repetition of content while ensuring musical decisions were spontaneous prioritised this intuitive kind of lead-taking:

"[I'm] trying not to repeat any kind of similar things [...] I think it's just happening [...] it's not like, oh, I'm going to keep this now for two seconds and then I'm going to do this" (S2.I3.VP).

M emphasised the importance of intuitive, moment-to-moment decision-making: "Every, every single situation is different. I don't have a specific timing of when I'm going to change material or anything. It's really in the moment that I make those decisions".

Decision making was aided by anticipation, close proximity, and keeping up (with speed), and again, "mirroring" *RAMI*'s gestures:

"I was able to anticipate or be as close as possible to the movements and try and sort of just follow them as quickly as possible" (S3.I1.VP), and "[I'm] just seeing if I can mirror perhaps some movement. But I'm also committed to not changing my decision" (S3.I1.VP).

Overall, decision making appears integral to the development of musical themes, form, and structure. Balancing the development of the music itself with feeling connected to *RAMI* is therefore closely associated with our musician's perception of their own decision making and, perhaps, how this is regarded by *RAMI*.

4.5 Connection, Agency and Ensemble

Instances of connection, agency, and ensemble specifically refer to *M*'s perception of both leading and following and the apparent contingencies for shifting sense of agency between them and *RAMI*.

The sense of feeling connected or disconnected to *RAMI* based on its perceived agency was explicitly apparent at times and showed a development from anticipation of gestural movement to an appreciation of having become familiar, and even comfortable at points: "I'm a bit more relaxed about the connection. It doesn't have to be. I don't have to be connected the whole [time]". Throughout the sessions, while *M* reflects they were actively looking for connections with *RAMI*, they were also not actively evaluating this as successful or not in the moment: "I'm just really trying to connect and not really evaluating... 'Thank goodness *RAMI*'s done that" (S2.I2.VP). It is therefore in retrospect that *M* remarks: "I think I'm trying to follow. I'm trying to find connections in some of the movement" (S2.I3.VP).

There was again a sense that *M* adopted the responsibility for establishing connection with *RAMI*. At one point in the first session, *M* acknowledged that they did not find the right material to connect with *RAMI*, suggesting this might have been their fault (S1.I1.VP2).

In the moments that *M* felt connection was achieved, they described it as a collaborative and joint effort with *RAMI*: "Well, now I feel a bit that there's a bit more collaboration going on. Somehow a bit more connection" (S1.I2.VP2).

"The music or the sound is more connected now, or we're just more [...] we're moving OK as an ensemble, we're working as an ensemble" (S1.I2.VP2).

More implicitly, agency was most apparent in the tracking of moments of leading and following, where *M*'s sense of independence and unison came to the fore, for example: "Sometimes when you're improvising you can present an idea that's completely counter to what's going on but I feel as if [I'm] not completely ignoring what's in front of me at this point".

Whoever—or whatever—was perceived as leading at specific moments of the performance, what appeared to matter most (at least in terms of the success of the improvisation) was the creation of space for the uncertainty of chance encounters—or as our musician put it, "happy accidents".

4.6 Schemas of RAMI

M employed at least 8 distinct schemas when describing their interaction with *RAMI*, suggesting variability in how they perceived its role. These schemas ranged from seemingly passive to highly active, in turn shaping *M*'s improvisational experience.

In its most passive state, *RAMI* was seen as a *restriction*, serving as a limitation to improvisational creativity and flow, e.g., *RAMI* is just "not picking up what I'm doing" (S1.I1.2VP). For the most part, *RAMI* was perceived as background *accompaniment*, rather than active performance partner. Tentatively, *M* suggests that "I think I'm accompanying [its] small little movement [and am] focused on that little bit and exploring it with my pitches and sound"; and furthermore, that "I'm perhaps leading a bit more here [as] a soloist [while] *RAMI* is accompanying" (S3.I1). This is supported by *M*'s perception of performing like a soloist at this moment.

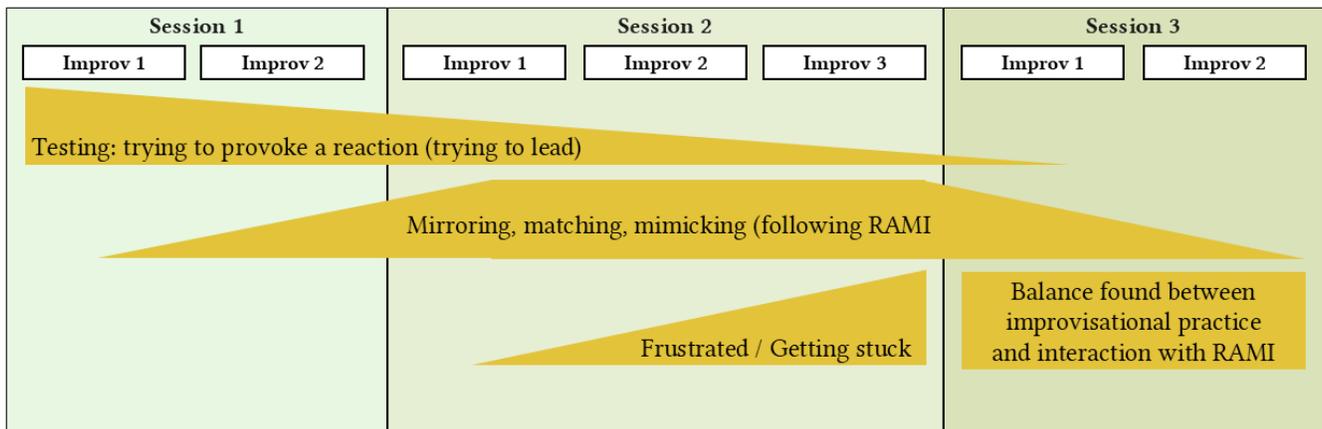


Figure 4: Musician's Journey Over Sessions

In another sense, though, and when *RAMI* was perceived as taking more of an active lead, it was likened to a *composer* in that our musician felt they were engaging in "compositionally tweak[ing]" their improvisation. This was best expressed by the perception of "add[ing] more layers to what I'm doing" (S3.I1). *RAMI* was also perceived as a *teacher*: "it might teach me to react in a different way or something will come out that I've no [previous] idea about" (S2.I2). Notably, this analogy points to the attribution of new material to *RAMI*, and perhaps by extension, an accreditation of creativity beyond the musician as human collaborator.

Furthermore, *RAMI* was—at times—perceived as a *performer*, specifically, a *dancer*. *M* described the interaction as evocative of "my experience with dancers" (S1.I1.2VP), with specific moments of *RAMI* "definitely dancing" (S1.I2), and leading eventually to a stronger conviction that, "I'm playing with a dancer there now" (S1.I2). Nevertheless, a sense of "expecting something else that I usually pick up with dancers that I'm not picking up [...] some kind of physicality" (S1.I1.2VP), stopped them short of maintaining this conviction.

In its most active role, *RAMI* emerged as a musical *provocateur*. While sound was not an intentionally designed dimension of our study, sounds of pneumatic limb movements—such as swishing and clicking—were perceived as invitations to respond. *M* noted, "[the sound is] an invitation to maybe match what I've just heard or try and play it [...] simultaneously" (S3.I1). *RAMI*'s movements and sounds "[give] me suggestions" (S3.I1) both intentionally, but also unintentionally: "[the] clicks and things are also inviting me to add even more layers" (S3.I1). *RAMI* is perhaps perceived as most intentionally provocative when *M* was "responding to [it's quick, sudden movement'] in this way" (S3.I1), especially when sound elicited congruent musical responses.

Finally, *RAMI* was actively embraced as a *score*, with *M* describing it as "a visual score. [I am] using small movements as a score" (S1.I1.2VP). In this sense, *RAMI* is seen but becomes embodied through a specific entrainment of movement (i.e. its smaller gestures). This schema allowed *M* to embody *RAMI*'s gestures as a framework for musical development, reflecting "there's some kind of development from [...] my part [...] But the thing is that it has

some kind of score in a way [...] it's sort of done something and [then] I've decided to do something." (S2.I2).

These schemas—*restriction*, *accompaniment*, *composer*, *teacher*, *performer*, *dancer*, *provocateur*, and *score*—were employed freely by *M* at different points throughout the sessions. We return to this topic in our final summary interview, to see, upon some passing of time and reflection, what *M* thought. They were clear in stating "*RAMI*'s a score", but adding, "I respond to *RAMI* as a dancer".

4.7 Enjoying *RAMI*

RAMI was explicitly enjoyable for *M* at certain moments and given that this experience is central to sustainable improvisation, we elaborate on the ways in which *RAMI*'s dynamic qualities, moments of synchronicity, and moments of precision, came together for *M* as notable experiential peaks.

The dynamic nature of *RAMI* and its expressive action as a robotic arm was manifest in speed and changes to tempo, ensuring that "the speed of it was more interesting to react to". *RAMI* was not simply more enjoyable when it was faster, rather, it was more enjoyable when these changes felt less predictable. Where tempo, pitch, or gesture were subjectively and objectively aligned, there was a sense this led to "satisfying moments [...] in that they're synched up accidentally, but it's because I have picked up something probably".

"I'm very aware of the sound that *RAMI* is making [...] it's just feeding into some of my musical decisions, perhaps [...] I'm conscious [and] listening to the sounds that it makes as well as watching the movement"

Perhaps most significantly considering the intersection of enjoyment and perceived togetherness, the perception of precision movements and the entrainment of those to musician actions was notable: "These kinds of moments of certainty led to evaluations that 'I am confident I am creating a piece at this point'".

4.8 Journey Development

We conclude the findings with a short summary of how *M*'s interaction with *RAMI* developed over the course of the three study

sessions, as illustrated in Figure 4. *M* summarised this journey in the final summary interview:

“Definitely the first session was testing. And then the second session is sort of trying to find sort of similarity, you know, common ground, if there can be. Perhaps the third session [...] I’ve found my way of responding.

M’s words highlight their application of different engagement strategies, as presented earlier in the findings. The testing process in the first session was typical of *M*’s experience in other musicking settings, “First encounter can be that, you know, it’s all about sort of just looking at ways through it.” In the second session, *M* pivoted away from trying to provoke responses from *RAMI* (or recognising *RAMI*’s responses), to forging connections to *RAMI*, largely by mirroring, mimicking and matching its physical and sonic gestures. Once more, *M* noted that this technique was a typical facet of their practice, “[it] is something that I do with my practice as well. So, it feeds into what we discussed about instability or something unexpected.” Towards the end of the second session, *M* (in the video protocol) expressed moments of getting stuck in the improvisations, and struggling to trying to find a way out. In the final interview, we returned to this topic where *M* noted that, “I suppose that can happen with other musicians, but I think maybe it was a bit of a shortcoming of *RAMI* that it was quite limited.” *M* expanded on what this meant:

“In an improvised session, somebody will build something up like a rhythm or unexpected thing that you can lock into [i.e., a musical theme or idea]. Perhaps that’s what I was missing, *RAMI* would do a quick movement, I’d follow, but then *RAMI* wasn’t able to sustain it”

M went on to describe that *RAMI*’s movements were quite similar and these movements were enacted with a tight range of velocity, but it lacked structure, discernible patterns and repetition, noting that if there’s no pattern or discernible structure from an accompanist, then it’s really hard to lock into it or disrupt it. Later noting, “by the third session I’d accepted, ‘well that this is what it is, basically’”.

5 Discussion

An initial reading of our findings would suggest that *M*’s journey of improvisation with *RAMI* was a linear one, with a progressive narrative that asserted enjoyable musicking as contingent on initial moments of familiarization Figure 5. As *M*’s process of making sense of their interaction was inherently reflexive though, we reconceptualize this continuum of interaction as grounded more in terms of (un)certainty of interaction perception. Consequently, we position *schemas* as the most certain of our themes of perception, on the basis that the eight instances offered were intrinsically exemplary of the category and required little further qualification. By contrast, perceptions of engaging in *decision making* were contingent on six identified sub-themes, which were not exemplary of particular kinds of decision in the same way (Figure 6).

We therefore focus our discussion on the nature of our human-AI interaction themes as relatively (un)certain in definition, on two relatively certain and enduring schemas of *RAMI* that persisted

beyond the immediate workshops, and on the implications of perceiving *RAMI* as on the one hand a tool for improvisation, and on the other, an improvising agent.

5.1 Making sense of (un)certainty: operationalizing narrative perceptions of *RAMI*

Implicit in our findings and during explicit self-reflection, *M* frequently returned to (un)certainty, ambiguity, and (un)predictability not as barriers to creativity, but as intrinsically enjoyable and as prerequisite for creative output. However, it is also apparent that these same qualities can stifle enjoyment and creativity at the same time. This suggests an underlying precondition for the unpredictable actions of *RAMI*. For example, unpredictability of the precise motion of the tip of the arm was viewed positively as it inspired entrainment of motion and precision when creating sound. Conversely, unpredictability of the intentionality of *RAMI*—the unknownness of the mode it may have been set to, for example—was perhaps more misaligned with conventional improvisation, given that human co-improvisers usually have shared background knowledge about the musical style and at least some informed expectations about how each other might play. Discerning the welcome and unwelcome contexts for (un)certain interaction on the part of *RAMI* is clearly vital, and so our findings help in this sense by providing a tentative, ordinal representation of perceptions based on their nature as standalone qualities, or qualities reliant on further knowledge.

5.2 Schemas of *RAMI*: Score and Dancer

As *M* became familiarized with *RAMI*, they relied on a series of schemas—preconceived frameworks shaped by their extensive experience improvising with musicians, dancers, objects, and their dual role as composer and performer—to interpret its behavior. These schemas were by far the most certain perceptions of *RAMI*, and dynamically shifted depending on the context of interaction. For example, when *M* felt stuck, they looked to *RAMI* to lead, likening it to a teacher or guide. When they aligned physically with *RAMI*’s expressive movements, they perceived it as a dancer. And when they prepared with specific gestures, they treated *RAMI* as a musical score. These shifting perceptions were immediate and tied to their mode of engagement at the time, drawing on familiar frameworks of musical interaction.

However, these schemas occasionally created tension, as *RAMI*’s behavior did not always align with expectations of how a collaborator, whether human or machine, should behave. This misalignment was particularly evident in the first session, when *M* noted, “And I suppose here I’m surprised if I was playing with a colleague, perhaps they would have taken it off onto somewhere different” (S1.I1.2VP). Such moments highlight the tension of applying human-centric schemas to an AI-driven system like *RAMI*.

At the start of *M*’s experience with *RAMI*, they initially followed their usual approach to improvisation, trying to establish a complementary language with *RAMI* by drawing on their prior experience. However, *M* struggled to connect with *RAMI*, perceiving it as unpredictable and disengaged, despite it actively listening and responding. This disconnect may have stemmed from the framing provided by the research team, which was deliberately vague, coupled with *M*’s

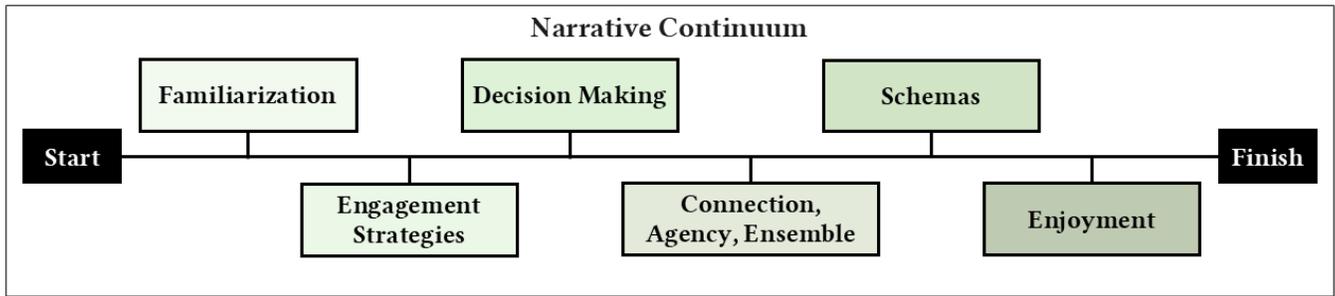


Figure 5: Themes as human-AI interaction narrative

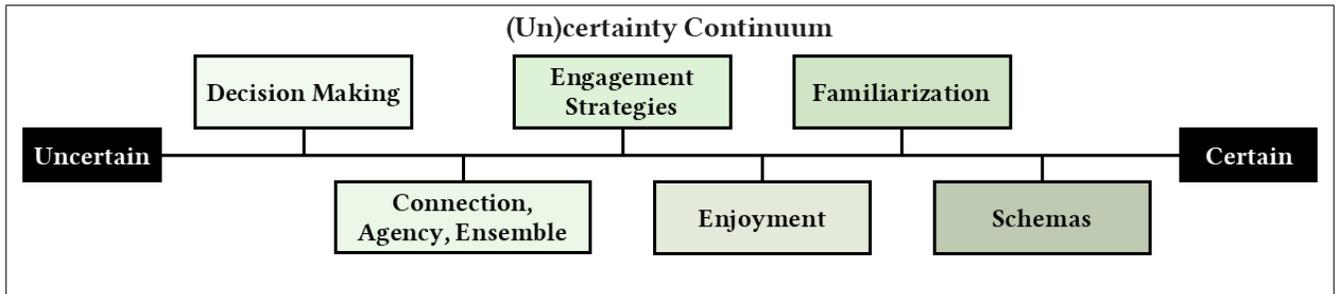


Figure 6: Themes as human-AI interaction (un)certainty

limited knowledge and skepticism about AI. Additionally, *RAMI*'s subtle gestures and behaviors may have been too understated for *M* to interpret as communicative.

According to Bishop's [4] model of musical togetherness, a shared understanding and belief in a bidirectional exchange are essential for meaningful interaction. Since *M* did not perceive *RAMI*'s responses as intentional or bi-directional, they did not form this belief. Bishop's model proposes several other factors that underlie togetherness experiences: perceived liveness, perceived responsiveness, and embodied alignment. In terms of perceived liveness, *RAMI* offers multimodal (auditory-visual) cues that emphasize its physical presence. *M* was explicitly receptive to multiple forms of communication from *RAMI* such as shapes, speed and sound, and there were times when *M* speculated that *RAMI* was observing them. Objectively, liveness was in place, as *M* and *RAMI* were situated, interacted together, were often synchronised and communicated, although *M* was not aware of how *RAMI* received communications. Consequently and despite this liveness, there remained no belief that *RAMI* was intentional.

In terms of perceived responsiveness, *M* did not readily perceive *RAMI* producing expressions that were complementary to their own sound or gestures. This contrasts with previous findings for *Jess+* [39], where musicians reported an awareness of whether or not *Jess+* was listening. In terms of embodied alignment, embodiment was a core facet of *M*'s experience: They worked hard to be in physical and gestural alignment with *RAMI* as a strategy to find musical connections and shape their improvisation approach, but this ultimately lacked a shared interpretation of where the performance was going. In short, it seems that the basic criteria that support the

experience of togetherness in human-human musical collaboration were not met in the case of *M*'s interaction with *RAMI*. This helps to explain why *M* did not establish an enduring perception of *RAMI* as a social musical partner.

We reflect that because we impaired *M*'s ability to know anything about *RAMI*'s belief system, their interpretation of *RAMI* is partially a result of the method. However, particular identification of *RAMI* as a dancer and score were prominent and, as mentioned, reiterated in the final interview. As a dancer *RAMI* was witnessed through *M*'s numerous description of its gestures and references to prior experiences of working with dancers. Reaching for these analogies was part of *M*'s process of getting to know *RAMI* and understanding its knowledge and expectations. *M* stated that in their previous experiences of working with dancers, they did not view them as scores, but rather as collaborators or performers. *M*, for the most part, followed and responded to *RAMI* using its gestures and sounds to stimulate or drive improvisational choices, a process typically referred to in HCI with robots as entrainment [15]. Despite uncertainty over *RAMI*'s role in interpreting gesture and sound, it was still capable of entraining similar means and ends in *M*'s musical performance and output. *M*'s sense of maintaining responsibility for this performance and output supports a sense of *RAMI* functioning as a motivational score, rather than as a first class agent in the ensemble.

In this study, we observe that *RAMI* actually functioned as a hybrid artifact: It was part score, part dancer, and thus its function and purpose were plastic and malleable, i.e., open-ended to the individual human to make sense of and form a relationship with.

5.3 RAMI: Improvisation tool or Improvisation agent?

Given that *RAMI* was described by *M* most specifically in terms of schemas of score and dancer, and that engagement with these schemas was means- versus ends-orientated, we might ask, to what extent was *RAMI* perceived as functional (a tool) versus embodied (agential)? While this work highlights the underlying variable of (un)certainly as a means of standardizing observations of qualitative perception—potentially by measuring (un)certainly of salience of perceptions in future works—it also points toward a need to evaluate schemas such as score and dancer on another speculative continuum; that of AI tool versus AI agent. This offers a provocative challenge to conventions in human-human improvisation, inviting the question: On what grounds can interaction with a digital score be considered agential? The quantification of our perceptions of *RAMI* in terms of (un)certainly here appears crucial to answering this.

In the literature on AI and creativity, there appears to be a preference for the agential application of AI in its collaboration with humans in artistic projects [33]. *Life 3.0* explores the profound implications of advanced AI on human creativity, suggesting that AI has the potential to function as a collaborative partner in the creative process. As observed with our musician, there were instances where *RAMI*'s behaviours stimulated their output and prompted them to respond to it as a collaborator—but only when they felt a sense of connection or togetherness, or when the uncertainty of the situation felt intriguing and stimulating enough to inspire their creative engagement. Thus, it seems there is more nuanced dynamic in play when a robotic AI agent comes in contact with an improvising musician, as the roles of interactions are not quite as clear cut as they might be in other creative forms.

5.4 Limitations and Future Works

What could seem as a limitation in our study is that we did not initiate a co-creative technical development process between *RAMI* and *M*, which could have enabled a more genuine collaborative engagement. As [24] highlights, the advantage of such an approach lies in actively involving the human collaborator in the development process, allowing them to shape the system in a way that aligns with their creative agency. By co-developing the system alongside the musician, the collaboration could have evolved into a more organic and mutually responsive partnership, enhancing the sense of shared creativity and agency. However, *RAMI* was designed as a digital score with its own belief system and a degree of agential freedom in interaction, even if this freedom might have been perceived as limited by our expert improvising musician. While it operated within defined parameters, its capacity to engage dynamically and make autonomous decisions introduced an element of unpredictability and agency, aiming to enable a sense of collaboration rather than merely serving as a passive tool.

Another potential limitation of this study is its focus on a single participant, which neglects statistically grounded generalization and potentially amplifies subjective biases. However, this experiment should be viewed as part of a broader lineage of research, building on the work of *Jess+* [39], and thus contributes to a growing corpus of evidence.

In terms of future works, our results point to critical thematic touch-points for engaging with *RAMI*. Beyond simply describing participants' experiences, these themes reveal expectations of *RAMI* as an AI improvisation partner that differ from those of a conventional human musician. To the extent that these expectations can be formalized, we proposed schemas for *RAMI* that characterize not only its affordances but also its identity as a recognized collaborator. These schemas draw from human-human roles, such as teacher, composer, accompanying artist, and dancer, suggesting that *RAMI* can be personified. At the same time, more mechanistic identities, such as provocateur and score, indicate that creative and collaborative agency can be embodied in new ways, leading to the emergence of novel improvisational collaborators. Schemas in particular also provide valuable design insights for in-the-loop musicking AI. The balance between certainty and uncertainty seems to mirror the dynamics of human-human ensemble playing, where these relationships are highly valued. Furthermore, digital scores created through human-AI co-creativity could be designed using the six themes identified in this study.

The continuum presented above can be reverse-engineered to align with findings from previous research, while also raising new questions for future experimental designs and studies. Together, the emerging set of schemas suggests insights into musicians' expectations of human-AI improvisation. These schemas also highlight the kind of unpredictability that consistently leads to desirable musical outcomes and experiences, which, in turn, raises important considerations about the boundaries of *RAMI*. This may offer a generalizable principle for studying human-AI musicking.

Finally, we acknowledge the continued tension between the live and the retrospective [31]; particularly when asserting theme significance. It was saturation and clarity of content that underscored our sessional themes for instance, yet the endurance of theme salience elicited in the final session provided a more authoritative validation. Ultimately, a technical challenge remains in how to bring the retrospective into the loop, such that we can increase the certainty of subjective musician classification of meaningful interaction. This is a work in progress.

6 Conclusions

In this study, we investigated further the meaning-making phenomenon reported in a previous study [39] with a digital score presented as an AI-enabled robot arm. By establishing a methodology that captured the subjective perceptions of a professional musician (*M*) through qualitative methods, and a systematic process of thematic analysis, we presented insights into how a professional musician perceives, and perhaps makes sense of, human-AI musicking. We found that *M* engaged in a process of familiarization, developing engagement strategies, and decision making that bore resemblance to established paradigms of human-human improvisation. Conversely, we also found that comparatively in-the-moment stages of forming *connection*, *agency* and *ensemble*, and maintaining *enjoyment* of interaction and musical output are frequently perceived as more human- than AI-led. Among these stages however, the identification of *schemas of RAMI* demonstrate discernible characteristics of a nascent improvisation prototype that were variably treated as

tools and agential, or as passive and active accessories to the improvisation. We propose that these schemas inform the acceptable bounds of future *RAMI* tools for human-AI musicking, and perhaps, the personas of *RAMI* itself when perceived as more fully agential.

Acknowledgments

This work was funded by the ERC under the European Union's Horizon 2020 research and innovation programme (ERC-2020-COG – 101002086), and the Research Council of Norway through its Centres of Excellence scheme, project number 262762.

References

- [1] James E Allen, Curry I Guinn, and Eric Horvitz. 1999. Mixed-initiative interaction. *IEEE Intelligent Systems and their Applications* 14, 5 (1999), 14–23.
- [2] Steve Benford, Adrian Hazzard, Craig Vear, Helena Webb, Alan Chamberlain, Chris Greenhalgh, Richard Ramchurn, and Joe Marshall. 2023. Five Provocations for a More Creative TAS. In *Proceedings of the First International Symposium on Trustworthy Autonomous Systems*. 1–10.
- [3] Steve Benford, Richard Ramchurn, Joe Marshall, Max L Wilson, Matthew Pike, Sarah Martindale, Adrian Hazzard, Chris Greenhalgh, Maria Kallionpää, Paul Tennent, et al. 2021. Contesting control: journeys through surrender, self-awareness and looseness of control in embodied interaction. *Human-Computer Interaction* 36, 5-6 (2021), 361–389.
- [4] Laura Bishop. 2024. Togetherness in musical interaction. *Routledge Open Research* 3, 16 (2024), 16.
- [5] Oliver Bown, Alice Eldridge, and Jon McCormack. 2009. Understanding interaction in contemporary digital music: From instruments to behavioural objects. *Organised Sound* 14, 2 (2009), 188–196.
- [6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3 (2006), 77–101. Issue 2. doi:10.1191/1478088706qp0630a
- [7] Mason Bretan and Gil Weinberg. 2016. A survey of robotic musicianship. *Commun. ACM* 59, 5 (2016), 100–109.
- [8] Rodney A Brooks. 1991. Intelligence without representation. *Artificial intelligence* 47, 1-3 (1991), 139–159.
- [9] David Byrne. 2022. A worked example of Braun and Clarke's approach to reflexive thematic analysis. 56 (2022), 1391–1412. doi:10.1007/s11135-021-01182-y
- [10] Carlos Cancino-Chacón, Silvan Peter, Patricia Hu, Emmanouil Karystinaios, Florian Henkel, Francesco Foscarin, Nimrod Varga, and Gerhard Widmer. 2023. The ACCompanion: Combining Reactivity, Robustness, and Musical Expressivity in an Automatic Piano Accompanist. In *In Proceedings of the 32th International Joint Conference on Artificial Intelligence (IJCAI-23)*. Macau, S.A.R., 5779–5787. doi:10.24963/ijcai.2023/641
- [11] Kerstin Dautenhahn. 2007. Socially intelligent robots: dimensions of human-robot interaction. *Philosophical transactions of the royal society B: Biological sciences* 362, 1480 (2007), 679–704.
- [12] Nicholas Davis, Chih-Pin Hsiao, Kunwar Yashraj Singh, Brenda Lin, and Brian Magerko. 2017. Creative sense-making: Quantifying interaction dynamics in co-creation. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*. 356–366.
- [13] Arne Dietrich. 2004. The cognitive neuroscience of creativity. *Psychonomic bulletin & review* 11 (2004), 1011–1026.
- [14] Théo Golvet, Louise Goupil, Pierre Saint-Germier, Benjamin Matuszewski, Gérard Assayag, Jérôme Nika, and Clément Canonne. 2021. With, against, or without? Familiarity and Copresence Increase Interactional Dissensus and Relational Plasticity in Freely Improvising Duos. *Psychology of Aesthetics, Creativity, and the Arts* (2021). doi:10.1037/aca0000422
- [15] Takamasa Iio, Masahiro Shiomi, Kazuhiko Shinozawa, Takaaki Akimoto, Katsunori Shimohara, and Norihiro Hagita. 2011. Investigating Entrainment of People's Pointing Gestures by Robot's Gestures Using a WOZ Method. *International Journal of Social Robotics* 3 (2011), 405–414. Issue 4. doi:10.1007/s12369-011-0112-0 nuances between experiments conducted 'in the lab' and 'in the wild'; to discuss natural language, alternative and augmented signal input, and to widen an understanding of the initial 'task analysis' stage to consider other participants as 'stakeholders' in technology propositions.
- [16] Miriam Iorwerth and Don Knox. 2019. Playing Together, Apart. *Music Perception* 36, 3 (2019), 289–299. doi:10.1525/mp.2019.36.3.289
- [17] Wendy Ju and Larry Leifer. 2008. The design of implicit interactions: Making interactive systems less obnoxious. *Design Issues* 24, 3 (2008), 72–84.
- [18] Harmanpreet Kaur, Eytan Adar, Eric Gilbert, and Cliff Lampe. 2022. Sensible AI: Re-imagining interpretability and explainability using sensemaking theory. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 702–714.
- [19] Steven Kemper. 2021. Locating Creativity in Differing Approaches to Musical Robotics. *Frontiers in Robotics and AI* 8 (2021), 647028.
- [20] George E Lewis. 2000. Too many notes: Computers, complexity and culture in voyager. *Leonardo music journal* 10 (2000), 33–39.
- [21] Todd Lubart, Dario Esposito, Alla Gubenko, and Claude Houssemand. 2021. Creativity in humans, robots, humbots. *Creativity: Theories—Research—Applications* 8, 1 (2021).
- [22] Jon McCormack and Mark d'Inverno. 2012. Computers and creativity: The road ahead. In *Computers and creativity*. Springer, 421–424.
- [23] Jon McCormack, Toby Gifford, Patrick Hutchings, Maria Teresa Llano Rodriguez, Matthew Yee-King, and Mark d'Inverno. 2019. In a silent way: Communication between ai and improvising musicians beyond sound. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–11.
- [24] Jon McCormack, Patrick Hutchings, Toby Gifford, Matthew Yee-King, Maria Teresa Llano, and Mark D'Inverno. 2020. Design considerations for real-time collaboration with creative artificial intelligence. *Organised Sound* 25, 1 (2020), 41–52.
- [25] Eduardo R Miranda, Wendy L Magee, John J Wilson, Joel Eaton, and Ramaswamy Palaniappan. 2011. Brain-computer music interfacing (BCMI): from basic research to the real world of special needs. *Music & Medicine* 3, 3 (2011), 134–140.
- [26] P Russel Norvig and S Artificial Intelligence. 2020. Design considerations for real-time collaboration with creative artificial intelligence. *Organised Sound* 25, 1 (2020), 41–52.
- [27] Eduardo R Miranda, Wendy L Magee, John J Wilson, Joel Eaton, and Ramaswamy Palaniappan. 2011. Brain-computer music interfacing (BCMI): from basic research to the real world of special needs. *Music & Medicine* 3, 3 (2011), 134–140.
- [28] P Russel Norvig and S Artificial Intelligence. 2020. Design considerations for real-time collaboration with creative artificial intelligence. *Organised Sound* 25, 1 (2020), 41–52.
- [29] Eduardo R Miranda, Wendy L Magee, John J Wilson, Joel Eaton, and Ramaswamy Palaniappan. 2011. Brain-computer music interfacing (BCMI): from basic research to the real world of special needs. *Music & Medicine* 3, 3 (2011), 134–140.
- [30] P Russel Norvig and S Artificial Intelligence. 2020. Design considerations for real-time collaboration with creative artificial intelligence. *Organised Sound* 25, 1 (2020), 41–52.
- [31] Eduardo R Miranda, Wendy L Magee, John J Wilson, Joel Eaton, and Ramaswamy Palaniappan. 2011. Brain-computer music interfacing (BCMI): from basic research to the real world of special needs. *Music & Medicine* 3, 3 (2011), 134–140.
- [32] P Russel Norvig and S Artificial Intelligence. 2020. Design considerations for real-time collaboration with creative artificial intelligence. *Organised Sound* 25, 1 (2020), 41–52.
- [33] Eduardo R Miranda, Wendy L Magee, John J Wilson, Joel Eaton, and Ramaswamy Palaniappan. 2011. Brain-computer music interfacing (BCMI): from basic research to the real world of special needs. *Music & Medicine* 3, 3 (2011), 134–140.
- [34] P Russel Norvig and S Artificial Intelligence. 2020. Design considerations for real-time collaboration with creative artificial intelligence. *Organised Sound* 25, 1 (2020), 41–52.
- [35] Eduardo R Miranda, Wendy L Magee, John J Wilson, Joel Eaton, and Ramaswamy Palaniappan. 2011. Brain-computer music interfacing (BCMI): from basic research to the real world of special needs. *Music & Medicine* 3, 3 (2011), 134–140.
- [36] P Russel Norvig and S Artificial Intelligence. 2020. Design considerations for real-time collaboration with creative artificial intelligence. *Organised Sound* 25, 1 (2020), 41–52.
- [37] Eduardo R Miranda, Wendy L Magee, John J Wilson, Joel Eaton, and Ramaswamy Palaniappan. 2011. Brain-computer music interfacing (BCMI): from basic research to the real world of special needs. *Music & Medicine* 3, 3 (2011), 134–140.
- [38] P Russel Norvig and S Artificial Intelligence. 2020. Design considerations for real-time collaboration with creative artificial intelligence. *Organised Sound* 25, 1 (2020), 41–52.
- [39] Eduardo R Miranda, Wendy L Magee, John J Wilson, Joel Eaton, and Ramaswamy Palaniappan. 2011. Brain-computer music interfacing (BCMI): from basic research to the real world of special needs. *Music & Medicine* 3, 3 (2011), 134–140.
- [40] P Russel Norvig and S Artificial Intelligence. 2020. Design considerations for real-time collaboration with creative artificial intelligence. *Organised Sound* 25, 1 (2020), 41–52.
- [41] Eduardo R Miranda, Wendy L Magee, John J Wilson, Joel Eaton, and Ramaswamy Palaniappan. 2011. Brain-computer music interfacing (BCMI): from basic research to the real world of special needs. *Music & Medicine* 3, 3 (2011), 134–140.