



# Amplified Connections

---

Matthias Jung

---

**Amplified Connections:**



Matthias Jung

Amplified Connections:  
Exploring Performative Co-Creation and  
Distributed Creativity in the Context of  
Accessible Music Creation Technologies

Doctoral Dissertation for the Degree *Philosophiae Doctor (Ph.D.)*  
at the Faculty of Fine Arts, Specialization in Popular Music Performance

University of Agder  
Faculty of Fine Arts  
2024

Doctoral Dissertations at the University of Agder 505  
ISSN: 1504-9272  
ISBN: 978-82-8427-223-8

©Matthias Jung, 2024

Printed by Make!Graphics  
Kristiansand

# Abstract

## English

Keywords

*Intelligent Music Performance Systems, Audience Participation, Distributed Creativity*

In recent decades, interactive music systems have begun to be integrated into popular music performance. Advances in digital technology have enabled the inclusion of a large number of co-creators, encompassing even the contributions of audience members at live concerts in real time. In addition, autonomous music technologies have transformed into a readily accessible tool for both composers and system creators. These developments motivated this thesis project, which investigates the field of participatory performance via a focus on system design, use, and experience within popular-music concert environments.

The thesis asks the following questions: How can interactive paradigms be integrated into live music performances, and how are they connected to the needs of both participating audience members and stage performers? What does the experience of these participatory systems depend on, and how can we typologize different ways of using them? Lastly, how do current autonomous concepts in music generation inform the design and use of interactive music performance systems? To address those questions, the thesis establishes its research context and approach in its opening theoretical part, which includes three framing chapters, then presents the results of its empirical work in the form of two preliminary studies and four articles that were completed within the research period.

The framing section opens with the chapter titled **Performance**, which is focused on musical impact and audience response via the notions of musicking, embodiment, and musical gesture. The chapter titled **Participation** explores the use of these audience gestures to contribute to concert performances. It theorizes various concepts of distributed creativity with the help of postphenomenology and human-computer interaction (HCI), which it applies to the case of "audience instruments." The chapter titled **Autonomy** reviews aspects of artificial creativity with respect to its use in interactive music performance systems.

Following two preliminary studies on audience gesture and performer motion, the empirical part of the thesis presents four articles. The first article is a study of audience participation involving the implementation of three interactive parts in an electronic live concert, followed by a self-report questionnaire, a group interview with audience members, and a performer interview. The second article turns to the further development of interaction software and analyzes the audience members' behavior in more depth, generating, in turn, a user typology for the participants. The third article introduces autonomous concepts in interactive performance systems and suggests a three-dimensional model with the categories *embodiment*, *participation* and *autonomy*. The model is then applied to the analysis of nine intelligent music performance systems (IMPS) and reframed from the perspective of speculative design research. The fourth article engages with the mechanics of the systems' intelligence and suggests certain alternative computation principles such as "symbiotic algorithms."

## Norsk

### Sammendrag

Nøkkelord

*Smarte Musikkframføringssystemer, Publikumsdeltakelse, Distribuert Kreativitet*

I de siste tiårene har interaktive musikkssystemer begynt å bli integrert i rytmisk musikkframføring. Utviklingen innen digital teknologi har gjort det mulig å inkludere et stort antall medskapere, inkludert bidragene fra publikum under livekonserter i sanntid. I tillegg har autonome musikkteknologier utviklet seg til å bli et lett tilgjengelig verktøy for både komponister og systemskapere. Denne utviklingen har motivert dette avhandlingsprosjektet, som utforsker feltet for deltakende framføring ved å fokusere på systemdesign, bruk og erfaringer innenfor konsertmiljøer for rytmisk musikk.

Målsetting for denne avhandlingen er å belyse følgende spørsmål: Hvordan kan interaktive paradigmer bli integrert i live musikkframføringer, og hvordan er de knyttet til behovene til både deltakende publikum og sceneutførere. Hva avhenger opplevelsen av disse deltakende systemene av, og hvordan kan vi inndele ulike måter å bruke dem på? Til sist, hvordan påvirker aktuelle autonome konsepter innen musikkgenerering designet og bruken av interaktive musikkframføringssystemer? For å svare på disse spørsmålene, etableres forskningskonteksten og tilnærmingen i den innledende teoretiske delen av denne avhandlingen, som består av tre rammekapitler. Deretter presenteres resultatene av det empiriske arbeidet i form av to pilotstudier og fire artikler som ble fullført innenfor forskningsperioden.

Den innledende delen begynner med kapittel **Performance** som fokuserer på musikalsk innflytelse og publikumsrespons gjennom begrepene musisering, utførelse, og musikalske

gester. Kapittel **Participation** utforsker hvordan bruken av disse publikumsgestene påvirker konsertframføringer. I dette kapitlet teoretiseres det ulike konsepter av distribuert kreativitet ved hjelp av postfenomenologi og menneske-datamaskininteraksjon (HCI; fra engelsk human-computer interaction), som anvendes i spørsmålet om "publikumsinstrumenter." I kapitlet **Autonomy** gjennomgås aspekter ved kunstig kreativitet med hensyn til bruken i interaktive musikkframføringssystemer.

Etter to forstudier om publikumsgester og utførerens bevegelser, inneholder den empiriske delen av avhandlingen fire artikler. Den første artikkelen inneholder en studie om publikumsdeltakelse som involverer implementeringen av tre interaktive elementer i en live elektronisk konsert. Data ble innhentet gjennom selvrappor-teringsskjema, gruppeintervju med publikumsmedlemmer og et intervju med utøvere. Den andre artikkelen tar for seg videreutviklingen av interaksjonsprogramvaren og analyserer publikums oppførsel mer grundig, og genererer på sin side en brukertypologi for deltakerne. I den tredje artikkelen introduseres autonome konsepter i interaktive framføringssystemer som foreslår en tredimensjonal modell med kategoriene *embodiment*, *participation* og *autonomy*. Deretter blir modellen anvendt på analysen av ni smarte musikkframføringssystemer (IMPS; fra engelsk intelligent music performance systems) og omformulert fra perspektivet til spekulativ designforskning. Den fjerde artikkelen analyserer mekanikken til systemenes intelligens og foreslår visse alternative beregningsprinsipper som "symbiotiske algoritmer."

# Publications

- Jung, M. and Kummen, V. 2024 (published). "Hacking the Concert Experience: Exploring Co-Creative Audience Interaction at a Chiptune Live Performance" in G ullo, J.-O. and Hepworth-Sawyer, R. and Paterson, J. and Toulson, R.: *Innovation in Music: Technology and Creativity*, Routledge. <https://doi.org/10.4324/9781003118817-8>
- Jung, M. and Clester, I. 2023 (published). "Hack the Show: Design and Analysis of Three Interaction Modes for Audience Participation" in Buffa, M. and Alexandraki, Ch. and Lebrun, J. and Roberts, Ch. (eds.): *Fostering Creativity through Web Audio*. *Journal of the Engineering Society (Special Issue)*. <https://doi.org/10.17743/jaes.2022.0111>
- Jung, M. 2023 (published). "Intelligent Music Performance Systems: Towards a Design Framework" in *Studia Musicologica Norvegica*. Universitetsforlaget. Vol.49, Iss.1: 28-44. <https://doi.org/10.18261/smn.49.1.3>
- Jung, M. 2023 (published). "Beyond Mutation: How Can We Acknowledge Symbiogenesis in Evolutionary Music Coding?". *The International Conference on AI and Musical Creativity*. Brighton. <https://aimc2023.pubpub.org/pub/if7rl6o5>

# Contents

<b>Abstract</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background . . . . .	1
1.2 Research Questions . . . . .	3
1.3 Research Aims . . . . .	5
1.4 Thesis Outline . . . . .	5
<b>I Theoretical Framework</b>	<b>9</b>
<b>2 Performance</b>	<b>11</b>
2.1 What Is Music Performance? . . . . .	11
2.1.1 Musicking . . . . .	12
2.1.2 Embodiment . . . . .	13
2.1.3 Gestures . . . . .	15
2.1.4 Instruments . . . . .	16
2.1.5 Technologies . . . . .	20
2.2 How Do We Listen to Performance? . . . . .	24
2.2.1 The Ecological Approach . . . . .	24
2.2.2 Emotional Response . . . . .	25
The Circumplex Model for Affect . . . . .	25
Other Models for Emotional Response . . . . .	26
Methods for Measuring Emotional Response . . . . .	26
2.2.3 Social Music Cognition . . . . .	28
Contagion and Empathy . . . . .	28
Synchronization and Attuning . . . . .	28
2.2.4 Performance as Staged Emotion . . . . .	29
2.2.5 Audience Gestures . . . . .	30
Embodied Gestures for Participatory Performance . . . . .	30
Audience Gesture as Compositional Material . . . . .	31
Audience Gestures and 4E Cognition . . . . .	31
2.3 Conclusion . . . . .	32

<b>3</b>	<b>Participation</b>	<b>33</b>
3.1	How Do We Interact during Performance?	33
3.1.1	Models for Musical Communication	33
3.1.2	Distributed Music Systems	36
	Structural Approaches	36
	Lean-in Consumption and the Idea of Prosumers	39
3.1.3	Concerns around System Design and Cultural Critique	40
	The Bandwidth Bottleneck	40
	Cultural Critique of Participatory Systems	42
3.2	How Can We Participate in Performance?	44
3.2.1	Distributed Creativity	44
3.2.2	Audience Instruments	47
3.3	Conclusion	48
<b>4</b>	<b>Autonomy</b>	<b>49</b>
4.1	What Are Autonomous Performing Machines?	50
4.1.1	Theory: Artificial Creativity	50
4.1.2	Intention: Creative Agency	54
4.1.3	Concepts: Underlying Principles of AI in Music	56
4.2	How Can We Build Co-Performing Machines?	59
4.2.1	Words: Terminology around Intelligent Music Performance Systems (IMPS)	59
4.2.2	Practices: Musical Interaction with Generative Systems	63
4.2.3	Touch and Immediacy: Embodiment and Human-Machine Integration	65
4.2.4	Models: Conceiving Interaction with Creative Agents	67
4.3	How Can We Use Autonomy for Participatory Performance Systems?	69
4.4	Conclusion	72
<b>5</b>	<b>Conceptual Frameworks Revisited</b>	<b>73</b>
5.1	Summary of Theoretical Perspectives	73
5.1.1	Performative Co-Creation	73
5.1.2	Participation and Distributed Creativity	74
5.1.3	Accessible Music Creation Technologies	76
5.2	Summary of Conceptual Approaches	79
5.2.1	Audience Gestures, Audience Instruments, and Intelligent Music Performance Systems	79
5.2.2	Co-Creation as a Working Concept	80
5.3	Methodological Approaches	83
<b>II</b>	<b>Empirical Studies</b>	<b>85</b>
<b>6</b>	<b>Research Design and Applied Methods</b>	<b>87</b>
6.1	Research Design	87

6.2	Quantitative Methods	88
6.2.1	Machine-Learning Methods	88
6.2.2	Motion-Capture Methods	88
6.2.3	Log Data Analysis	88
6.2.4	Questionnaire Analysis	89
6.3	Qualitative Methods	89
6.3.1	Practice-based Research Design	89
6.3.2	Interviews	90
6.3.3	Case Studies	90
6.3.4	Speculative Design	91
<b>7</b>	<b>Preliminary Studies</b>	<b>93</b>
7.1	<b>Preliminary Study 1: Clustering Audience Reaction Audio</b>	93
7.1.1	Application and Relevance	93
7.1.2	Machine-Learning Algorithm and Feature Extraction	94
7.1.3	Dataset	95
7.1.4	Implementation	96
7.1.5	Evaluation	96
7.1.6	Discussion, Conclusion, and Further Work	99
7.2	<b>Preliminary Study 2: How Does Audience Participation Affect Performer Motion?</b>	100
7.2.1	Motivation, Background, and Previous Work	100
7.2.2	Experimental Setup	100
7.2.3	Data Collection	101
7.2.4	Qualitative and Quantitative Data Analysis	104
7.2.5	Discussion	107
7.2.6	Conclusions and Further Work	109
7.3	Conclusion of the Preliminary Studies	110
<b>8</b>	<b>Article Summaries</b>	<b>111</b>
8.1	Article Overview	112
8.2	Article 1	113
8.2.1	Background, Research Aims and Questions	113
8.2.2	Methods and Research Design	113
8.2.3	Findings and Contributions	114
8.3	Article 2	115
8.3.1	Background, Research Aims, and Questions	115
8.3.2	Methods and Research Design	116
8.3.3	Findings and Contributions	116
8.4	Article 3	117
8.4.1	Background, Research Aims, and Questions	117
8.4.2	Methods and Research Design	117
8.4.3	Findings and Contributions	117

8.5	Article 4 . . . . .	119
8.5.1	Background, Research Aims, and Questions . . . . .	119
8.5.2	Methods and Research Design . . . . .	119
8.5.3	Findings and Contributions . . . . .	119
<b>9</b>	<b>Discussions and Conclusions</b>	<b>121</b>
9.1	Final Discussions . . . . .	121
9.2	Research Questions Revisited . . . . .	128
9.3	Overarching Contributions of the Thesis . . . . .	133
9.4	Limitations . . . . .	135
9.5	Further Work . . . . .	136
	<b>Bibliography</b>	<b>141</b>
	<b>Appended Articles</b>	<b>157</b>
	Article 1 - Hacking the Concert Experience: Exploring Co-Creative Audi- ence Interaction at a Chiptune Live Performance . . . . .	157
	Article 2 - Hack the Show: Design and Analysis of Three Interaction Modes for Audience Participation . . . . .	174
	Article 3 - Intelligent Music Performance Systems: Towards a Design Framework . . . . .	186
	Article 4 - Beyond Mutation: How Can We Acknowledge Symbiogenesis in Evolutionary Music Coding? . . . . .	204
	<b>Appended Questionnaires and Interview Guides</b>	<b>211</b>
	<b>Acknowledgements</b>	<b>221</b>

# List of Figures

2.1	Birnbaum et al.’s dimension space for musical devices, adapted from (Birnbaum et al., 2005, 194) . . . . .	18
2.2	Thor Magnusson’s categorization of musical instruments, adapted from (Magnusson, 2009, 208) . . . . .	19
2.3	Hauser’s visualization of postphenomenology based on Verbeek’s analyses, adapted from (Hauser et al., 2018, 466) . . . . .	23
2.4	Russell’s proposition of the circumplex model for affect, adapted from (Russell, 1980, 1164) . . . . .	26
2.5	Self-report data of Russell’s study sorted into the two-dimensional space of <i>principal components</i> , adapted from (Russell, 1980, 1174) . . . . .	27
3.1	Leman’s model of musical communication between performer and listener, adapted from (Leman, 2008, 160) . . . . .	34
3.2	Leman’s model of social musical communication, adapted from (Leman, 2008, 171) . . . . .	35
3.3	Centralized social perspective for organizing interconnected musical networks, adapted from (Weinberg, 2005, 33) . . . . .	36
3.4	Decentralized social perspective for organizing interconnected musical networks, adapted from (Weinberg, 2005, 33) . . . . .	37
3.5	Motivations for a process-centered musical network, adapted from (Weinberg, 2005, 32) . . . . .	37
3.6	Interaction Topologies in Mobile-Based Situated Networks, adapted from (Matuszewski et al., 2019, 4) . . . . .	38
3.7	Mazzanti et al.’s six-dimensional evaluation model, adapted from (Mazzanti et al., 2014, 30) . . . . .	39
3.8	Rebelo’s typology of network dramaturgy, adapted from (Rebelo, 2009, 390) . . . . .	47
4.1	The development of Glăveanu’s 5 A’s framework from Rhodes’s 4 P’s model, adapted from (Glăveanu, 2013, 71) . . . . .	52
4.2	Illustration of the 5 A’s framework as suggested by Glăveanu, adapted from (Glăveanu, 2013, 72) . . . . .	52
4.3	Illustration of the five C’s as suggested by Kantosalo Takala, adapted from (Kantosalo and Takala, 2020, 22) . . . . .	53
4.4	Kirke and Miranda’s generic model for most current CSEMPs, adapted from (Kirke and Miranda, 2021, 526) . . . . .	63

4.5	Paradigms in HCI suggested by Mueller et al., adapted from (Mueller et al., 2020)	66
4.6	Levels of human-machine embodiment as suggested by Lopes, adapted from (Lopes, 2017)	66
4.7	Interaction with digital musical instruments, adapted from (Miranda and Wanderley, 2006)	67
4.8	Human-machine interaction loop in sonic interaction design, adapted from (Franinović and Serafin, 2013, 86)	68
4.9	Human-machine interaction loop in sonic interaction design, adapted from (Bongers, 2007)	68
4.10	Creative practice and AI system design informing each other in human-machine collaborations, adapted from (McCormack et al., 2020, 42)	69
4.11	McCormack’s environmental approach to musical interaction with AI, adapted from (McCormack, 2021)	70
4.12	Interaction layers of Kantosalo et al’s framework for human-computer co-creativity, adapted from (Kantosalo et al., 2020, 59)	71
5.1	Working concept of performative co-creation	80
5.2	Overview of theories, concepts, methods, and studies of the thesis project	83
7.1	Manual annotation of the dataset using four categories	95
7.2	Plots for use of 3 low level features	97
7.3	Scatterplots for different clustering approaches	98
7.4	Loss function of MLP Regressor	99
7.5	Interaction interface for audience participation (experiment 3)	101
7.6	Performing drummer during the participatory concert (experiment 3)	102
7.7	Audience and stage at the participatory concert	102
7.8	Audience interaction data	103
7.9	Accelerometer data of the bracelet on the drummer’s right wrist	104
7.10	Cumulative plot of all interactions by audience members	105
7.11	Cumulative interaction data compared to acceleration data of right wrist	106
7.12	Cumulative interaction data compared to cumulative acceleration of right wrist	106
7.13	Cumulative interaction data compared to acceleration data of left wrist	107
7.14	Cumulative interaction data compared to cumulative acceleration of left wrist	107
7.15	Cumulative interaction data (snare decay up and down buttons only) compared to acceleration of left wrist	108
7.16	Cumulative interaction data (snare decay up and down buttons only) compared to cumulative acceleration of left wrist	108
9.1	Comparison of Quantity of Participation	122

9.2	User clusters . . . . .	123
9.3	Schematic graph of genetic algorithms for music . . . . .	126
9.4	Suggestions for symbiotic concepts applied to the context of IMPS . . . . .	126
9.5	Integration of entanglement with symbiosis . . . . .	127
9.6	Concept for further study with emotion-based audience interface . . . . .	137

# Chapter 1

## Introduction

Music matters because (...) it is a necessary part of what we are as humans, as feeling, empathetic beings, interested in and engaged with other people. To study music is to study what it is to be human – biologically, cognitively, culturally; to play music is to experience what it is to be human – physically, mentally, socially, in an aesthetic, playful, sensual context. Music matters, in short, because without it we wouldn't know who we are and what we are capable of being.

---

*Prof. Simon Frith in his inaugural lecture at the University of Edinburgh (Frith, 1998, 178)*

### 1.1 Background

Music is inherently social; researchers have even suggested that it arose as a means of bringing people together (known as the social bonding hypothesis) (Savage et al., 2021). Popular music, for example, is often performed by smaller groups of individuals in front of bigger groups of individuals; it also involves various technologies dedicated to inspiring, composing, practicing, performing, and distributing it. This PhD thesis takes an exploratory approach to its questions about how technological developments impact those concert situations in particular, and how they enable new interpersonal entanglements, social bonds, and connections to both other humans and technological representations.

The title of this thesis, *Amplified Connections*, points to the idea that the con-

nections established through musical activities can be amplified by technology. My use of the term "amplification" is inspired by Don Ihde's idea of the amplification-reduction structure that is connected to technology (Ihde, 2009). Ihde suggests that the experience of the world through technology is non-neutral and that, with "every amplification, there is a simultaneous and necessary reduction" as part of that experience (Ihde, 1979). For example, while mobile phones amplify our experience by allowing us to talk to a person in another place, they reduce the experience by turning the person into a disembodied voice or even losing them altogether in the case of a bad audio signal. Ihde notes that "the amplification tends to stand out, to be dramatic, while the reduction tends to be overlooked, or may be forgotten, particularly when the technology is truly good" (Ihde, 1979). Another effect Ihde assigns to his amplification-reduction structure is that the transformation initiated by technology "also alters what may be called 'the distance' of the phenomenon being experienced" (Ihde, 1979)<sup>1</sup>. Ihde calls his technological approach to phenomenology *postphenomenology*; while it "clearly owes its roots to phenomenology, it is a deliberate adaption or change in phenomenology that reflects historical changes in the twenty-first century" (Ihde, 2009, 5).

The subtitle of the thesis, *Exploring Performative Co-Creation and Distributed Creativity in the Context of Accessible Music Creation Technologies*, focuses on specific aspects of those amplification-reduction structures, which will be unpacked in the three framing chapters of the thesis: *Performance*, *Participation*, and *Autonomy*. To start with, the project is framed within the field of popular music performance, so it is important to establish concepts related to composing and performing music with the help of current technologies. Distributed creativity - sometimes also referred to as distributed music - will be explained in the second theoretical chapter in the context of the ways in which creative acts can involve individuals other than the music performers themselves. The last part of the subtitle foregrounds the fact that technologies for live performances can make performances more accessible, and the third framing chapter presents current forms of technology and their affects in relation to performing music.

Consequently, this project's research questions address technological backgrounds, scenarios, and forms of composition and performance involved in producing and performing music. Technology neither completely changes performance nor merely represents established ideas of performance in different ways - its impact is complicated and subtle, as is our experience with and behavior around it in co-creative scenarios. The overarching question for this PhD is as follows:

How does interactive performance technology both amplify and reduce the concert experience?
---

---

<sup>1</sup>The experienced distance is further explored from the perspective of proxemics (Hall, 1966) in article 3

## 1.2 Research Questions

This section reviews the research-leading questions of this thesis. In addition to the main question, each of the subordinate research questions will be supplemented by a short description of the area of focus the question raises together with further sub-questions that will inform the article summaries in chapter 8 and the conclusion of the thesis itself in chapter 9.

**Main question:** How does interactive performance technology both amplify and reduce the concert experience?

**Research question 1:** How do audience members and stage performers experience different concepts of audience participation at live concerts?

**Research question 2:** How does the design of audience instruments affect the participatory behavior of audience members at live music concerts?

**Research question 3:** What are the main design aspects of Intelligent Music Performance Systems (IMPS) and how can their past, current, and future developments be understood?

**Research question 4:** What are the underlying principles of IMPS and what might preferable alternatives look like?

### Research question 1

The first research question engages with the experience of audience participation scenarios from the vantage points of the audience members and the stage performers, respectively. More specifically, it addresses the individual, social, and technological aspects of those participant experiences, including *Enjoyment* (individual), *Arousal* (individual), *HCI Design* (technological), *Effectiveness and Motivation* (social), and *Surprise and Connectedness* (social). Beyond the consideration of these experiential aspects, this first area of focus also revisits concepts used for the design of participatory live performances. Consequently, the investigation is led by the following sub-questions:

How do audience members experience the performance when they are invited to participate in it musically?

How do participatory paradigms change the performance from the perspective of the performers on stage?

Which interaction paradigms are best suited to chiptune live music<sup>2</sup>?

---

<sup>2</sup>The genre of chiptune music will be explained in article 9.5

## Research question 2

The second research question engages with the use of the audience interface and the behavior of the audience members. Besides considering the interactive paradigms and the collective audience engagement, this line of questioning dwells upon the level of individual participation and looks for participatory patterns and the participation types that emerge from them. Finally, this line of questioning frames the design of participatory systems as a consequence of collective and individual participatory behavior. The following sub-questions guide this investigation:

How do different interaction modes affect collective audience behavior?

How can the audience members' contributions be structured on an individual level?

How can the findings be applied to the design of participatory systems?

## Research question 3

The third research question engages with theorizing Intelligent Music Performance Systems. Informed by the findings of the first two research questions, this question seeks the underlying concepts in the design of real-world cases of IMPS, and particularly aspects of proximity and autonomy in their constitution. Moreover, it attempts to understand not only what steers the system design but also what determines the development of future systems and how we can make those systems into accessible tools for collaborative creative engagement. The following sub-questions apply here:

What are the main aspects to be considered when designing IMPS?

What tendencies in their design appear in a review of past and current cases of IMPS?

How can we use the suggested model and conducted case study to inform speculation about future iterations of IMPS?

## Research question 4

The fourth research question centers upon alternative approaches to implementing participatory scenarios. After this thesis's investigation of the system design of IMPS in general, and audience instruments in particular, this line of questioning presents a critique of current coding principles and attempts to develop alternative approaches with audience instruments in mind. With the help of theories from symbio(gene)sis, it looks at how audience gestures could be conceptualized and dealt with in the context of audience instruments, and how to connect those adaptations to the performers' gestures within IMPS. The sub-questions here are as follows:

What are beneficial ways to simulate short and longer-term symbiotic relationships that would be helpful for musical applications?

How can we combine interdependent populations in a way that advances the creative outcome?

How can environmentally inspired concepts such as the Gaia theory be helpful?

## 1.3 Research Aims

Given that this set of research questions might appear overwhelming and even somewhat scattered in direction, it is important to restate the main aims of the project here. In general, it seeks to explore how changes in (our understanding of) technology impact our creative musical forms and our concept of music itself. The thesis project explores what happens when audience members are not only guided through a performance but also given tools to explore performance systems themselves, both individually and collectively, so as to participate in the performances in real time. Since perception is always itself an active process (Noë, 2004), this question invites a heuristic approach, and the related research aims are listed here:

- To understand the compositional affordances of participatory technology for staged musical performance.
- To observe different user behaviors of participatory technology in different scenarios.
- To understand how performers adopt and use participatory concepts in performance.
- To review and theorize performative technologies and their potential future developments.
- To explore alternative concepts of the contemporary implementation of intelligent performance technologies.

## 1.4 Thesis Outline

The thesis has two parts: the first establishes the theoretical framework, and the second presents empirical work carried out as preliminary experiments and studies that produced four published articles. The articles are then appended at the end of the thesis. The following paragraphs provide a short overview of the structure of the thesis.

### Chapter 1. Introduction

The first chapter provides an introduction to the research field and brief background on the related disciplines, the research questions, and the aims of the thesis, then briefly describes the content of the thesis.

## Part I. Theoretical Framework

### Chapter 2. Performance

The first of the three framing chapters introduces the reader to the main concepts of musical performance used in the thesis, then turns to theories of music listening

and responses to music. It then combines theory addressing musical gesture with concepts of audience response to establish the concept of *audience gesture*.

### **Chapter 3. Participation**

The second framing chapter aligns the concept of audience gesture with that of *audience instruments* by drawing upon aspects of musical communication, distributed creativity, and participatory design. It also addresses certain concerns brought over from cultural studies and discusses the limitations of audience instruments.

### **Chapter 4. Autonomy**

The third framing chapter extends the concepts of musical interaction and co-creation through a reading of technological developments including artificial creativity and autonomy within interactive music systems. Before relating those concepts to audience gesture and instruments as established in chapters 2 and 3, the chapter looks at existing models and terminology around the co-creation between humans and non-human musical agents. This chapter is a precursor to the model developed in article 3.

### **Chapter 5. Conceptual Frameworks Revisited**

This chapter revisits the theoretical perspectives presented in the three preceding chapters and discusses the resulting conceptual approaches to be used in the following empirical part of the thesis.

## **Part II. Empirical Studies**

### **Chapter 6. Research Design and Applied Methods**

This initial chapter presents the research design and quantitative and qualitative methods used in the thesis.

### **Chapter 7. Preliminary Studies**

This chapter reports on the studies undertaken as part of my PhD coursework that led to the ensuing concert studies and published articles. This chapter expands upon the theoretical input of chapter 2 (Performance) and especially the concept of audience gesture (preliminary study 1), as well as chapter 3 (Participation) and the concept of audience instruments (preliminary study 2).

### **Chapter 8. Article Summaries**

This chapter supplies an overview of the articles included in this thesis, including background, research aims and questions, research design and methods, and summary of findings, contributions, and limitations.

**Chapter 9. Discussions and Conclusions**

The last chapter discusses the findings presented in the articles and attempts to relate those findings to one another and put them in a broader context. Furthermore, it outlines the main limitations and contributions of the thesis and suggests further work on its topics.



# Part I

## Theoretical Framework



# Chapter 2

## Performance

I remind myself to listen so that I  
may be here now even though now  
has already gone.

---

*Pauline Oliveros*  
(*Oliveros, 2010, 249*)

*This chapter introduces concepts of performance with a focus on music reception, including models for music listening at concerts, and suggests the term "audience gesture" as a way to conceptualize audience response.*

### 2.1 What Is Music Performance?

In his book *Performance Rites*, popular-music sociologist Simon Frith advocates for a broad understanding of the term *performance*, one that is not limited to the staged presentation of musical text but instead draws its meaning from the analysis of the cultural context in which musical events are embedded. He suggests understanding performance as an "'experience (or set of experiences) of sociability'" that includes "'listening' itself (...) [as a] performance" (Frith, 1998, 203-204). Frith claims that performance is a "social - or communicative - process" and that performance art is a form of rhetoric, a "rhetoric of gestures in which, by and large, bodily movements and signs (including the use of the voice) dominate other forms of communicative signs, such as language and iconography" (Frith, 1998, 205). Further, he stresses the fact that "*all* live performance involves both spontaneous action and the playing of a role" (Frith, 1998, 207). Seen in the wider context of the field of performance studies, Frith could be placed in the analytical tradition of performance that Barbara Kirshenblatt-Gimblett presents as the "aesthetic communication approach" (Kirshenblatt-Gimblett, 2004, 45), a product of the Northwestern University school of thought and one of three main tendencies to emerge in the field during the 1980s (Kirshenblatt-Gimblett, 2004). The other two tendencies are the "broad spectrum approach" and "ethnoscenology." The broad spectrum approach—originating with researchers at New York University—understands performance on a basic level as

connected to human activity beyond the arts. Schechner advocates for this "happy alternative (...) to expand our vision of what performance is, to study it not only as art but as a means of understanding historical, social, and cultural processes" (Schechner, 2004, 9). Ethnoscenology, on the other hand, "rejects mind/body dualism and integrates the cognitive and the somatic." In contrast to Northwestern's paradigm, ethnoscenology does not take text as its point of departure but instead the "knowing body" and the "corporal dimension of performance" (Kirshenblatt-Gimblett, 2004, 46).

In the context of this thesis, performance is used in Frith's sense of describing the broad range of activities that occur on and around a staged music event as well as the musicians, audience listeners, and other involved parties that make this event happen. While a broader understanding of musical performance would certainly encompass practices of composing and recording music, for example, the use of the term here remains focused on real-life music events. Still, the experiments carried out in this project make use of musical elements that are both prerecorded and created in real time. This means that the delineation between performance and composition, traditionally viewed as discrete stages in the process of creating a musical work to be performed, is softened in the sense that precomposed elements are adjusted and new ones are added, so that the work is effectively recomposed on the spot in a co-creative manner.

In conclusion, the term performance is used in relation to music played from a center of attention (such as a stage), and *performers* are the individuals at that center of attention on stage (that is, they are *stage performers* unless otherwise indicated). In the case of participatory or distributed music, audience members can act as performers and access the center of attention, but the concert experiments conducted for this thesis made use of a staged performance by musicians playing in front of an audience in a conventional concert setting. This is why the divide between (stage) performers and audience members persists below, even if those audience members act like musical performers while musicking during a participatory performance.

### 2.1.1 Musicking

Within the field of music, Small develops the concept of *musicking*, which elaborates on the idea that music essentially implies "doing things." Like Frith's understanding of performance, Small's notion is not limited to activities that directly produce music, such as playing instruments, but instead encompasses the manifold activities that contribute to a musical event. Small concludes that the *meaning* of music ultimately resides among the relations between individual actions that occur during the making of music. In turn, he stresses a social dimension: the "act of musicking establishes in the place where it is happening a set of relationships, and it is in those relationships that the meaning of the act lies" (Small, 1998). These relationships need not be only those literally taking place between the performers during a performance but rather can be the "desired relationships [that] are brought into virtual

existence so that those taking part are enabled to experience them as if they really did exist" (Small, 1998, 183). Small concludes that a performance is best described thusly: "members of a certain social group at a particular point in its history are using sounds that have been brought into certain kinds of relationships with one another as the focus for a ceremony in which the values — which is to say, the concepts of what constitute right relationships — of that group are explored, affirmed, and celebrated" (Small, 1998, 183). To best unpack the workings of a musical performance, Auslander suggests the use of frame analysis (Goffman, 1974) and sees "musical performance as framed in at least four ways, which could be labeled the music frame, the performance frame, the genre frame, and the social frame" (Auslander, 2021, 91). This analytical model inspired the later empirical experiments shared in this thesis.

Musicking as a social concept is applicable to the research field of participatory concerts in that the real and the desired relationships in a performance space can be explored via instruments, which enable all the individuals who are conventionally distinguished as performers versus listeners to express their creative ideas and emotional states within that space. As those contributions are explored in turn, passively or actively, musicking adds to the abiding web of social relations in the collective performance. Unless otherwise indicated, I will be referring to *human* musicking performers and activities performed by means of *physical* bodies (see below).

### 2.1.2 Embodiment

Initially formulated as *embodied cognition* in the 1990s, embodiment in music points to the fact that the realized musical text will always depend upon a body to perform it. Likewise, in fact, the musical text to be requires a composing body with the idea for it (or, again, a performing body, in the case of improvised music). Ultimately, embodiment in music dismisses the mind-body duality in favor of the idea that we perform music as holistic beings. The question of human versus non-human "bodies" is a topic to which I will return. For people, in any case, music will always imply an embodied and enacted form of performance; it cannot exist as an abstract stream of information that is unrelated to the bodies that create and interact with it. In the field of music, scholars such as Erik Clarke (Clarke, 2005) and Marc Lehman (Leman, 2008) have developed the idea of *embodied music cognition* in the 2000s, building upon Gibson's so-called ecological approach to cognition from the 1970s (Gibson, 1979). In its forthright acknowledgment of the fundamental bodily dimension of musical activities and music perception, embodied music cognition should be understood as a performative act, as Varela et al. explain:

By using the word *embodied* we mean to highlight two points: first, that cognition depends upon the kinds of experience that come from having a body with various sensorimotor capacities, and second, that these individual sensorimotor capacities are themselves embedded in a more encompassing biological, psychological, and cultural context. By using the word *action* we mean to emphasize once again that sensory and

motor processes, perception and action, are fundamentally inseparable in lived cognition. Indeed, the two are not merely contingently linked in individuals; they have also evolved together (Varela et al., 1992, 172-173).

This entanglement of action and perception has also been described as the *action-perception loop* and has in turn informed both the playing of musical instruments and the designing of new interfaces for musical expression.<sup>1</sup> In the field of participatory performance, this action-perception loop has implications for both the individual level of the stage performers and audience members *and* the instruments and interaction interfaces used for the participatory performance. For example, when one uses a smartphone app to contribute to the music played by the performers, the visual feedback from that phone, the altered sounds and sights from the stage, and the attendant behavior of the performers and other audience members are all sources of sensory input that will influence the motor activities involved in the next participatory action. In terms of embodiment, Varela's distinction between an initial, somewhat "tighter" action-perception loop and a subsequent "looser" one (in the sense that direct interaction and feedback are complicated or even hindered by the conditions) is very relevant to participatory concert settings. This thesis is most interested in the second loop, which involves the biological, psychological, and cultural context of a concert setting; individual sensorimotor capacities, on the other hand, are important when one is designing musical instruments for individual players.

Embodiment in music performance can be regarded as one of the E's in so-called 4E cognition—embodied, embedded, enacted, and extended<sup>2</sup>—which assumes that cognition is always characterized by those four aspects (Newen et al., 2018). Along with the first two dimensions, participatory performance relies upon enactment to prompt audience members to realize the affordance of instruments and initiate a perception-action loop that shares their listening experience with others. In terms of this research project, the final dimension of extension is also central in that those audience contributions are made with the help of instruments and other technological devices—a condition to which we will return later in this chapter. Still, Burnett et al. criticize the overemphasis on extension in 4E cognition (Burnett and Gallagher, 2020) and claim that many theories are working with a simplified idea of perception as "intrinsically pragmatic or action-oriented to the claim that in the relevant cases artifacts like tools or pieces of technology are used to scaffold or offload cognition" (Burnett and Gallagher, 2020, 158). This has bearing upon the experiments for this thesis because they use smartphones as extensions of the audience's agency in displaying and enacting their musical intentions as part of a performance. Following Burnett et al., those contributions are in fact limited in many ways—for example,

---

<sup>1</sup>For more, see (Wessel, 2006) and (Noë, 2004), who dig more deeply into the terms sensation, perception, experience, and action.

<sup>2</sup>It could also be argued that the abovementioned notion of embodiment as suggested by Varela already encompasses the fact that cognition is embedded in a biological, psychological, and cultural context.

some participants might not be willing to contribute; they might be psychologically held back; or they might experience technical difficulties while using these interfaces. Bodily interactions with both musical instruments and technological devices are seldom straightforward extensions of the mind; they are instead complex processes that encompass existing knowledge and experience as well as a host of intended and unintended limitations due to instrument design, technical difficulties, and social aspects of the communal experience, among other things.

While embodied musicking always involves some kind of communication, most musical settings involving groups of performers and listeners center communication as the reason for the attendant bodily acts. While the next section therefore focuses on musical *gesture*, we will return to 4E cognition toward the end of this chapter to develop the concept of audience gesture further.

### 2.1.3 Gestures

Starting with the idea that we use our human bodies to perform music (here understanding performance to include composition, playing instruments, listening to music, and so forth), what we do with the body in this context can be theorized via the concept of musical gestures, which are at once body performances and performed bodies, one could say. Godøy defines gesture as "a movement of part of the body, for example a hand or the head, to express an idea or meaning. In the context of musical performance, gestures are movements made by performers to control the musical instrument when playing a melodic figure, to coordinate actions among musicians (conducting gestures), or to impress an audience (for example, moving the head during a solo performance). In the context of listening, such as tapping along with the beat, swaying, or dancing" (Godøy and Leman, 2010, 5). This understanding, Godøy admits, focuses on physical movement, which he calls the primary focus of the gesture, or its *extension*, "whereas the secondary focus is on *intention*, namely that which is imagined or anticipated (...). In order to call a movement a genuine gesture, it is required that this movement is in some way a carrier of expression and meaning" (Godøy and Leman, 2010, 5). Wanderley distinguishes between instrumental and ancillary gestures, the former being "applied to a concrete (material) object" and the latter described as "gestures that are part of a performance but are not produced in order to generate sound" (Miranda and Wanderley, 2006, 10).

As highly contextualized carriers of meaning, gestures are often used to convey intention to other performers (here used in the broadest sense to include listeners). Sound-related gestures reveal that "musical communication is fundamentally driven by movement. In particular, musical communication is steered by bodily motion, which is encoded by the player, transmitted through audio, and decoded by the listener" (Godøy and Leman, 2010, 6). Godøy ultimately frames gesture as a cognitive category by defining it "as a pattern through which we structure our environment from the viewpoint of actions. Gesture conceived that way is thus a category, or structural feature, of our perception-action system. In this approach, gesture is both a mental and a corporeal phenomenon" (Godøy and Leman, 2010, 8). For a more

comprehensive discussion of embodied gestures in music, the reader is referred to Tomás et al., who illuminate the topic from a variety of perspectives (Tomás et al., 2022).

Musical gestures in the context of participatory performance often appear simultaneously: while stage performers use gestures as part of their work, audience members complement their own ongoing instrumental gestures with ancillary listening gestures. For example, they might attune their body movements to some aspect of the stage performance while at the same time abiding by certain collective parameters in their own space among their peers, as we will later see <sup>3</sup>. Before we return to listening gestures and performer-audience interaction, we will consider some aspects of musical instruments and technologies that are critical to conveying musical gestures during participatory performances.

### 2.1.4 Instruments

The act of playing a musical instrument is often at the core of musical performance. But what are musical instruments in the first place? What role do they play in the context of embodied music cognition and musical gestures? And how can they be theorized with regard to the technologies penetrating and surrounding them?

An early means of categorizing acoustic musical instruments was the Hornbostel-Sachs system, which classifies them as idiophones, membranophones, chordophones, and aerophones (Von Hornbostel and Sachs, 1914). These categories were later complemented by the so-called electrophones, which use loudspeakers as sound-producing entities instead of physical materials or air (Kvifte, 1989)<sup>4</sup>. With the rise of digital technology, yet another category arose: digital musical instruments (DMI). The DMI uses "computer-generated sound (...) and consists of a control surface or gestural controller, which drives the musical parameters of a sound synthesizer in real time" (Miranda and Wanderley, 2006, 1). Jensenius contests this term, arguing that any DMI must also have a sound-producing entity (such as a corpus or loudspeaker) to transmit its digital information to the physical world by making it audible. He prefers the category *electro-acoustic* instead (not to be confused with the genre of electroacoustic music) (Jensenius, 2022, XI). Most recently, the more inclusive category "new instruments for musical expression" (NIME) has evolved since the early 2000s into a viable designation in turn closely connected to the NIME conference. As part of the first NIME conference, instrument designer Perry Cook proposed a set of ten principles for NIME design (Cook, 2001). These principles are closely connected to instruments as interfaces to computers and many of them

---

<sup>3</sup>Mazzola theorizes the merging of gestures by multiple musicians as "hypergestures." While he claims that gestures are semantic generators rather than semantic carriers, he insists that a network of multiple gestures should always be seen as a hypergesture (Mazzola et al., 2011, 244-245).

<sup>4</sup>For an in-depth discussion of those categories and the belonging instruments the reader is directed to (Campbell et al., 2004). Kvifte argues for a less precise definition in the favor of acknowledging the multidimensionality in the concept (Kvifte, 2008, 55)

impact their musical use<sup>5</sup>.

Thor Magnusson proposes a taxonomy of instrumental functions in three instrument classifications—acoustic, electric, and digital—as shown in Fig. 2.2. This taxonomy is inspired by the abovementioned Hornbostel-Sachs system, which is based on the way instruments physically produce sound. Magnusson applies this to his digital musical instrument classification by way of the categories physical controllers, patches, and simulators. Examples of such instruments include sensor-based systems, Max/MSP patches, or single software solutions such as Cubase or Reason (Magnusson, 2009, 209). In this way, they can be regarded as discrete tools for creating music. Magnusson suggests two other taxonomies of musical instruments according to activities and interactions, respectively. Activities relate to stages in the musical process such as composition, performance, or analysis and editing, whereas interactions depart from his notion of "metaphors of action." In traditional instruments, these would be musical gestures such as stroking, blowing, plucking, and so forth, whereas in the digital realm, these would be grouping tracks using software, navigating a computer game (which produces sound), or interacting with artificial-intelligence music systems (Magnusson, 2009, 211).

In contrast to Birnbaum et al.'s dimension space focused on the musical control of digital instruments, as depicted in Fig. 2.1 (Birnbaum et al., 2005), Magnusson outlines an epistemic dimension space for digital musical instruments with eight parameters: *integrated composition* (the degree of musical composition inscribed in the tool), *expressive constraints* (the limitations of the tool's design), *autonomy* (the degree to which the tool affords expressivity), *music-theoretical mass* (the amount of culturally specific music theory contained in the tool), *required foreknowledge*, *creative simulation* (whether the tool is novel or instead intentionally imitative of established tools or practices), *improvisation* (the degree to which the tool lends itself to free improvisation), *generality* (how open in expression the instrument is and how well it copes with musical situations), and *explorability* (how much depth the instrument offers) (Magnusson, 2009, 213-214). Ultimately, Magnusson supports the idea of *instrumentedness* introduced by Bertelsen et al. that challenges the belief that human-computer interaction (HCI) designs should offer transparent and clear points of interaction to their users, and that "can be seen as a HCI perspective addressing how the interface mediates the user's relation to the software and to the outcome of the interaction" (Bertelsen and Bødker, 2003). In other words, it is more important that digital instruments are exploratory and able to adapt to the user's general relationship to the instrument than that they execute predictable outcomes.

---

<sup>5</sup>While DMIs reflect a longer tradition of development and use than NIME and encompass digital versions of traditional instruments such as digital pianos as well, they are nevertheless considered to be part of the NIME class. This is because digital interfacing converts instruments with acoustic paradigms into "new instruments" which afford novel means of "musical expression." Of course, NIME also include interfaces that do not rely upon digital interaction paradigms. In addition, what is considered "new" will eventually give way to its integration into musical performance practice, at which point it becomes something else. DMI, NIME, and their particularities for the performance space are further discussed in [article 3](#).

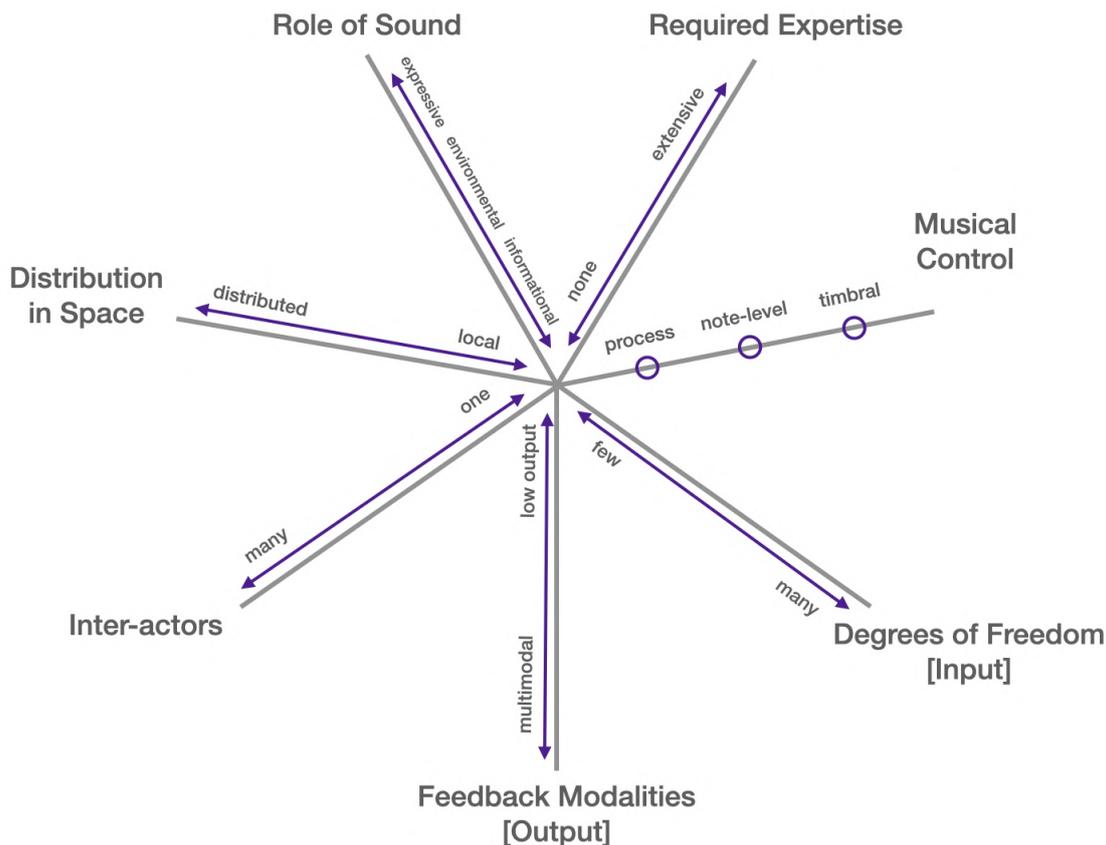


Figure 2.1: Birnbaum et al.'s dimension space for musical devices, adapted from (Birnbaum et al., 2005, 194)

Rebecca Fiebrink also stresses the importance of this human-machine relationship over the goal of offering the user maximum control, and this seems especially relevant to AI music instruments (Fiebrink, 2017, 3). Her supervised machine learning-based system *Wekinator* uses the principle of "showing" the system gestures that the user needs to produce certain sounds (or other media). This training data allows the learning algorithm to then map similar gestures to the intended outcome or create combinations of them and so forth. Fiebrink thinks of such interactive machine-learning tools as design tools for composers that offer the benefit of a layer of tangibility (including an action-perception loop) during the composition process rather than simply an abstract programming interface (Fiebrink, 2017, 12). In contrast to designing instruments by creating direct action-to-sound mappings, which can be time-consuming, her supervised learning approach is much faster and offers designers and users of those instruments unforeseen results and a level of surprise that can be helpful to the creative process and also applicable in performance (Fiebrink, 2017, 9). One important aspect in the latter regard is that *Wekinator* uses on-the-fly learning, which makes it a valid real-time instrument for performance. In this case, the most fitting category might be *meta-instrument*, to capture the machine-learning layer that exists for the complex mappings that the system learns during the training process (Fiebrink et al., 2009).

Musical instruments have journeyed over millennia from basic acoustic tools to

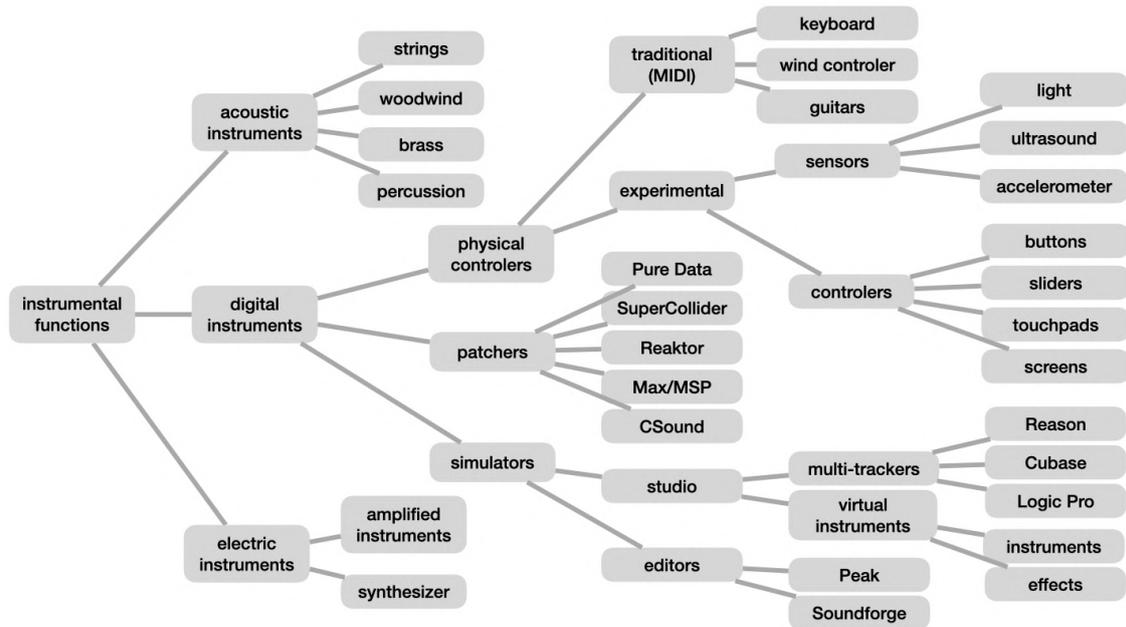


Figure 2.2: Thor Magnusson’s categorization of musical instruments, adapted from (Magnusson, 2009, 208)

highly complex machines. A key design aspect of DMI now seeks to move beyond the bodily affordances and gesture-to-sound mappings and account for how performers of these new instruments relate to them. This aspect has also informed acoustic instruments, of course (Doğantan-Dack, 2022), but now it has taken on renewed importance, such that HCI is now turning into the "human-computer relationship" that ties into the emotional and aesthetic understanding of the user rather than the instrument’s various functional refinements or elaborate action-to-sound mappings. This relationship appears likely to govern the emerging classes of instruments that will amplify inherent aspects of, or even penetrate, the human body—for example, brain-computer interfaces use electrodes or other brain-recording devices for sound generation, so that the performer does not consciously control the instrument as such but instead allows it access to some of their bodily processes for music making. Related instrument types include musical wearables such as sound-generating clothing, wristbands, or other gesture-recording sensors, or musical tools that use muscle tension to generate sound and give tactile feedback or even actuate the muscle to cause contractions, as in the so-called electric muscle stimulation (EMS). Pedro Lopes suggests the term human-computer integration instead of human-computer interaction (or, presumably, human-computer relationship) since these instruments overlap with the body rather than interact with it as distinct entities (Lopes et al., 2015). The next step could be instruments that become part of the human body, including implanted technology such as the Neuralink<sup>6</sup>. While their impact remains to be seen, it is clear that our relationship with and use of technology will take new

<sup>6</sup>The Neuralink interface uses a brain implant that directly connects to nerve cells to control the sensorimotor system and, as such, could be used for playing and learning instruments. For more, see <https://neuralink.com>.

new forms that transcend the traditional input-output paradigm. Before I return to the musical performance space to take a closer look at audience perception and listening, I will supply some general background on human-centered technologies with the aim of applying some of this theory to the field of musical performance.

### 2.1.5 Technologies

In his 1977 essay *The Question Concerning Technology*, Heidegger advocates for developing a theory of technology via questioning rather than defining it. By doing so, he writes, we "should like to prepare a free relationship to it. The relationship will be free if it opens our human existence to the essence of technology"<sup>7</sup> (Heidegger et al., 1977, 3). This essence, he continues, "is by no means anything technological," and "we we are delivered over to it in the worst possible way when we regard it as something neutral" (Heidegger et al., 1977, 4). His starting points for a critique of the technological are what he calls the *instrumental* (technology as means to an end) and the *anthropological* (technology as a human activity) definitions, which, he insists, belong together (Heidegger et al., 1977, 5). Heidegger's interest in rejecting the causality of technology leads him to conclude that the essence of technology is something *revealing* rather than instrumental as such:

Instrumentality is considered to be the fundamental characteristic of technology. If we inquire, step by step, into what technology, represented as means, actually is, then we shall arrive at revealing (...) Technology is therefore no mere means. Technology is a way of revealing. If we give heed to this, then another whole realm for the essence of technology will open itself up to us. It is the realm of revealing, i.e., of truth (Heidegger et al., 1977, 12).

With regard to modern AI, this definition aligns well with the fact that we put together lots of training data in order to arrive at an essence via datasets that would be too big to explore manually. Especially using non-supervised techniques, machine learning can reveal the essence of the data, giving us clusters and other forms of representation, as we will see later.

Furthermore, Heidegger continues, the word *technē*, stemming from the Greek, is closely related to the word *epistēmē*, in that both mean knowing in the broadest sense. Further, *technē* "is a mode of *alētheuein*. It reveals whatever does not bring itself forth and does not yet lie here before us, whatever can look and turn out now one way and now another" (Heidegger et al., 1977, 13). In relation to nature's resources, the revealing force of modern technology is claimed to be "the challenging" (Heidegger et al., 1977, 14), in the sense, for example, that humans challenge the exploitation of natural resources using technology. Thinking about this in the realm of music technology reminds me of a talk I recently attended at the AI and Musical

---

<sup>7</sup>Interestingly, Heidegger sees this essence as something that pursues its own course, or the way in which something remains what it is through time and is "enduring as presence" (Cerf, 1961, 59) as cited in (Heidegger et al., 1977, 3)

Creativity conference in Brighton, when CJ from Databots suggested that one way to respond to the perceived threat of AI to people making music is to embrace the challenge and show AI that we can do better than it can (Carr, 2023).

The last theoretical concept in Heidegger's essay is the self-revealing as standing-reserve, or "Ge-stell" (Enframing): "Enframing means the gathering together of that setting-upon which sets upon man, i.e., challenges him forth, to reveal the real, in the mode of ordering, as standing-reserve" (Heidegger et al., 1977, 20). Enframing is used in the sense of not only challenging but also producing and presenting "in the sense of *poiēsis*" (Heidegger et al., 1977, 21). Heidegger's definition can be seen as a useful counterweight to the currently popular action-based theories and anthropological definitions that focus on human activities rather than the "standing-reserve." Heidegger, on the other hand, assures us that the "merely instrumental, merely anthropological definition of technology is therefore in principle untenable" (Heidegger et al., 1977, 21). The technological advances yet to come could be seen as just such a standing-reserve—an Enframing that is yet to be revealed by human activity (of instrument designers, performers, and so forth). In the end, Heidegger concludes that the "essence of technology is in a lofty sense ambiguous. Such ambiguity points to the mystery of all revealing, i.e., of truth (...) The question concerning technology is the question concerning the constellation in which revealing and concealing, in which the coming to presence of truth, comes to pass" (Heidegger et al., 1977, 33). In 2023, this statement is rather anachronistic in the way it reduces technology to a methodology for revealing things that are already there. Whether technology really is only a "revealer" and not an actor that influences the course of things, art, and life, and whether the standing-reserve is simply another way to articulate the destiny-driven, ultimately deterministic view of technology, remain up for debate.

In his work *Understanding Media: The Extensions of Man*, McLuhan argues for media technology as an extension of humans, in the sense that "extension, whether of skin, hand, or foot, affects the whole psychic and social complex" (McLuhan, 1964, 4). McLuhan's analyses mainly concern media and as such are inspired by mediation technologies, yet his view is arguably human centered at heart: "[The] medium is the message. This is merely to say that the personal and social consequences of any medium—that is, of any extension of ourselves—result from the new scale that is introduced into our affairs by each extension of ourselves, or by any new technology" (McLuhan, 1964, 7). With regard to the societal impact of technology, McLuhan is a technological determinist who views technology as an ultimately uncontrollable force that directly affects (even threatens) society. John Culkin summarizes this position quite succinctly: "We shape our tools and, thereafter, our tools shape us" (Culkin, 1967, 53). Since the late 19th century, technological determinism has informed many techno-philosophical discourses, and its Marxist adherents include Horkheimer and Habermas. Today, especially with AI entering the mainstream media, it is experiencing a revival.

Technological constructivism, on the other hand, tries to see technology through the eyes of its human users and as part of their social systems, which are subject

to change at any time. Rather than at the mercy of or victimized by technology, humans are positioned as though they were in control of it, including its autonomous manifestations. Central texts in the 1980s such as *The Social Shaping of Technology* (MacKenzie and Wajcman, 1985) and *The Social Construction of Technological Systems* (Bijker et al., 1987) gave rise to the discipline we refer to as SCOT (Social Construction of Technology) today. Magnusson argues that those texts are also connected to socio-cultural interactionism, since they stress the fact "that technology and the context in which it is used are always already interdependent" (Magnusson, 2009, 54). He concludes that technology has always interacted with the human body and mind, and as such it should be seen as something that inherently conditions our thinking (Magnusson, 2009, 76). The thesis of SCOT, which was "strongly influenced by the British social construction of science" in the 1970s, is "to reject determinism, and demonstrate that 'it could have been otherwise'" (Eglash, 2006, 333). Books such as Pierre Lemonnier's *Technological Choices* (Lemonnier, 2013) remind us that humans design, interact, and co-evolve with technology rather than simply watch it grow outside their control. Eglash notes that social constructivism is "mainly regarded as a framework of the postmodern era" though only a few of the aforementioned authors would regard themselves as postmodernists (Eglash, 2006, 334). Postmodernism, he continues, "takes the 'semiotic turn' in analyzing material objects as sign systems", and it is only authors such as Haraway (Haraway, 2006) who point to "technology as a 'material-semiotic hybrid' " (Eglash, 2006, 335). A similar challenge to any postmodern theory of technology is offered by Latour et al.'s famous Actor-Network Theory (ANT), which goes beyond constructivism by questioning the longstanding divide between nature and society, or, in other words, things and humans.

ANT states that "things are also actors (or at least actants) which must be recruited as allies, refuted as enemies, or otherwise dealt with in the web of relations that constitute scientific and technological development" (Eglash, 2006, 335). Latour therefore allows technology itself "to be a voice in the narrative. Thus the constructionist simply gives more causal emphasis to the social side, whereas a description in terms of ANT undermines the very division" (Eglash, 2006, 335). ANT as a theory of relational materialism points to the limitations of any individual actor and focuses on the negotiation between the different interests of multiple actors as a critical part of the process of creation in science and art (Sismondo, 2010, 69). Consequently, these single aspects or actors are not to be ordered a priori as good or bad; instead, technological innovations take place via a "multi-dimensional development" (Sismondo, 2010, 70). While ANT clearly embraces ecological (and embodied) approaches and sees technology as material culture, it has been criticized for failing to grasp the translations (or interactions) between different human and non-human actors, recalling, in effect, the logical positivism of the 1930s (Sismondo, 2010, 73). Especially with regard to the field of human-computer interaction in the context of music, there might be better approaches to understanding those relationships.

As already introduced in the first chapter, Don Ihde has developed a now well-known approach that combines the ideas of embodied cognition with technology

studies. As a phenomenological perspective (that is, one framed according to the experience of consciousness), it incorporates two distinct parts. Its first phenomenology is the immediate experience of a new phenomenon, such as a new instrument or interactive process, wherein there is sensorial input to which our senses react. Its second phenomenology is the subsequent phase of reflection—that is, when "the conscious mind starts to assign it [the new experience] names, expectations, bias, understanding, memories, labels, perspectives, subjectivity" (Vear et al., 2023). Equally, Vear argues that a music performance is a double embodied process: first, human and instrument form a layer of embodiment; second, this process happens inside the music performance (or musicking) (Vear et al., 2023). Ihde calls his phenomenology *postphenomenology* to account for what technical advances do to our bodies and how technology relates to the world: "Postphenomenology is the concrete and empirical study of the social and cultural roles of technologies in human existence and experience" (Hauser et al., 2018, 460) Fig. 2.3 shows a visualization of the idea of postphenomenology by Hauser that is based on Verbeek's analyses of the mediation implied by technology (Verbeek, 2005). With regard to how we view our world, in short, "technologies work to amplify and reduce human perception and experience, and invite and inhibit human action and practices" (Hauser et al., 2018, 466).

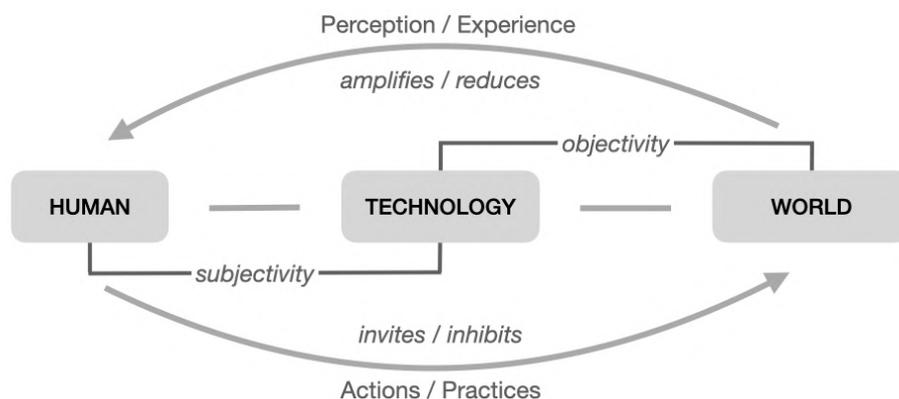


Figure 2.3: Hauser's visualization of postphenomenology based on Verbeek's analyses, adapted from (Hauser et al., 2018, 466)

Just as performers are known to develop emotional relationships with their musical instruments, they could also do so with technologically more advanced tools as well. The postphenomenological view of the world through technology insists upon the possibility of more emotional interactions with its tools. As mentioned above, Bickmore and Picard propose the term human-computer relationship rather than of human-computer interaction because we are now interacting with systems that "understand" and address our emotional states (Bickmore and Picard, 2005).

Lastly, our relation to technology must accommodate embodied cognition because these tools are now as likely to be a literal as a figurative part of us. Helena De Preester suggests three categories of technologies that extend the human body: limb extensions (technologies that primarily seek to alter our motor capacities), perceptual extensions (technologies that primarily seek to alter our sensory capacities),

and cognitive extensions (technologies that primarily seek to alter our cognitive capacities) (De Preester, 2011, 121). She also makes an important distinction between *extensions* and *incorporations* (the latter being more profoundly integrated into the human body) as she differentiates among these categories: "In the case of limb prostheses, the crucial factor is a change in the feeling of body ownership. In the third case, cognitive prostheses, it is also a feeling of ownership that is at stake, but this time it is about ownership of thoughts. In the second case, perceptual prostheses, a change in subjective perceptual experience is the central issue" (De Preester, 2011, 135). The divide between extension and incorporation perhaps recalls the line drawn between posthumanism (technology is an extension of humans) and transhumanism (technology is a part of humans). Applied to the case of performer-audience interaction, Haraway's notion of performing cyborgs has interesting implications for the future interplay between stage performers, audience members, and autonomous mediators. (Haraway, 2006).

While the preceding introduction engaged with the main aspects of musical performance (Musicking, Embodiment, Gestures, Instruments, and Technologies), the next part of this chapter will focus on the *listening* aspect of musical performance and introduce central theories concerning musical listening, the emotional response to music, and aspects of social music cognition before returning to the concept of audience gestures in the conclusion.

## 2.2 How Do We Listen to Performance?

What happens when we listen to a musical performance? What are the layers of listening and perceiving? How do the environment and the social surroundings impact that experience? And what are the active parts of the listening process? If we were to understand the listening process better, would there be a way to "record it" and use it to train systems to become creatively useful to both musical performers and audiences?

### 2.2.1 The Ecological Approach

As already mentioned earlier in the subsection "Embodiment" 2.1.2, Gibson's ecological approach to perception offers an action-oriented understanding of the act of listening (Gibson, 1977) (Gibson, 1979) that Clarke has adapted to the act of listening to music (Clarke, 2005). Relatedly, Jensenius concludes that there "is no such thing as a passive listener" and insists that listeners recognize *action-sound mappings* when listening to a performance (Jensenius, 2022, XV). This underlines the listeners' sensitivity for active gestures<sup>8</sup> and extends to theories of contagion and empathy, as we will see later in this chapter.

---

<sup>8</sup>This sensitivity to performative gestures is often mentioned in the same breath with the hypothesis of the so-called mirror neurons, which activate when we perceive another human in action. Cox has discussed this phenomenon for the case of music with her mimicry hypothesis and uses the term physical empathy (Cox, 2011).

In terms of the ecological approach, a musical performance affords certain ways for listeners to react to it. Gibson proposes that the "affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. The verb to afford is found in the dictionary, the noun affordance is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment" (Gibson, 1977, 127). For example, when a song comes to an end, audience members are afforded the ability to applaud. At the same time, of course, this affordance is simply a learned ritual in the sense of Bourdieu, in that we have already agreed to this behavioral norm in attending the performance in the first place. Affordances are therefore limited or at least influenced by ritualized behavior, and rituals themselves can be seen as affordances in a performative context. Further, musical performances afford emotional responses as a listening behavior, depending on the audience member's personal state. The next section introduces relevant models for our emotional responses when listening to music.

## 2.2.2 Emotional Response

Music is said to evoke emotional responses in humans, and several models have been developed to unpack this particular affect. Because this thesis's articles rely on those models to capture the audience members' responses in the questionnaire, I will endeavor to explain them here.

Today, studies of affective response regularly differentiate between *felt* and *perceived emotion* (Sloboda and Juslin, 2001), or emotion in an *internal* or *external* locus (Eerola and Vuoskoski, 2012, 310). The former is what an individual listener feels as a product of their background, emotional state, or other factors beyond those introduced by the music to which they are listening. Perceived emotion is the emotion the listener thinks is intended by the artists behind the music to which they are listening (the composer or performers, for example). Most of the models presented below refer to perceived emotion, not felt emotion, because the latter can differ profoundly among individuals and groups.

### The Circumplex Model for Affect

One often-used model for collecting emotional responses is derived from Russell's circumplex model for affect, which was not specifically developed for the field of music but works quite well as a self-report-based means of collecting the overall music-induced emotional states of one's subjects (Russell, 1980). Fig. 2.4 shows the model in its original form., and Fig. 2.5 adds certain self-reported affection words along its axes. This dimensional model has also been adapted by the Geneva Emotion Wheel (see, for example, (Sacharin et al., 2012)). The circumplex model and its offshoots have been criticized, however, for their inability to capture the mixed emotions that reside at different points on one or more of the model's dimensions (Hunter et al., 2008).

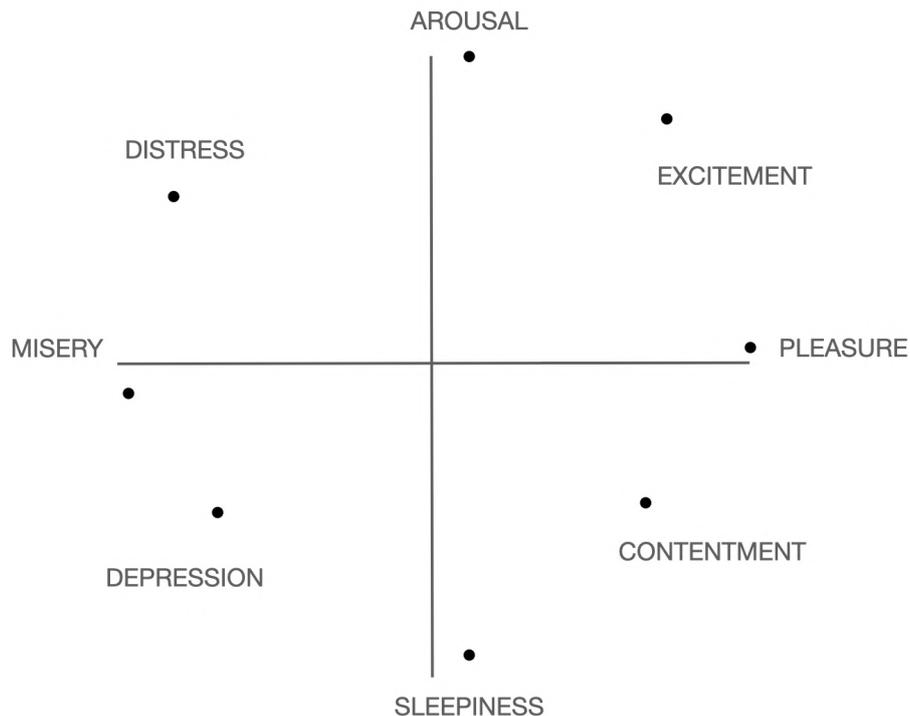


Figure 2.4: Russell's proposition of the circumplex model for affect, adapted from (Russell, 1980, 1164)

### Other Models for Emotional Response

Eerola and Vuoskoski have suggested four categories of emotional model for musical studies of emotion: discrete, dimensional, miscellaneous, and music-related (Eerola and Vuoskoski, 2012, 309). Discrete models of emotion suppose that all emotional responses can be derived from a basic, finite set of emotions, whereas miscellaneous models for emotion in music "include a motley collection of emotion concepts such as intensity, preference, similarity, tension, or some other construct closely linked to emotions in general" (Eerola and Vuoskoski, 2012, 310). One miscellaneous model that has been used for musical studies is the Geneva Emotional Music Scale (GEMS) by Zentner et al. (Zentner et al., 2008). This model relies upon nine emotion categories: tension, power, joy, wonder, tenderness, peacefulness, transcendence, nostalgia, and sadness (Zentner et al., 2008, 5). Based on fMRI measurements, these categories were grouped into four quadrants but only partly fit into the circumplex model. Instead, the nine initial emotions fit into three main classes (Vitality, Unease, and Sublimity) (Zentner et al., 2008, 13). While the circumplex model of affect supplied the basis for the questionnaires used in the empirical studies executed as part of this thesis, the GEMS model informed the further development of the interaction technology used for those studies.

### Methods for Measuring Emotional Response

When it comes to retrieving emotional responses, the "self-report approach has provided the cornerstone of the results in music and emotion research" (Eerola and Vu-

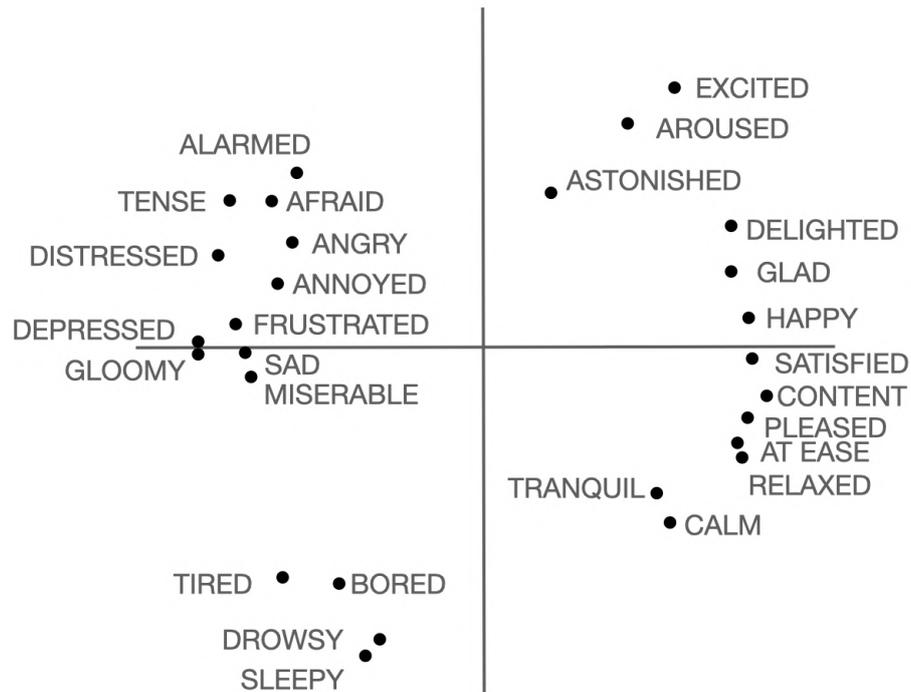


Figure 2.5: Self-report data of Russell's study sorted into the two-dimensional space of *principal components*, adapted from (Russell, 1980, 1174)

oskoski, 2012, 314). In addition, biological approaches have "become frequent, but the analysis methods and self-report methods have not developed at the same pace" (Eerola and Vuoskoski, 2012, 314). In Eerola and Vuoskoski's opinion, "nonverbal, behavioral measures are too infrequently used to measure music-induced emotions" (Eerola and Vuoskoski, 2012, 324). These measures would include the observation of audience behavior during performances. In the end, of course, any method must capture "the dynamic aspects of music-induced responses" rather than reducing them to static measurements or statements, and this "would require more sophisticated time-series analyses" (Eerola and Vuoskoski, 2012, 314). It is not easy to accomplish this in listening experiments, even if some continuous self-report measures have already been employed, such as for example those studied by (Sloboda and Lehmann, 2001). Yet it remains a necessary approach because "music unfolds in time, and emotional experiences also change as a function of time" (Eerola and Vuoskoski, 2012, 323). Notably, almost all of the 251 studies reviewed by Eerola and Vuoskoski "have been conducted in laboratory settings" and "do not see the laboratory experiments as problematic for perceived emotion studies, as emotion recognition is probably little affected by the artificiality of the music listening situation" (Eerola and Vuoskoski, 2012, 323). This position is particularly counterproductive in performance studies, which must account for their actual contexts, especially given that contemporary music concerts are often staged as immersive, multimodal experiences whose visual components contribute to the musical stimulus. In our own work, we used retrospective self-reports but measured audience engagement on a continuous scale for the interactive parts of the performance, as will be seen in article 1 and 2.

### 2.2.3 Social Music Cognition

So far, emotional response has been framed as an individual reacting to a musical stimuli in the context of their individual background and reading of the musical content, environment, and passing mood. But how do people listen differently in social surroundings such as performance spaces, concerts, or other group settings? Is there such a thing as a collective emotional response in an audience? If so, how do groups affect the individual's listening experience?

#### Contagion and Empathy

Musical contagion describes the arousal of emotions evocative of the ones perceived during the act of listening to music, which "leads the listener to mirror the emotional expression internally" (Juslin, 2019, chapter 20). Hatfield et al. define emotional contagion as the "tendency to automatically mimic and synchronize facial expressions, vocalizations, postures, and movements with those of another person's and, consequently, to converge emotionally" (Gibbs, 2010, 5). Juslin argues that musical contagion is especially applicable to the expressive performance (rather than the musical composition), but that it is *not* limited to facial or bodily gestures and can also occur through auditory perception alone (Juslin, 2019, chapter 20.2). This contagion seems connected to empathy in the sense that the more a listener is empathetic in their listening (or, in other words, affectively open), the more likely it is that they will mirror or mimic the emotions expressed internally. Still, Maibom distinguishes contagion from empathy by pointing out that contagion relates to the expression itself, whereas empathy applies to the people performing the expression. She speaks of artists as persona who exhibit their emotions through the work of art, and it is to these persona that we are emotionally empathetic (Maibom, 2020, 119). This empathy, in turn, might foster contagion in the sense that once we are empathetic, it becomes easier to be moved by the creative expressions of those persona.

#### Synchronization and Attuning

Leman understands empathy not only as an internal emotional openness or involvement in another's emotional and expressive world but also as one of "three levels of corporal imitation: synchronization, attuning, and empathy" that "all involve imitation, but in different degrees of participation and identification" (Leman, 2008, 112). Synchronizing to music has been shown to have a positive emotional effect in individuals, and this is also true for *interpersonal* synchrony, which "can serve as a cue for prosocial behavior: when we move together in time, we become more likely to trust and cooperate with one another, and to like and help each other" (Juslin, 2019, chapter 19.5). Synchronizing to a beat (tapping along to it, for example) is regarded as a low-level resonance of the motor system, while embodied attuning "implies corporeal movements in accord with music, such as (...) singing along with the music" (Leman, 2008, 115). In contrast to synchronization, which is "based on low-level sensorimotor activity, attuning aims at addressing higher-level features such as

melody, harmony, rhythm, and timbre, or patterns related to expressiveness, affects, and feelings." As such, attuning could be regarded as "a kind of playing together with the music" that "involves a more active role for the subject and an engagement in higher-level intentional processes" (Leman, 2008, 115). Leman concludes that attuning "brings the human body into accordance with a particular feature of music. It can be seen as navigation with or inside music" (Leman, 2008, 115).

## 2.2.4 Performance as Staged Emotion

While it seems clear that the above forms of entrainment (in the sense of *mutual adaptive behavioral resonance* (Leman, 2008, 171)) "occur in response to music," it "remains to be demonstrated fully (...) that this mechanism is able to arouse *emotions* in music listeners" (Juslin, 2019, chapter 19.4). When audience members synchronize their movements while dancing or attune to the music while singing along, a strong emotional bond develops. Performance might then be seen as a form of *staged emotion* constructed by the composer, executed by the performer, and offered to the audience to tune into (or attune to). In contrast to other forms of emotional response that happen spontaneously, the emotional response to a concert is partially predefined and often something the audience anticipates. Nevertheless, audience members can respond to the performance's emotional affordances in different ways by mirroring, developing, or even rejecting them. Juslin wonders what types of affect are induced by musical entrainment and whether that entrainment acts to regulate the audience's arousal levels (for example, when members are rhythmically entraining to a groove). Of course, entrainment can "be a source for both excitement and calm, two common emotional reactions during musical events" (Juslin, 2019, chapter 19.5). Juslin considers an entrainment-induced emotional response to be a *secondary response* and notes that entrainment between individuals has been reported to have positive effects of "social bonding" (Levitin, 2008) and "feeling connected" (Juslin, 2013).

In April 2023, I took part in a study by Doug VanNort's group at York University that included a performance of Pauline Oliveros's work titled *Tuning Meditation* (1971). The group of eight people, connected by a JackTrip server, tried to match one another's humming voices by listening closely while singing. It turned out to be a very special, even spiritual experience for the performers. Oliveros coined the term *deep listening* to capture this form of focused and committed listening while actively performing and eventually tried to develop deep-listening machines. This attuning of individuals in groups of either ensemble players or audience members has been labeled *allelo-imitation*, wherein "each subject's corporeal imitation of her or her neighbor" generates "global effects of group behavior" (Leman, 2008, 111). These behavioral attunings of audience listeners at musical performances will be referred to as *audience gestures* in what follows.

### 2.2.5 Audience Gestures

Recall Godøy's aforementioned definition of the gesture from Sec. 2.1.3 as "a movement of part of the body, for example a hand or the head, to express an idea or meaning" (Godøy and Leman, 2010, 5). If we extend our understanding of embodied music cognition to the social realm of an audience, we can view the audience gesture as a collective expression of meaning by multiple bodies that are responding to performed or staged emotional affordances. In light of Leman's ideas about corporeal imitation, this response could be both low-level (that is, simple synchronization such as clapping, dancing, jumping, and so on) or high-level (that is, attuning via singing along and so on).

#### Embodied Gestures for Participatory Performance

Audience gestures depend on a variety of individual, collective, and environmental factors, and participatory music systems initiate both lower and higher levels of audience response to staged performances. Often, low-level responses such as synchronizing the body to the beat, dancing, or clapping along arise less consciously than high-level responses such as attuning the voice or moving sliders in an app in relation to what the musicians are doing on stage. In addition, the physical audience gesture is not the same kind of straightforward intentional body movement reflected by the stage performer's musical gesture. Audience gestures derive from the *collective* state of the audience and contribute to what is often referred to as the "atmosphere" at a concert event. This atmosphere is ultimately determined by many individuals responding to the staged performance based upon their personal and musical backgrounds, their emotional constitutions, their relations to one another, the staged situation and conditions, and, of course, the performance itself. These various conditions and their impact upon audience gesture and atmosphere bear measurement and analysis, given the appropriate means of doing so. And those means, in turn, can afford creative participation in their own right, acknowledging once and for all that listening is an active (enacted) process rather than a passive state to be simply observed.

Audience gestures, that is, are not simply responses to performers but also purposeful expressions of one's likes or dislikes as well as other kinds of meaning. For example, clapping or booing indicate a collective verdict (but not nuanced feedback from any given individuals). As mentioned above, audience gestures also result from group effects that need not be music- or concert-related. Still, audience gestures are most often directed at someone or something—the overall performance, individual stage performers, the environment or situation surrounding the concert, or other members of the audience (or the audience as a whole). In short, they can be performance directed, peer directed, or co-creation-directed. As we will see, this profile is very relevant to the design of audience-participation instruments.

## Audience Gestures as Compositional Material

From the perspective of composition, audience gestures can be used in a number of ways. In a traditional concert setting, a composer could try to incorporate the audience's response into the work—applause at the end of a movement could be used to transition into the next movement or blended directly into the music or shifted to the performers on stage as part of the score. In genres such as electronic dance music, for example, audience responses to earlier sets might enable a DJ to shape a set differently next time.

Audience gestures can be produced through bodily reactions such as dancing or other corporeal synchronization, clapping or other hand gestures, or singing or other vocal utterance, any of it aligned with others in the audience, either simultaneously or consecutively (as a wave, for example). Audience gestures can also be generated by instruments or other technological devices—LED wristbands can signal enthusiasm or occasion synchronized movements during a concert, while small handheld instruments or personal smartphones can contribute to the sonic output<sup>9</sup>. Smartphones as audience instruments bring with them both opportunities and limitations during a concert, especially if they are also serving as data recording devices. I will return to this at the end of the next chapter after a consideration of musical interaction in performance in section 3.2.2.

## Audience Gestures and 4E Cognition

This last part returns to embodied cognition and gesture in the context of audience studies with a focus on participatory composition. In particular, the notion of 4E cognition, as presented in section 2.1.2, provides an interesting framework for conceptualizing participatory performance.

Embodied cognition in participatory performance recognizes the fact that audience members' corporeal presence and interaction inform their emotional response and their participatory behavior during a performance. Embedded cognition recognizes the environment of the performance space, including the venue and staging of the performance and the social settings impacting the audience, as well as the acoustic features of the performance space. Enactive cognition recognizes the fact that the use of audience gestures to actively participate in a performance furthers the engagement and impacts how the co-creation is experienced. Lastly, extended cognition recognizes how the individual experience of a performance can extend to digital interfaces that incorporate other audience members' gestures, in turn enabling more connected and collaborative experiences.

Audience gestures seen through the lens of 4E cognition provide a good base framework for audience members' responsive behaviors at the individual level (though embedded cognition does acknowledge the social context of the performance space). To shed light upon the design of audience instruments and audience-performer interaction in general, chapter 3 will introduce musical communication and interaction

---

<sup>9</sup>Clément et al. investigated gestures with handheld music devices and classified them for a participatory performance (Clément et al., 2021).

more generally, then return to audience instruments at the end.

## 2.3 Conclusion

In this chapter, I discussed performance as an active, embodied, and multifaceted phenomenon that has incorporated technology in various ways. I stressed the importance of the act of listening as an embodied, cognitive part of musicking, or the total of all the activities that occur during a performance. I introduced gesture as a vehicle for communicating embodied musical ideas to other musical agents, whether instruments, other musicians, or audience members. I then argued that this is a two-way process between performers and audience members and introduced the concept of *audience gestures*. In short, I conclude that the idea of the musical gesture as presented by Godøy and others should be extended to the responses of the audience and framed as an interaction among gestures rather than a one-sided coding/decoding process. In this way, we will be able to better recognize the audience's agency and understand performer-audience interaction as a core phenomenon of musical performances.

# Chapter 3

## Participation

...there is nothing more difficult to grasp than social ties.

---

*Bruno Latour*  
([Latour, 2005](#), 159)

*This chapter will discuss concepts of musical communication and distributed creativity among performers and audiences. It analyzes models for distributed creativity and participatory instruments and addresses concerns about the general concept of distributed music.*

### 3.1 How Do We Interact during Performance?

Now that we have arrived at the concept of audience gestures as attuned, collective ways of responding to performances under the influence of group listening effects, we will now develop that idea from a communication perspective. We will begin with a general account of musical communication.

#### 3.1.1 Models for Musical Communication

In their article about intra-ensemble communication, King and Gritten distinguish between the terms dialogue, communication, and interaction with regard to music in particular. Dialogue, they suggest, refers to "[d]eliberate simultaneous exchanges of information and evaluations (verbal, non-verbal) between two (or more) musicians which may reflect communication among other facets of conversation." Communication is a "[d]eliberate one-way process of discourse between two (or more) musicians as illustrated by turn-taking." Interaction is a "[d]eliberate two-way process of mutually constructive discourse between two (or more) musicians as illustrated by reciprocity and co-constructed musical identities" ([King and Gritten, 2017](#), 307). In most of the empirical work connected to this thesis, interaction best describes the two-way processes that are present, but it is also true that communication is used to describe interactive processes that exchange information by ways other than

turn-taking, as we will see in the following model by Leman.

Leman's model of musical communication introduces musical mediators to the communication between performer and listener (see Fig. 3.1). A mediator (such as a musical instrument) is controlled by the performer "in a closed loop with haptic, sonic, and perhaps visual feedback," and "in the mind of the performer, this physical interaction can be enhanced by corporeal imitation processes that translate the sensed energy back into the action-oriented ontology, giving meaning to the interaction" (Leman, 2008, 161). The mediator transmits the underlying intended articulations of the performer via "sonic and visual energy to the listener, who, through mirror processes, can make sense of it" (Leman, 2008, 161). Finally, "the listener's processing of musical information is likely to be reflected in corporeal articulations, which can be seen by other listeners" (Leman, 2008, 161). The arrow on the bottom right of the graph does not feed back into the communicative model, but the bio-mechanic energy of the listener could be considered part of the communicative process as well. Once captured and made available to both performers and audience members, as we will see, it would be accessible creatively as well.

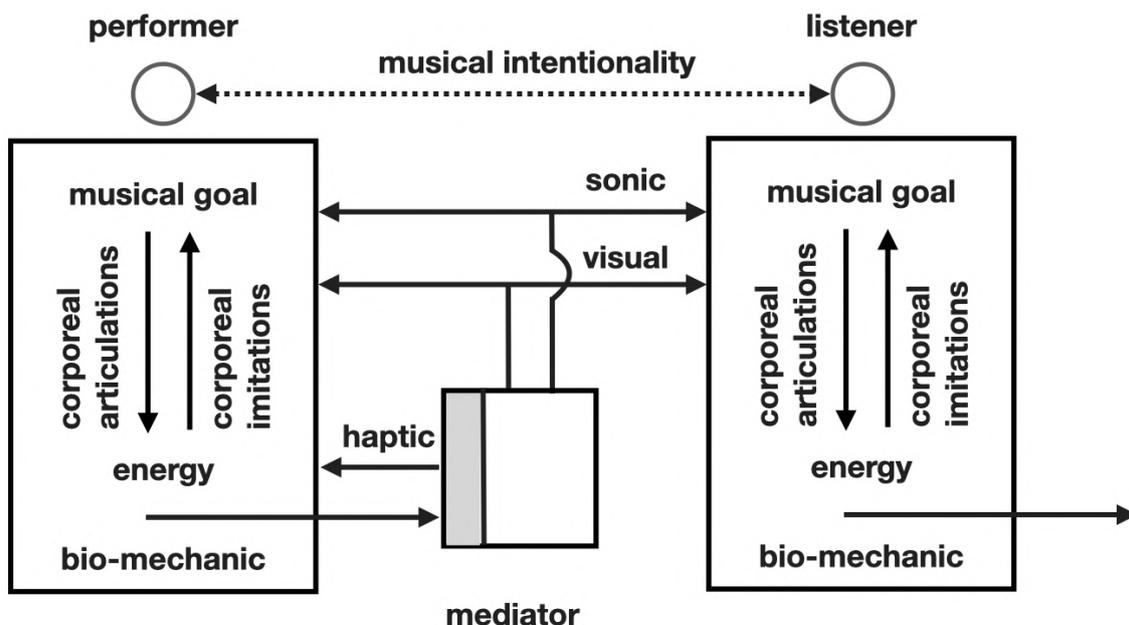


Figure 3.1: Leman's model of musical communication between performer and listener, adapted from (Leman, 2008, 160)

Leman also offers a model of social interactions at the more global scale of the entire performance that involves performer and listener groups (see Fig. 3.2). He distinguishes between four interaction channels that can be multimodal—that is, include "sonic, visual, olfactory, and tactile information" (Leman, 2008, 171). Interaction channel (1) refers to the performers generating music that is transmitted to the listeners as interaction channel (2) and "may drive the movements of the listeners in response to the music" (Leman, 2008, 171). Interaction channel (3) describes the influence listeners have on one another, as described in section 2.2.4, here also called *social entrainment energy*. Interaction channel (4) points to the

"musical/sonic energy, haptic feedback, and social entrainment energy coming from the audience" (Leman, 2008, 172), and influencing the performers. Leman notes that "performers report that the audience tends to be perceived as a gestalt whose action is likely to have an effect on their playing," and that they perceive the effects of allelo-mimetism "as a global emerging effect of the audience, which can be very stimulating for their performance" (Leman, 2008, 172). Forms of entrainment in this context<sup>1</sup> can produce "musical magic, a peak experience of a group of people" that is likely "related to empathy, which is strongly associated with feelings of social connectedness" (Leman, 2008, 172).

Leman considers "the development of interactive music systems that can deal with group effects of social musical cognition to be one of the major challenges of future music research" and suggests that interactive music systems "could be designed to enhance this type of social resonance communication, in contexts where many subjects are confronted with many machines." These machines, in turn, become "the social agents with which it becomes possible to exchange intentions" (Leman, 2008, 172-173). He ultimately finds "mediators that incorporate mirror processes (...) of great potential interest for interactive music systems" (Leman, 2008, 173) – an observation that also inspired this thesis project.

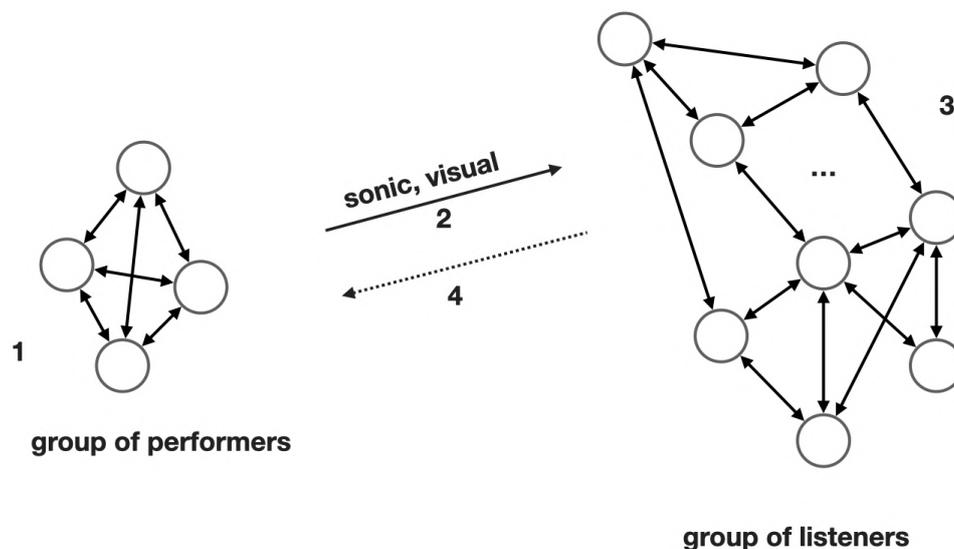


Figure 3.2: Leman's model of social musical communication, adapted from (Leman, 2008, 171)

<sup>1</sup>To add an example for a different context, entrainment was also used more specifically in relation to rhythm. Clayton proposes three levels of musical entrainment: intra-individual (occurs within a person), inter-individual (occurs between individuals in a group), and inter-group (occurs between different groups)(Clayton, 2013, 30). While this study frames rhythmic entrainment as a product of independent rhythmic processes that eventually become synchronized, these levels could be adapted to participatory performances, including the entrainment effects between the respective groups of stage performers and audience members.

### 3.1.2 Distributed Music Systems

Since Leman published his book in 2008 (and, in fact, for many years before that), artists, composers, programmers, and system designers have attempted to do exactly what his model describes: develop interactive music systems—"from musical instruments to music environments"—that "open a new world of expressive possibilities in art." They seek these systems despite the persistent challenge of adapting the "electronic equipment to the action-oriented ontology of the human subject, so that technology-based mediation can create an illusion of non-mediation." This "illusion may form the basis of an interaction between minds (real/or artificial), which is the ultimate goal of musical communication" (Leman, 2008, 182-183). Other interesting questions include how the individual audience contributions are dealt with, how they relate to the staged performance in question, and who controls what at what point in the performance. Below, we will review the most prominent approaches to distributed music systems.

#### Structural Approaches

Gil Weinberg separates *interconnected musical networks* into two main categories of social organization: centralized and decentralized. The "main conceptual axes at play when designing a 'social philosophy' for a network are the level of central control desired and the level of equality provided to the different participants in the interaction" (Weinberg, 2005, 33). Centralized networks, as depicted in Fig. 3.3 , "tend to be governed by a computerized hub that is responsible for generating the musical output based on input from participants," whereas in decentralized systems (see Fig. 3.5), "players communicate directly with each other through instruments or software applications that are computationally self-contained" (Weinberg, 2005, 33). To describe the social structures implemented in systems, Weinberg suggests that political metaphors (democratic, monarchic, decentralized, and anarchic) "may be appropriate" (Weinberg, 2005, 33).

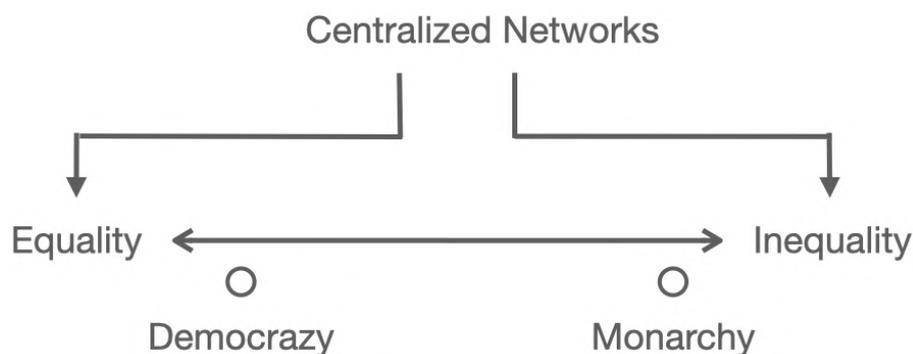


Figure 3.3: Centralized social perspective for organizing interconnected musical networks, adapted from (Weinberg, 2005, 33)

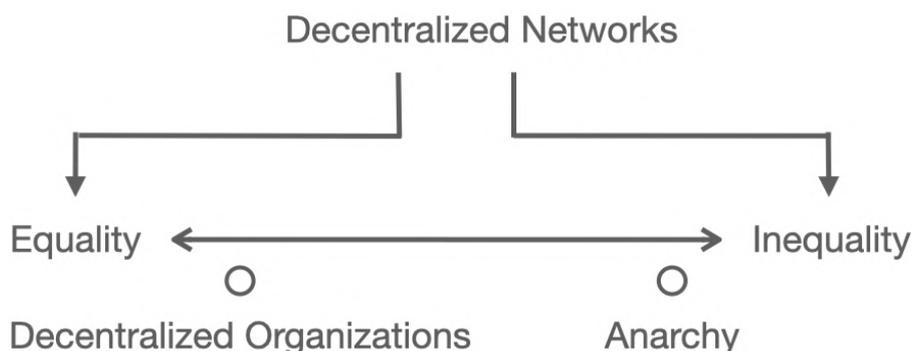


Figure 3.4: Decentralized social perspective for organizing interconnected musical networks, adapted from (Weinberg, 2005, 33)

Weinberg also discerns two kinds of participant *activities*—structure centered and process centered—while noting that most networks "combine process and structural elements in different levels and balances" (Weinberg, 2005, 32). He sees process-centered networks built for creative experiences as either goal oriented (that is, "designed to encourage players to achieve specific objectives, musical or non-musical") or exploratory (that is, they "do not impose specific directions or goals for the participants" (Weinberg, 2005, 32). Further, he observes that in goal-oriented democratic systems within which participants "would have to collaborate to create a noticeable and significant musical effect (...) only the collaborative act of the majority defines the final musical result" (Weinberg, 2005, 33). He continues, "In decentralized systems, on the other hand, interaction occurs directly between participants without a central control to govern the experience" (Weinberg, 2005, 33). This type of interaction can lead to the extreme of the anarchic network, which provides "minimal central control and maximum freedom for players to generate and manipulate their musical material" (Weinberg, 2005, 33).

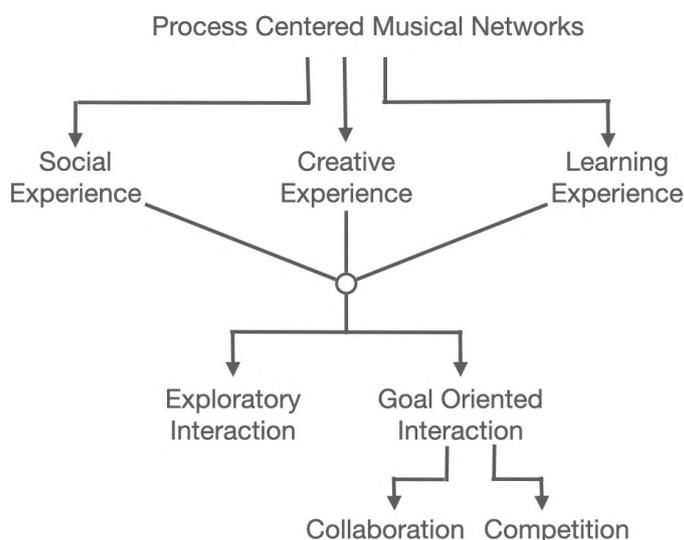


Figure 3.5: Motivations for a process-centered musical network, adapted from (Weinberg, 2005, 32)

Renwick, on the other hand, categorizes network music systems based on the strategies they apply: local music networks, composition support systems, collective creation systems, telematic performance, distributed performance, live coding and internet-based systems, and shared sonic environments (Renwick, 2017, 28-58). Matuszewski et al. present six topologies for participatory composition as depicted in Fig. 3.6 that relate to the social structure of the participatory systems (Matuszewski et al., 2019). The idea of topologies proposes "a set of basis figures that can be used to describe several levels of interactions without focusing solely on technical aspects," and the researchers' "graphs represent the actual possible interaction between each entity, human and technical artifacts" (Matuszewski et al., 2019, 4). Mazzanti et al. suggest a six-dimensional model for evaluating audience-participatory system designs as depicted in Fig. 3.7 (Mazzanti et al., 2014). He includes the audience's activity and the distribution and transparency of the audience interaction as three of those dimensions.

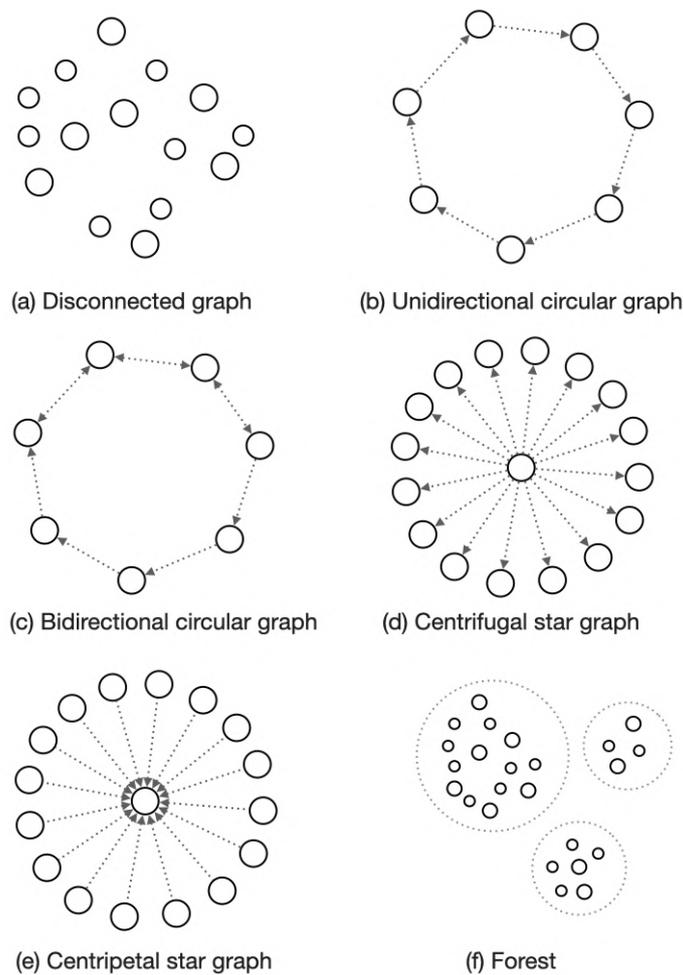


Figure 3.6: Interaction Topologies in Mobile-Based Situated Networks, adapted from (Matuszewski et al., 2019, 4)

The systems applied in the experiments in this thesis and later articles make use of a performative center and in that regard could be considered centralized networks (even though Weinberg's central hub is a computer rather than a group of

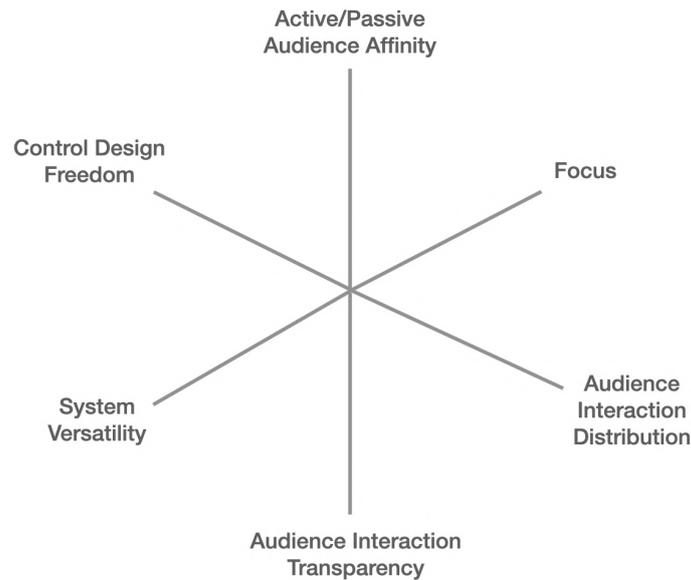


Figure 3.7: Mazzanti et al.'s six-dimensional evaluation model, adapted from (Mazzanti et al., 2014, 30)

stage performers). Yet, they also offer exploratory interactions in terms of sound and light parameters that do not advance specific goals as such (though it is a goal of sorts for individual audience members "to make the light red and move it towards the audience" or "to put maximum delay time on a synthesizer." Seen outside of the stage performers, the audience's contributions could also be framed as a decentralized network with anarchic tendencies, given that the individual contributions are limited only by their step-size for changing parameters and there is no intermediary layer for controlling the output of the system. The audience's contributions are intended to inspire the stage performers and offer a layer of novelty for both performers and audience participants. Lastly, the connections among performers and audience members are considered to be different in strength, given that the physical distances among all of the participants involved in a performance play a role in terms of both performing on stage and resonating within the audience. For example, the rhythm section on stage has stronger ties while they are playing, and a group of friends attending a performance might interact more intensely because they would be standing closer together than a group of strangers would.

### Lean-in Consumption and the Idea of Prosumers

On a more general note, scholars have observed a trend among music consumers toward more active engagement since the 2000s with the rise of social media and other participatory technologies. Attali anticipated this trend as early as 1985 and, upon the arrival of the internet, recategorized these active consumers as "prosumers" (Attali, 1985). Wikström has discerned a similar trend toward active participation (not necessarily in a real-time concert environment) among music fans (Wikström, 2009, 147-169). From a music business perspective, Mulligan has measured this trend and what he calls its "lean in" (as opposed to "lean back") experiences. He argues

that "AI is competing for lean back, while (...) virtual artists compete for lean in (fandom)" (Mulligan, 2023). Here, AI is understood as generating background music, but AI in live interactive performance is understood as navigating sonic spaces, training models with examples and other data (for example, interaction data), and using those models for search algorithms. Clearly, all of the above examples reflect the trend of active participation via technology that has emerged in recent decades.

### 3.1.3 Concerns around System Design and Cultural Critique

Creating an interactive experience for audience participants can be seen as a process of system design, or more specifically the system's compositional design. As Perry Cook suggests in *Principles for Designing Computer Music Controllers*, the time-based art of music demands "a piece, not an instrument or controller," which should produce "instant music, subtlety later" (Cook, 2001, 1). This appears to advise us to create affordances rather than think about systems as rigid structures (Norman, 1999). On a more general design note, Norman suggests in his book on the design of everyday things that it is the designer's mental model and system image that interact with the user's mental models (Norman, 1988) as read in (Biles, 2007, 31). Similarly, in the case of musical instruments, the design work takes place at the level of musical ideas rather than the level of physical interfaces.

Lee and Freeman found that despite their system design, which featured a master musician to pre-define the musical keys, audience members felt more connected to the instrument players, even though they did not interact with them directly via the system (Lee and Freeman, 2013, 454). This likely reflects the fact that the corporeal articulations of the acoustic musician are more relatable than the abstract (albeit more musically impactful) interaction-design aspects of the musical interface. They also found that transparency was a key design factor for audience musicians to feel a sense of collaboration during the participatory experience (Lee and Freeman, 2013, 454). To identify issues in the design of technology-mediated audience participation (TMAP) in live music, Hoedl et al. developed 16 design criteria from the results of a questionnaire offered to potential audience members: "skilfulness, expressiveness, diversity, objects, readiness, masterfulness, exposure, communication, openness, creativity, distinctiveness, obtrusiveness, sociability, mood, appropriateness, and contradiction" (Hödl et al., 2017, 33). I will return to questions of system design in more detail in both articles 9.5 and 9.5. Below, I will address one of the biggest problems of participatory systems: the bandwidth bottleneck.

#### The Bandwidth Bottleneck

In his 2010 article "Web-based collaboration, live musical performance and open-form scores," Freeman raises general concerns about stage performers' real-time collaboration with live audiences, given the fact that the "technical, logistical and design constraints at live venues, including simultaneous participation and single-stream output, make such live participation systems challenging to create and deploy effectively" (Freeman, 2010, 149). As an alternative, his article "instead focuses

on linking web-based collaborative experiences with more traditional live musical performances" (Freeman, 2010, 149). He presents two hybrid musical compositions where "website users participate before the performance and out of real time, with ample opportunities to practice, develop and edit their creative ideas" (Freeman, 2010, 150). Freeman's argument centers upon his hypothesis that the use of internet-based social listening, composition, and improvisation applications by concertgoers before a performance would enable them to "better understand musical structure, gesture, composition and performance" (Freeman, 2010, 151). He also points out the practical realities of live musical performances that "can make it challenging to incorporate ideas from the social web within the concert hall," as well as the fact that "participants have no opportunity to edit or revise their contributions, nor can they delve into the structure of participation by exploring the evolution of the music out of real time" (Freeman, 2010, 152). He notes that many real-time participation pieces use only one output stream for their many contributions even though such a many-to-one scenario poses a "critical design challenge" in the need to balance "individuality and coherence" (Freeman, 2010, 153). All of this can lead to a reduction in "the depth and transparency of the creative collaboration with the audience" (Freeman, 2010, 153).

In the end, he accepts the reality that "live performance venues impose daunting technical, logistical and economic constraints" and looks for alternatives to real-time collaboration with audiences (Freeman, 2010, 153). Freeman's later works make use of what he calls the *open-form score*, which is offered to audience composers in advance of a performance to create the material that performers then interpret on stage. He acknowledges his frustration with the fact that concertgoers "rarely visited the website in advance of the concert, and they sometimes understand little about the open-form score when they first hear the piece in concert" (Freeman, 2010, 153).

While I applaud the activation of audience members as co-composers of a musical work, I am not surprised by the challenge of involving them in advance of a performance (even if the technology for doing so is as accessible as a smartphone). As Freeman notes above, attentive listening during the event is more important than a traceable contribution in the score. The desired sense of audience collaboration in the event does not have to be driven by the audience members' composing or performing skills. Still, Freeman's compositions are appealing because they incorporate audiences at the level of composition rather than appointing them to a role in the performed work. The question then becomes: Is the design of the concept and interface for the participatory work itself participatory or does it remain, in the end, the case that a single composer controls the musical event? Our work related to this thesis tries to keep the interaction parameters as simple as possible in order to allow audience members a level of co-creation and hopefully some investment in a participatory performance that offers a higher level of design control as well.

### Cultural Critique of Participatory Systems

The concept of interactive technology, or interaction with technology in a more general sense, has been critiqued from various angles. Saltz describes interactive technology as "one of the hot concepts of the 1990s" and wonders whether "interactivity, or the interactivity we see in computer art, [is] a new phenomenon, and more specifically, is it new to art?" (Saltz, 2003, 395). He disputes Lucy Lippart's claim that interactive computer art sustains a general dematerialization of art, countering that "the work of art must project some kind of stimulus - sound, image, kinetic movement - back into the world of the audience to perceive" (Saltz, 2003, 396). He concludes that "a participatory interaction is performative when the interaction itself becomes an aesthetic object; in other words, participatory interactions are performative to the extent that they are *about* their own interaction" (Saltz, 2003, 404). The extent to which a work is performative is a function of the relationship "between live interactor and computer-controlled media" since "each work of interactive computer art establishes a particular kind of relationship" (Saltz, 2003, 399).

Media critic Paul A. Taylor questions the "excessively uncritical attitude to new technologies" and worries that the "majority of social science literature dealing with hacking and hacktivism is premised upon the conventional wisdom whereby technology is deemed to be essentially neutral," whereas there is actually only "the myth of neutrality" (a term initially coined by Langdon Winner) (Taylor, 2013, 247). Taylor discerns an "experiential narrowness of computer-based interaction" and concludes that "interactivity serves a ubiquitous, yet equally generally unacknowledged, ideological function" that from "critical theory's perspective (...) should more accurately be labeled 'interpassivity'—a situation in which technology undermines human agency in the very guise of claiming to enhance it" (Taylor, 2013, 247-248). This interactivity, he argues, "has become synonymous with point-and-click digital mediations," and those hackers and hacktivists "who proactively engage with digital technology risk becoming false alibis for a more general loss of autonomy in the rest of society's technologically mediated interactions—meta-level interpassivity" (Taylor, 2013, 248). He continues, "not only do we over-rely upon technologically savvy groups to compensate for our own techno-deficiencies, but we also fail to recognize the more negative, disempowering aspects of those groups' own experiences" (Taylor, 2013, 248). Ultimately, Taylor critiques mediated interactions by echoing "Jean Baudrillard's dogged insistence that, in practice, contemporary communications systems fabricate non-communication" (Taylor, 2013, 249).

Another critical voice from the field of cultural studies arrives with Nicolas Bourriaud's book *Relational Aesthetics* (Bourriaud et al., 2002). It is mainly concerned with approaches to interactive exhibition art from the 1990s but offers relevant insights into other art forms as well. For example, Bourriaud defines relational art as "an art taking as its theoretical horizon the realm of human interactions and its social context, rather than the assertion of an independent and *private* symbolic space", which he sees as consequence of the growing urbanization of the mid-20th

century and, in particular, its "system of intensive encounters" (Bourriaud et al., 2002, 5). This system, he claims, "has ended up producing linked artistic practices: an art form where the substrate is formed by intersubjectivity, and which takes being-together as a central theme, the 'encounter' between beholder and picture, and the collective elaboration of meaning" (Bourriaud et al., 2002, 5). He worries that the general mechanisation of social functions that evolved with the development of communication tools "reduces the relational space," and that contemporary art in this sense is "developing a political project when it endeavours to move into the relational realm by turning it into an issue" (Bourriaud et al., 2002, 6). He proposes that a relational aesthetics "does not represent a theory of art. This would imply the statement of an origin and a destination, but a theory of form" (Bourriaud et al., 2002, 6). He sees form itself as an unstable concept and suggests that technologies may enable the human spirit to recognize "other types of 'world-forms' (...) for example, computer science put forward the notion of program, that inflect the approach of some artists' way of working" (Bourriaud et al., 2002, 8). He ascertains that "forms are most often defined as an outline contrasting with content. But modernist aesthetics talks about 'formal beauty' by referring to a sort of (con)fusion between style and content, which is why we ought to talk of 'formations' rather than 'forms'" (Bourriaud et al., 2002, 8). Ultimately, as "part of a 'relationalist' theory of art, intersubjectivity does not only represent the social setting for the reception of art, which is its 'environment', its 'field' (Bourdieu), but also becomes the quintessence of artistic practice" (Bourriaud et al., 2002, 9).

The artist's practice or behaviour as producer, Bourriaud continues, "determines the relationship that will be struck up with [their] work. In other words, what [they] produce (...), first and foremost, is relations between people and the world, by way of aesthetic objects" (Bourriaud et al., 2002, 19). In relational aesthetics, the sphere of inter-human relations becomes a genre in itself with which artists can be associated, and Bourriaud determines that it "arises from an observation of the present and from a line of thinking about the fate of artistic activity" and is neither the "revival of any movement, nor the comeback of any style" (Bourriaud et al., 2002, 20). Relational art, he explains, "encompasses in its working process the presence of the micro-community that will accommodate it. A work thus creates, within its method of production and then at the moment its exhibition, a momentary grouping of participatory viewers" (Bourriaud et al., 2002, 27). Bourriaud sees a need to reinterpret the idea of community in the face of modernity's strive for individualism, noting that reinventing "the idea of plurality (...) means inventing ways of being together, forms of interaction that go beyond the inevitability of the families, ghettos of technological user-friendliness" that are offered by institutions; in our "post-industrial societies, the most pressing thing is no longer the emancipation of the individual, but the freeing up of inter-human communications, the dimensional emancipation of existence" (Bourriaud et al., 2002, 27). That is why he talks about the "community effect in contemporary art," arguing that its "aura (...) no longer lies in the hinter-world represented by the work, nor in the form itself, but in front of it, within the temporary collective form that it produces by being put on show"

(Bourriaud et al., 2002, 28). Because "the technological machines of information operate at the heart of human subjectivity," we must learn to "seize, enhance and reinvent" subjectivity, for "otherwise we shall see it transformed into a rigid collective apparatus at the exclusive service of the powers that be" (Bourriaud et al., 2002, 41).

## 3.2 How Can We Participate in Performance?

As we saw in the previous chapter, interactive participatory real-time settings can be challenging, and there are valid concerns and limitations with regard to the use of interactive technology in the first place. Still, most of the researchers mentioned here remain convinced that we should continue to work on concepts, conduct experiments, and reframe the theoretical backgrounds of interactive artistic experiences in terms of ongoing technological developments. Important questions persist: How can we meaningfully incorporate the audience's creative engagement into a real-time performance? And how can we enable participants to share in collective affective responses and open relational forms that incorporate audience gestures (in the sense of audience instruments)?

### 3.2.1 Distributed Creativity

Critical to these considerations is the notion of *creative engagement* itself and, in turn, the ways in which we might creatively engage more people. We will therefore begin by reviewing understandings of social creativity, then move on to the concrete work of audience-participation experiments. Sociologies of creativity often highlight the word *distribution*—for example, we hear of distributed creativity (Sawyer and DeZutter, 2009), distributed performance (Renwick, 2017, 46), and distributed dramaturgy (Rebelo, 2009, 390). The hope that everyone can seamlessly contribute to a collective performance is idealistic at best. What, then, does the distribution of "the creative" even mean, how might this agency be shared, and what does it mean to perform in a distributed setting?

We will start with a short review of the concept and study of creativity before focusing on musical, distributed, and artificial creativity in particular.

Guildford was a pioneer in the researching of creativity as a general human trait in the 1950s (at least the Western understanding of the term); previous to that, creativity was reserved for the impenetrable workings of geniuses. Guildford modeled creativity (Guilford, 1950), as 150 mental capacity factors connected to five basic operations (divergent production, convergent production, cognition, memory, and evaluation), also known as the Structure of Intellect Problem-Solving, or SIPS model (Guilford, 1967). Inspired by this link between creativity and divergent production, the widely recognized Torrance Test of Creative Thinking then began to evaluate the aspects of fluency, flexibility, originality, and elaboration as measures of creativity in individuals (Torrance, 1966)<sup>2</sup>.

---

<sup>2</sup>The Torrance test was recently done by ChatGPT-4, which scored in the top percentiles (Guzik

"Given the general failure of divergent thinking and cognitive trait models to account for creative productivity," Brown concludes (Brown, 1989), more inclusive models of creativity arose during the 1980s. Amabile's component framework theorizes creativity as "behavior resulting from particular constellations of personal characteristics, cognitive abilities, and social environments" (Amabile, 1983), rather than from a set of personality traits.<sup>3</sup> (Getzels, 1987). While most researchers acknowledge that creativity depends on social environments, it has long been associated with the creative *individual* in particular. Only recently have some begun to suggest that creativity is a social phenomenon in the first place.

Sawyer, for example, insists that to "explain creativity, we must not only *include* sociocultural approaches; in many cases, we must *begin* with them" (Sawyer, 2011, 279). Performance is more social than composition, but even composition is a lot more social than we usually realize" (Sawyer, 2011, 355). Sawyer sees the term *distributed creativity* as derived from distributed cognition, which is used in cognitive science to describe "situations where collaborating groups of individuals collectively generate a shared creative product" (Sawyer and DeZutter, 2009, 82). Based on his studies of music and theater performances, Sawyer discerns three main characteristics of group creativity: improvisation, collaboration, and emergence. Improvisation is the real-time process that happens "in the moment of the encounter"; collaboration addresses the way in which "members contribute and their interactional dynamics result in the performance"; and emergence describes "collective phenomena in which, as it is said, 'the whole is greater than the sum of the parts,'" and the product is "unpredictable, contingent and hard to explain in terms of the group's components" (Sawyer, 2006, 148). The latter also shades into something called *collaborative emergence*, which "is more likely to be found as a group becomes more aligned with the following four characteristics: (1) The activity has an unpredictable outcome, rather than a scripted, known endpoint; (2) There is moment-to-moment contingency: each person's action depends on the one just before; (3) The interactional effect of any given action can be changed by the subsequent actions of other participants; and (4) The process is collaborative, with each participant contributing equally" (Sawyer and DeZutter, 2009, 82). Sawyer wonders whether an "understanding of the interactional mechanisms that occur when creativity is distributed" would "contribute to our understanding of group creativity" (Sawyer and DeZutter, 2009, 82). He also suggests that distributed creativity is highly dependent on the "characteristics and constraints of the performance genre"—constraints that can even derive from "prior interaction in the ensemble activity" (Sawyer, 2003, 86-7)<sup>4</sup>. In line with "Csikszentmihalyi and Howard Gardner, who have advocated for a 'systems approach' that combines individualist perspectives with analyses of the social organization of creative fields and the symbolic structure of creative do-

---

et al., 2023)

<sup>3</sup>Still, Amabile's iterative, motivation-based framework was criticized for its adherence to the idea of creativity as related to problem-solving (Brown, 1989, 27) rather than problem-finding, as suggested by Getzels

<sup>4</sup>as read in (King and Gritten, 2017, 311)

mains," Sawyer advocates for a "sociocultural approach" stressing the "necessity to incorporate group-level laws and concepts" in order to avoid the reductionist views of individualist psychology (Sawyer, 2019, 584). Consequently, he extends Csikszentmihalyi's concept of creative flow to the group level via the term *group flow* (Sawyer, 2003, 43). Note that group flow is related to a real-time performance process rather than a compository one—Sawyer states, "the performance is its own goal. In improvisational creativity, the process is the product, and the researcher is forced to focus on the creative processes of group creativity." (Sawyer, 2003, 5). This recalls the differences discussed above between creativity in composition and creativity in performance, the latter of which is more improvisational due to its real-time character. Mazzola refers to them as "two forms of musical creation" and concludes by distinguishing between *compositional creativity* and *performative creativity* using the creation of a lead sheet in jazz (compositional) and its (improvisational) interpretation as respective examples (Mazzola et al., 2011, 233). Could audience instruments be similarly regarded as audience "lead sheets"<sup>5</sup>

Rebelo approaches the notion of distributed creativity by combining the idea of a network with dramaturgy as a performative concept. He notes that networks as complex models of interaction were understood in the 18th century as connectivity tools featuring multiple knots (or nodes) that had to be solved mathematically (Euler, 1741). Today, the word *network* is strongly associated with the internet, and it has been investigated as part of economic inquiries (Castells, 2010) and analyses of power distribution, among other things (Deleuze and Guattari, 1988). Meyas advocates for the significance of the space between the nodes as opposed to that of the nodes themselves (Mejias, 2007). Regarding real-time performance and the network, Rebelo emphasizes the importance of the place of performance, which is culturally specific and drives shape-giving decisions about what goes on, and cautions us that to "think that the network will do without these basic questions of artistic practice is to romanticise the technology" (Rebelo, 2009, 388). He suggests viewing the network itself as a culturally specific space for relationships, adding that "dramaturgy provides a useful framework for addressing these relationships" and a "method for understanding artistic practice in the network" (Rebelo, 2009, 388). Still, he also acknowledges the importance of the morphology of the network space to these interactions - it is a "key element in the reorganization of power relationships" (Rebelo, 2009, 388) - and proposes three broad distributed dramatic structures of performance networks derived from issues of "authorship and collaboration as well as presence and environment" (Rebelo, 2009, 389)<sup>6</sup>. He labels

---

<sup>5</sup>Compositional versus performative creativity is also connected to the given musical genre and its compositional and performative traditions—think of the differences between jazz and electronic music in these regards, for example. As we will see in article 9.5, the musical genre involved in the concert study demands designs that satisfy its socio-cultural conventions to ensure the approval of the participants.? Before returning to such questions below, we will first look at different conceptions of distributed creativity and the resulting balances between composition and performance.

<sup>6</sup>The term *dramaturgy* is used as an umbrella term "for a number of aspects that characterise performance practice" (Rebelo, 2009, 389).

these three structures of dramaturgy *projected*, *directed*, and *distributed*, as seen in Fig. 3.8. In projected network performances, the performance design is "suggested by one node" that acts "as author and the others as contributors" (Rebelo, 2009, 389). This node might also "be responsible for the physical/technical organisation of the network" (Rebelo, 2009, 390). In directed network performances, on the other hand, one "artist or group is in charge of the overall performance," but it "becomes the responsibility of each node to support the dramaturgy as developed by the director" (Rebelo, 2009, 390-391). In distributed network performances, "each node retains authorship while contributing specific content and expertise to a shared production," and "media and data are constantly distributed to participating nodes" (Rebelo, 2009, 390-391). Importantly, Rebelo sees these three dramaturgical categories "as a continuum rather than in isolation" (Rebelo, 2009, 390-391) and suggests that distributed dramaturgy in particular "present[s] an opportunity for radical re-thinking of the composer-performer-audience relationship" (Rebelo, 2009, 392). Regarding the research presented in this thesis, the question is what a node might actually represent: Is it the performing group or the individual performer(s), a single audience member or the audience as a whole? This question in turn suggests a nested understanding of dramaturgy because these processes happen on different levels in different micro-groups that come together during a real-time networked performance.

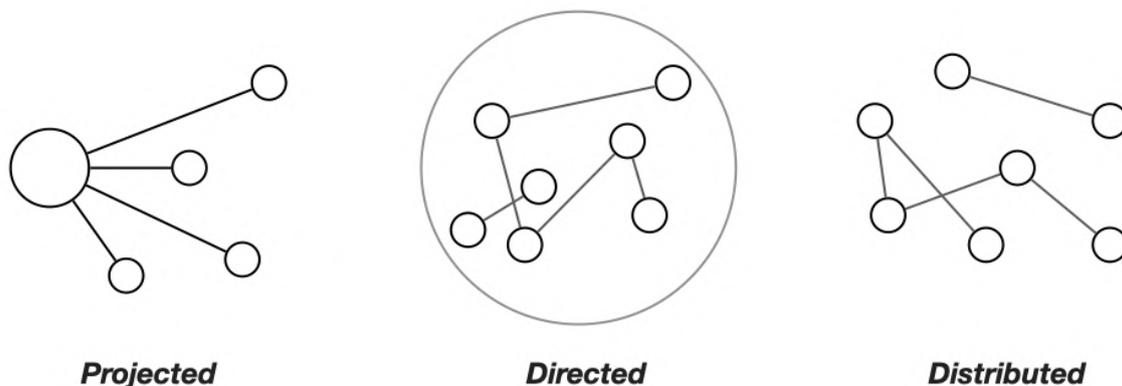


Figure 3.8: Rebelo's typology of network dramaturgy, adapted from (Rebelo, 2009, 390)

### 3.2.2 Audience Instruments

The idea that the audience's affective responses can be seen as collective gestures, coupled with the idea of distributed creativity as a network with agential structures, cultural preliminaries, and various limitations (as discussed in the last section), raises certain new questions. How can audience gestures be included in a creative performance? What would this mean for the design of such systems and the ability to offer an experience to audience members that goes beyond the mimicry of participation to encompass creative emergence in Sawyer's sense of the term? What are the requirements for collective instruments in the first place?

As discussed in the last chapter, audience gestures emerge from many individual contributions arising in a shared performance space and underlying a number of group effects that are characteristic of musical performance, such as musical entrainment, emergence, and group flow. Mazzola suggests the term *hypergestures* for such merged gestures within a network of performing musicians (Mazzola and Cherlin, 2008, 92-94). Consequently, the construction of audience instruments should respect both the individual and networked levels in their design in order to allow group flow to emerge. At the individual level, Csikszentmihalyi (1996) suggests that the use of the instrument should offer neither underwhelm nor overwhelm the performer to allow them to enter a flow state (Csikszentmihalyi, 2014) as cited in (Cook, 2001). At the same time, the instrument should provide a layer of social affordance—that is, an appropriate amount of control of the collective state, as suggested by the above discussion of distributed dramaturgy. For smaller groups of players, Jorda coins the term *multi-user instrument* and recommends a playful approach (Jordà, 2005). A collaborative music instrument would ideally afford collective emergence, whereby the individual interactions would allow for collective gestures that in turn could be accessed by the stage performers. Given that there remains a performative center in any staged performance, we might then speak of the *participatory staging of interactive compositions* or *participatory compositional design*. This terminology points to the fact that the creation of participatory works always encompasses the design of the interface and an idea about how that interface integrates, musically and aesthetically, into the work.

### 3.3 Conclusion

This chapter has considered models of musical communication, distributed music systems and their critiques, and the idea of distributed creativity to make the case for the instrument category of *audience instruments*. The participatory compositional design of such systems often involves a layer of so-called artificially creative parts that act autonomously during performance, which will be discussed in the next chapter.

# Chapter 4

## Autonomy

You don't have a chance...  
...but use it!

---

*Herbert Achternbusch*  
([Achternbusch, 1978](#))

*This chapter follows upon the previous one by engaging further with the issues involved in interactive system design for bigger groups of contributors. It introduces the concepts of artificial creativity and agency, reviews their main principles and terminology, and examines concepts and models for interaction with creative systems. Its final section addresses the use of AC in participatory systems.*

In 2024, big base models and co-performing musical robots are part of our work as performers, and it is hard to keep up to date with the learning speed and practical refinement of these systems. At the same time, much of their acclaimed virtuosity and creative potential is the result of reactive systems and limited style imitations and transfers. Human creativity continues to elude computers, and some argue they cannot get there. The following sections offer an overview of the concepts, developments, limitations, and terminology surrounding AI music systems. In addition, the term *Artificial Intelligence* (AI) is switched over to *Artificial Creativity* (AC) in the field to acknowledge the fact that intelligence is not the same as creativity. Other such terminological shifts include *smart mediators* ([Leman, 2008](#)), *meta-creativity* ([Eigenfeldt et al., 2013](#)), and *computational creativity*. The latter is a science-related complement to the term *generative art*, which describes the creative practice of computationally creative systems in particular.

Much has been written on the relationship between creativity and intelligence. For a long time, creative production was tied to problem solving and seen as a subcategory of human intelligence ([Guilford, 1967](#)). While some scholars have by now suggested that it is the other way around (intelligence is a subset of creativity), while others take the extreme position that they are the same thing (or completely separate), it remains the case that creativity and intelligence are generally regarded as overlapping and intertwined, even symbiotic concepts ([Sternberg and O'Hara,](#)

1998). Presumptions tend to trouble both labels depending upon what one wants to address—a particular creative skill set versus an intelligence capacity, an individual versus a collaborative process, human or non-human agents? Boden pins his hopes on a Turing Test for Artistic Creativity made by machines (Boden, 2010). However, propositions like this confine the notion of creativity to the genesis of creative artifacts and the entities that are able to do so. Creative intelligence, I would argue, goes beyond those artifacts to represent the ontological level of the entity generating them. It is this human or non-human individuality, inclusive of its collaborative abilities, that is what we would describe as either creative or intelligent. For a current and in-depth investigation of artificial intelligence in art, the reader is referred to (Voigts et al., 2024). In what follows, I will turn to autonomous and co-creating machines in the context of musical performance.

## 4.1 What Are Autonomous Performing Machines?

Blackwell et al. define the autonomy of computationally creative systems as follows:

An autonomous system, in contradistinction to an automatic system, is able to act and respond to unknowable and unforeseen inputs, and in ways that have not been completely prescribed. Autonomy is one quality that might enable a machine improviser to become accepted as an equal partner in a group setting (Blackwell et al., 2012, 149).

They list four attributes of creative interaction as requirements of autonomous creative collaborators. The first is the autonomous agent itself, an "embodied system that satisfies its own internal and external goals by its own actions while in continuous interaction with the environment (Beer, 1995)", while that "agent, as used in Artificial Intelligence, refers to a device that perceives its environment through sensors, and takes action on the environment by means of actuators (Russel and Norvig, 2003)" (Blackwell et al., 2012, 150). The other three requirements for autonomous creative collaborators are novelty, participation, and leadership (Blackwell et al., 2012, 151).

### 4.1.1 Theory: Artificial Creativity

Responding to the question "How can creativity be computational?" Bown first acknowledges that the concept of creativity "has historically been something of a moving target and one that is still very much in motion" (Bown, 2021, 63). Theorizing artificial creativity regularly leads back to the field of human creativity, and many scholars approach computational systems through the lens of their human counterparts and expect the machine to live up to those capacities. We will therefore consider selected concepts of human creativity before turning to a discussion of artificial creativity.

Beneath her famous distinction between P(psychological)- and H(historical)-creativity, Boden posits three main sources of creativity that "correspond to the

three ways of surprise" (Boden, 2004, 3). The first arises through the "unfamiliar combinations of familiar ideas" (Boden, 2004, 3) that are understood and valued by the receiver. The second arises through the exploration of conceptual spaces or "structured styles of thought" that produce creative acts: "someone who comes up with a new idea within that thinking style is being creative in the second, exploratory, sense" (Boden, 2004, 4). The third, *transformational creativity*, arises when creative agents reach the limitations of the conceptual space, then go beyond them and extend that space. Her fundamental question regarding artificial creativity bypassed the (at the time) conventional objections regarding creative machines (including their lack of consciousness, intentionality, shared human experience, and so on) to ask "whether it's true that computers could, in fact, come up with ideas that at least *appear* to be creative" (Boden, 2004, 7)<sup>1</sup>. Almost two decades and multiple examples of staged or performed artificial creativity later, most would respond yes to that question. Today, that is, attention has shifted to the interactional spaces, the phenomenological relationships to creative technology, and the social integration of creative systems - in short, to the applicable practice of the creative machines within human performance practice and to interfaces and their autonomy.

As stated earlier, scholars have framed creativity as a profoundly socially grounded phenomenon (Csikszentmihalyi, 1999) (Saunders and Bown, 2015)<sup>2</sup>. Glăveanu's concept of distributed creativity likewise sees creativity as an effect of social interaction - one not confined to the individual cognitive level. He bases it on the 4 P's, as suggested by (Rhodes, 1961)<sup>3</sup>, and his "5 A's framework—actor, action, artifact, audience, affordances—is grounded in current literature from sociocultural and ecological psychology as well as theories of the distributed mind and tries to achieve a more comprehensive and unitary perspective on creativity" (Glăveanu, 2013, 1). Fig. 4.1 and fig. 4.2 show his shift away from Rhodes's model and visualization of the 5 A's framework. In contrast to Rhodes, Glăveanu incorporates audiences and affordances as integral parts of the creative act rather than external evaluators.

In the field of live music, Kantosalo et al. suggest a model of collective, performative creativity they call "the 5 C's" of creativity (Kantosalo and Takala, 2020). This model also derives from Rhodes's four perspectives but adapts it to the field of computational creativity using Jordanous's use of the term *producer* instead of computer (Jordanous, 2016). They also rely upon Glăveanu in their 5 C's model, which is described as follows and depicted in Fig. 4.3.

The creative human–computer *collective* consists of at least one human

---

<sup>1</sup>Wiggins has explored Boden's framework for creativity in her account of implemented creative AI systems and argues that, when we formalize this framework, we must implement rule sets for creative agents, since only that way can we discuss real-life applications (Wiggins, 2006, 214).

<sup>2</sup>For example, Saunders and Gero "implemented a computational model of social creativity using 'curious design agents' able to both generate novel artefacts and evaluate the novelty of artefacts generated by other agents" (Saunders and Gero, 2002). Saunders and Bown theorized this concept into what they call *computational social creativity* (Saunders and Bown, 2015, 10).

<sup>3</sup>Rhodes considers creativity a "multifaceted phenomenon, which includes aspects of the creative *person*, the creative *process*, the creative *product* and the person's relations to their environment, the *press*" (Rhodes, 1961) as read in (Kantosalo and Takala, 2020, 18).

The four P's of creativity		The five A's of creativity	
Focus on: Internal attributes of the person	<b>Person</b> → <b>Actor</b>	Focus on: Personal attributes in relation to a societal context	
Primarily cognitive mechanisms	<b>Process</b> → <b>Action</b>	Coordinated psychological and behavioral manifestation	
Features of products or consensus around them	<b>Product</b> → <b>Artifact</b>	Cultural context or artifact production and evaluation	
The social as an external set of variables conditioning creativity	<b>Press</b> → <b>Audience</b> ↘ <b>Affordances</b>	The interdependence between creators and a social and material world	

Figure 4.1: The development of Glăveanu's 5 A's framework from Rhodes's 4 P's model, adapted from (Glăveanu, 2013, 71)

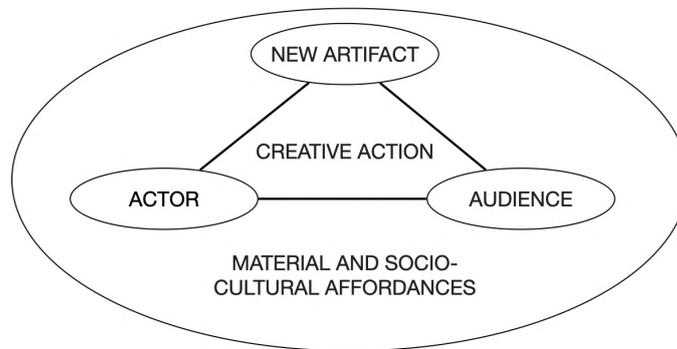


Figure 4.2: Illustration of the 5 A's framework as suggested by Glăveanu, adapted from (Glăveanu, 2013, 72)

and one computational collaborator. The *collaboration* of the collective consists of individual and collaborative creative processes and interactions that support them. The collaboration results in an artefact or a product that represents the *contributions* of the collective. These contributions are communicated to and shared with a wider *community* of peers, audiences, and other social influences. The co-creative collaboration takes place in a *context* representing the environment of the creative act, including e.g. cultural artefacts and conventions, and more immediate factors such as material affordances and shared mental resources, such as the creative task (Kantosalo and Takala, 2020, 19).

They conclude: "collective and context perspectives offer a way to situate co-creativity in a wider sociocultural and physical setting, while offering a way to analyse the effect that individuals and materials outside the collective may have on co-creativity" (Kantosalo and Takala, 2020, 23).

For music performance, those sociocultural and contextual aspects must be related to the domain of time, and Tanimoto suggests six levels of liveness for interactive performance systems. These levels were initially developed for the field of live coding but apply to other live performance systems as well. While levels 1-3 refer

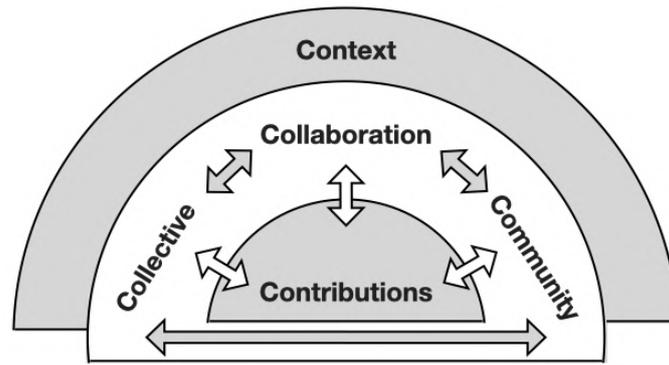


Figure 4.3: Illustration of the five C's as suggested by Kantosalo Takala, adapted from (Kantosalo and Takala, 2020, 22)

to different scenarios for computer calculation times, level 4 describes a real-time system that "would keep running the program, modifying the behavior as specified by the programmer as soon as changes were made" (Tanimoto, 2013, 34). In level 5, the computer not only runs commands in real time but also anticipates the users' actions and runs some of these predictions (or responses to them) in a fashion that Tanimoto describes as "tactically predictive." This "incorporation of the intelligence required to make such predictions into the system is an incorporation of one kind of agency," understood as "the ability to act autonomously" and as "commonly associated with life and liveness" (Tanimoto, 2013, 34). With a "more intelligent inference of the programmer's intentions or desires" rather "than simply making tactical predictions, a system might be capable of successfully making strategic predictions" (Tanimoto, 2013, 34). This is the "strategically predictive" level 6, and it requires "a large knowledge base" (Tanimoto, 2013, 34). Since most of the case studies conducted in article 9.5 are in fact real-time systems, those levels of liveness offer an important perspective upon them.

Artificial creativity is fraught with bottlenecks related to training, audio display, fitness, and agency (who decides what is being played or controlled). In addition, it is hard to evaluate artificial creativity (like creativity in general), because music composition, for example, is an open-ended domain rather than a concrete problem with a corresponding solution. Simon Holland echoes Peña (Pena, 1987) and (Cook, 1994) when he describes music composition as *problem seeking* rather than *problem solving*: "there are in general no clear goals, no criteria for testing correct answers, and no comprehensive set of well-defined methods" (Holland, 2013, 240). Nevertheless, Ariza has attempted to evaluate artificial creativity in relation to human creativity using a test based on the Turing test (Turing, 1950) that is both interaction based and musical (Ariza, 2009, 66).

### 4.1.2 Intention: Creative Agency

Questions of artificial creativity engage agency, intention, and consciousness. The current understanding of a generative AI, gained through prompting ChatGPT with the question "What is agency in music?," produces the following answers:

In music, the term "agency" can have multiple meanings depending on the context. Here are a few different interpretations of agency in music:

1. **Musical Agency within Performance**: This refers to the performer's ability to make choices and decisions while interpreting a musical piece. Musicians often have a degree of agency in how they play a piece, including choices related to dynamics, tempo, articulation, and phrasing. This agency allows for individual expression and interpretation within the framework of the written music.
2. **Composer's Agency**: Composers have agency when creating music. They make decisions about melody, harmony, rhythm, instrumentation, and more. This creative agency allows composers to express their artistic vision and style through their compositions.
3. **Production and Recording Agency**: In the context of music production and recording, agency relates to the choices made by producers, engineers, and artists during the recording and production process. This includes decisions about sound engineering, mixing, mastering, and the overall sonic qualities of a recording.
4. **Music Industry Agency**: This aspect of agency pertains to the decisions and actions taken by individuals, groups, or organizations within the music industry. It could involve choosing which artists to sign, marketing strategies, distribution methods, and other business-related decisions.
5. **Audience Agency**: The listeners of music also have a form of agency. They can choose what music to listen to, how to interpret it personally, and even participate in the music-making process through fan engagement, cover songs, and online communities.
6. **Sociocultural and Political Agency**: Music can also be a tool for social and political agency. Musicians may use their music to convey messages, advocate for causes, or express their views on social and political issues. In this context, music becomes a means of agency for individuals or groups to effect change or raise awareness.

Overall, agency in music refers to the ability of individuals or groups involved in the creation, performance, promotion, and consumption of music to make choices, express themselves, and have an impact on the music and its cultural context. It can encompass artistic expression,

decision-making, and the power to influence the music industry and society at large. <sup>4</sup>.

Interestingly, the ChatGPT output does not mention artificial agency or agency claimed by or distributed to non-human collaborators in music. This reflects the common understanding that agency is something inherently human, or at least can only be attributed to living beings with a conscious mind displaying some form of intentionality. However, this understanding has recently been challenged. Based on Malafouris's suggestion that agency and intentionality are "properties of material engagement, that is, of the grey zone where brain, body and culture conflate" (Malafouris, 2008, 22), Bown suggests superseding "the privilege of the human actor, making a place for the idea of humans and other actors forming temporary networks of interaction that produce things" (Bown, 2015, 21). In line with Latour's Actor-Network Theory (ANT) (Latour, 2005), which suggests incorporating non-human actors (or actants) in our understanding of creativity, Bown introduces *agency networks* to the notion of distributed creativity (Bown and Brown, 2018, 76). This recalls Clark's extended mind theory (Clark, 2003), which implies that "the human organism is linked with an external entity in a two-way interaction, creating a coupled system that can be seen as a cognitive system in its own right" (Clark and Chalmers, 1998, 12). Bown et al. then explore the idea of *behavioral objects* that can be attributed to "two types of agency that are relevant to interaction with behavioural objects: *performative agency*, which refers to the capacity for autonomy in musical software specifically in performance contexts, and *memetic agency*, which refers to the potential for musical software to exert influence in musical change over time" (Bown et al., 2009, 189 (my italics)). In conjunction with the distributed agency of which artificial agency is a part, (overall) performative agency can be seen as a dynamic attribute of the interaction among all the actors in a creative network of human and non-human participants. Malafouris takes this a step further by deciding that agency is not merely the attribute but also a state of flow:

Agency is a temporal and interactively emergent property of activity, not an innate and fixed attribute of the human condition. The ultimate cause of action in this chain of micro and macro events is none of the supposed agents, humans or non-humans; it is the flow of activity itself (Malafouris, 2008, 35).

The question then arises as to how this distribution of agency is actually organized in creative practice. Johnson and Verdicchio describe a *triadic agency* consisting of the designer, the user, and the artifact (or technology) (Johnson and Verdicchio, 2019). Dahlstedt prefers the term *influential agency* to describe the impact human actors have on non-human actors (Dahlstedt, 2021). Influential agency, that is, extends the concept of *causal agency*, which is seen to "pass through" non-human actors such as objects or tools from its origin point in the intentional agency of a

---

<sup>4</sup>Prompted on chat.openai.com on September, 3, 2023.

human actor. Dahlstedt thus counters those AI scholars who suggest the attribution of intentional agency to artificial agents, such as (Boden, 2010). In his concrete explanation of the application of a recursive neural net for music generation, Dahlstedt attempts to show that "there are many layers of human influential agency at play" (Dahlstedt, 2021, 886). This agency is obvious in the multiple roles humans assume: they "invented the general concept of artificial neurons" and "programmed a particular implementation" but also "chose which particular generative algorithm to use," "what hardware to run it on," and "how the material was to be represented to the algorithm." In addition, they usually "choose the training set which has crucial aesthetic implications" and "select the best examples from a large set of generated outputs" (Dahlstedt, 2021, 886). In the end, Dahlstedt also insists upon influential agency because he concludes that mediated agency originating in human intentional agency cannot "carry over" directly and in that way be controlled by the human user with regard to AI tools (Dahlstedt, 2021, 886).

It is precisely at this intersection of humans and machines where certain questions of agency become interesting: What does "carrying over" actually mean and how is it processed? What mechanisms, procedures, and hierarchies are at play here, and what steers this distribution of interaction? Distribution implies a limiting of agency in terms of capacity; it can even deny agency altogether. But is agency as a construct in fact limited from the start? Must agency be balanced among actors as a basis for their creative acts? What happens to agency as flow during these creative scenarios? How can it be measured or described without resorting to a discussion of power structures? Is intentionality a premise for agency, or is it itself an entangled flow state that is distributed across all actors (including artificial ones)? What is the process for extending the intentionality from humans to objects, or does the objects' agency exclude intentionality, as suggested by Latour? Davis has discussed the idea of mediated intentionality using postphenomenology (Davis, 2019) and points to Verbeek's understanding of the role of interaction: "We should not see interaction as interaction between two fixed objects. The humans and technologies are not 'two poles' between which there is an interaction; rather, they are the result of this interaction" (Verbeek, 2015, 28).

### 4.1.3 Concepts: Underlying Principles of AI in Music

One way to categorize musical AI and its autonomous systems is generative versus non-generative - that is, those that generate sound and those that do not, at least directly. Generative systems include those that generate the same outcome from the same prompt (they are deterministic) and those "truly" generative systems that generate different outcomes from the same prompt. This binary separation between music-related (generative) and non-music-related (non-generative) applications should be more refined with regard to AI, depending upon whether the AI system is being used for live or recorded music, or for generating melodies, longer musical structures, or timbres. AI systems can also be structures in and of themselves, depending on the types of inputs of the user and the types of outputs the

system creates (inputs can be audio, symbolic data, text prompts, or sensory data such as movements or other measurements).

When reviewing artificially creative applications for musical metacreation, Bown and Brown suggest "four general human-computer interaction activities that commonly appear in metacreation: (1) generating forms, (2) searching and finding, (3) helping machines learn, and (4) evaluation and iteration" (Bown and Brown, 2018, 70). These activities will be used to structure this section reviewing conceptual approaches to designing creative algorithms. When reviewing algorithms for computational creativity, Brown distinguishes between those generating creative artifacts (generators) and those evaluating them (evaluators) (Bown, 2021, 150). He further argues that while the underlying principles of musical metacreation are often similar (such as learning, optimization, etc.), the algorithms tend to be very domain specific (Bown, 2021, 151). Given the widely ranging concepts that inform various art forms, he expresses "some doubt as to whether a general creativity algorithm is a useful concept" (Bown, 2021, 152). Bown claims that the creative potential of generators is not less than those systems that include both generators and evaluators, even though the latter might display a higher level of autonomy (Bown, 2021, 153).

Taking a step back in the history of the field of artificial intelligence, Marvin Minsky suggested already in 1961 five main areas of heuristic computing: "Search, Pattern-Recognition, Learning, Planning, and Induction" (Minsky, 1961, 8). He noted that intelligence "seems to denote little more than the complex of performances which we happen to respect, but do not understand" and in this way remains "untangible" (Minsky, 1961, 27). As Minsky discusses in his later book *Society of Mind*, it is the co-evolution of subsystems that makes up the intelligence of an individual because those subsystems mainly execute trivial tasks (Minsky, 1988), an idea that has returned recently with the reimportation of AI into the cognitive sciences (Hawkins, 2021). The combination of principles rather than the individual principles themselves likewise denotes the intelligence of an artificially creative system.

## Search and Pattern-Recognition

One set of commonly used search algorithms for music is made up of genetic algorithms, a form of evolutionary algorithm inspired by Darwin's theory of evolution. Loughran recently reviewed work in evolutionary music from 2000 to 2020 and analyzed 47 systems using genetic algorithms or genetic coding (Loughran and O'Neill, 2020). One early system that used audience members as experts to evaluate the fitness of melodies is GenJam by Biles (Biles and Eign, 1995). For a thorough discussion of the fitness bottleneck that bedevils genetic algorithms and potential solutions to them using interactive fitness, see also Biles's contribution to (Miranda and Al Biles, 2007, 42), where he examines the genetic operators *initiation*, *selection*, *crossover*, *mutation*, and *replacement*. More recently, genetic algorithms have been discerned in the context of live coding by (Brown and Sorensen, 2009) and (Hickinbotham and Stepney, 2016).

Another interesting approach for music is the multi-objective search algorithm, which allows the user to distribute and diversify the fitness measures or goals of the algorithm. In addition, pattern recognition can be framed as a hybrid of search and learning—for example, clustering techniques (as unsupervised learning approaches) search for similarities in data while using data as a training source.

## Learning

As an AI field that has attracted much attention, neural networks have also occupied those communities interested in generating sound with AI. One of the first convolutional neural networks is Wavenet (Oord et al., 2016), and the music community has worked with it as well as its successors, including recursive neural networks (RNNs), generative adversarial networks (GANs), long-term short-term memory networks (LSTMs), and variational autoencoders (VAEs), each focusing on a specific aspect of artificial creativity. For general introductions and discussions of deep learning in music, see (Cope, 1991) and (Miranda, 2021). Often, users combine different types of networks, as in RAVE, which uses a combination of VAEs and GANs (Caillon and Esling, 2021), or different types of learning strategies, such as unsupervised, supervised, and reinforcement learning.

## Planning and Induction

In the abovementioned article from 1961, Minsky pointed to planning and induction as principles that acknowledge the environment of creative agents and in that way transcend the mere simulation of individual intelligence. Inductive inference describes a machine's ability "to construct general statements about events beyond its recorded experience" (Minsky, 1961, 27) and in this way go beyond training a system with data. In music, various environment-focused systems have also been explored. Blackwell and Bentley look at swarms that are based on simple patterns but, as larger entities, then generate patterns of emergence—an effect that is well known in music (Blackwell and Bentley, 2002) (Dahlstedt and Nordahl, 2001) (Yee-King, 2011).

More recently, the trend toward implementing agents that are connected to real environments, or understanding creative humans in relation to that environment, has produced interesting projects. For example, Bulley and Jones created a system that connects their installation *Living Symphonies* (Bulley and Jones, 2014) to forest environments. By using a sensor-based infrastructure that collects site-specific data (weather, humidity, pressure, wind direction, wind speed, and so forth), they give agency to the natural environment. They also incorporate observations from people who know the environment in question, such as flora and fauna maps, which they call *warm data* in contrast to the sensor-collected *cold data*. This warm data represents a form of audience participation in the sense of a citizen science, as was explored by Bulley in his PhD thesis *Sounding Materiality* (Bulley, 2018).

Note that it is common in generative AI music to combine principles in applied systems. Often, people use learning to train systems with data that is intended

by the composer or system designer for use in tandem with algorithms that are more suited to real-time application, such as fast search algorithms. For example, Shimon, a marimba-playing robot, uses a combination of a CNN and a genetic algorithm (Savery and Weinberg, 2022). While this brief overview cannot capture all of the new approaches in this active field, it does shed additional light upon the design of intelligent music performance systems such as those discussed below.

## 4.2 How Can We Build Co-Performing Machines?

This section describes what methods are used to generate autonomous musics. Tatar historicizes musical agents as the "modern-day equivalent of the automaton," which he describes as "a self-operating machine that carries out pre-defined procedures" (Tatar and Pasquier, 2019, 56). This definition points to a twofold process in the design of autonomous music technologies. First, system designers predefine certain procedures, then the machine carries them out on its own. Before looking into these processes, we will review terminologies to better understand various notions of autonomy and make visible the intentions of the designers behind those notions.

### 4.2.1 Words: Terminology around Intelligent Music Performance Systems (IMPS)

Many terms describing interactive systems include some form of creative computational autonomy. They reflect a certain way of thinking about how we co-create with machines and some level of speculation about how we might do so in the future. Thus, there is a difference between reviewing the concepts of the systems and reviewing the systems themselves. As we will see, assumptions about human-machine collaboration rarely accommodate the true breadth of performative cultures and contexts. Whether one sees a computer as a partner or a tool makes a profound difference in how one talks about one's system, names it, and ultimately creates it. In addition, concepts must be framed according to the milieu from which the authors emerge. The following is an overview of the most prominent concepts and their corresponding terminology in chronological order.

**Interactive Music Systems (Rowe 1993)** Robert Rowe describes interactive music systems as "those whose behavior changes in response to musical input," noting that such "responsiveness allows these systems to participate in live performances, of both notated and improvised music" (Rowe, 1993, 1). Already in the early 1990s, he discerns that "computer music systems are able to modify their behavior as a function of input from other performing musicians," situating it alongside the abilities to generate new timbres and "implement algorithmic methods for generating musical material" (Rowe, 1993, 2). He claims such "changes of behavior in response to live input are the hallmark of interactive music systems" (Rowe, 1993, 2). He sees those systems as useful for "enriching the performance situations in

which humans work" and advocates for programmers to "concentrate on improving the musicianship of our emerging partners" (Rowe, 1993, 262).

**Improvisational Music Companions (Thom 2000)** A similar idea about human-machine collaboration underpins Thom's use of the term *improvisational music companions* as a subcategory of believable agents (i.e., agents to which human-like behavior has been attributed) in the field of interactive music systems. With her system named Bob (Band-out-of-the-Box), Thom introduces an improvisational companion in the genre of jazz that uses unsupervised machine-learning techniques to supply its "aesthetical musical sense" (Thom, 2000, 309). Its companionship surfaces in the form of four-bar traded solos that the system generates based on offline pre-training (Thom, 2003, 316).

**Interactive Music Improvisers (Blackwell and Bentley 2002)** In the context of their system based on swarm behavior, Blackwell and Bentley introduce the term *interactive music improviser*. Their system, called SWARMUSIC, generates sound "through the attraction of the particle swarm to a target" and claims to be "the first application of swarm intelligence to music" (Blackwell and Bentley, 2002, 1462). It draws upon musical events represented in a three-dimensional space (pitch, loudness, and pulse) to interpret the outcome of the interactive swarm and flock algorithms it uses (Blackwell and Bentley, 2002, 1467).

**Live-coding Systems (Collins et al. 2003)** While live-coding is not exactly a descriptor for interactive collaboration with creative agents, it may still be regarded as a real-time performance approach toward human-machine creativity. As a sub-form of generative coding, "real-time scripting during laptop music performance" moves "beyond graphical programming languages to the command-line antics of interpreted text-based programming languages," (Collins et al., 2003, 321) and it remains a dynamic practice to this day. In the early 2000s, languages for live-coding such as SuperCollider were still in their infancy, but the list of possibilities and dangers presented by Collins et al. holds true for them all today (Collins et al., 2003, 322). Collins even describes live-coding as the "dramatization of a program" (Collins, 2011). Since live-coding performances often share the performer's screen, they at least create a visual connection with their work even if they do not solve "the laptop-performer-stuck-behind-the-laptop dilemma" (Collins et al., 2003, 329). Live-coding runs command lines in real-time and embraces the incomplete over the thoroughly thought-through, and audience members embrace these conditions and offer a culture of understanding and appreciation for those performers who take this on. Connected to live-coding is the *live algorithm*, defined by Blackwell et al. as an "autonomous machine that interacts with musicians in an improvised setting"; research on this algorithm is a multidisciplinary field encompassing "Music Technology, Artificial Intelligence, Cognitive Science, Musicology and Performance Studies" (Blackwell et al., 2012, 147). In addition, the term *algorave* describes events where

live-coding algorithms are used in performance (Collins and McLean, 2014). A current overview of the field's developments can also be found in (Blackwell et al., 2022).

**Interconnected Musical Networks (Weinberg 2005)** Already cited in the section 3.1.2, Gil Weinberg uses the term *network* to capture the structural approaches and architectures of music-creation systems that involve more than one creative party. His graphical topologies mapping the connections and relations among multiple nodes apply to either smaller or bigger groups of musical collaborators and can also be used to theorize non-human players as nodes (even if they are introduced as human entities by Weinberg). While his article mainly focuses on the structure of the networks, Weinberg also includes recommendations regarding the relation between the structure and the content of the network. For example, he notes that "allowing players to influence and control parameters such as pitch or contour of a peer's melody may lead to an incoherent experience for the peer who may lose control over some of the fundamental aspects of the melodic production," and this can lead "into an anarchic experience that is difficult to follow for participants and viewers" (Weinberg, 2005, 37). Weinberg often says that losing control of the musical parameters can cause problems for performers. This, in turn, suggests the need to limit the possibilities for influencing one another's play in the interests of a more coherent musical result.

**Intelligent Music Instruments** Developed from the expression DMI (digital musical instruments), Miranda and Wanderley suggest the term *intelligent musical instruments* for those that boast human-like levels of intelligence (Miranda and Wanderley, 2006). The term has mainly been applied to physical instruments and embedded systems; see, for example, the [Intelligent Instruments Lab](#). This thesis explains and employs the term *Intelligent Music Performance Systems* (IMPS) in [article 4](#).

**Machine Learning as Meta-instrument (Fiebrink et al. 2010)** Fiebrink et al. use the term *meta-instrument* for their real-time interactive machine systems, which use multimodal gestures for mapping onto target sound parameters in a supervised context (Fiebrink et al., 2009). The [Wekinator](#) system has proven a successful means of including learned gesture-sound mappings in musical performances based on classifier and regressor architectures (Fiebrink and Cook, 2010). A broad theoretical framework for applying machine-learning techniques to the making of music via both unsupervised and semi-supervised techniques appears in (Fiebrink and Caramiaux, 2018).

**Virtual Improvisation Partners (Nika and Chemillier 2012)** Nika and Chemillier use *virtual improvisation partner* to describe a "music generation system which is able to understand the logic of the horizontal and vertical associations in a live musical improvisation performance to become itself a source of proposals

by developing its own aesthetics close to that of its partners" (Nika and Chemillier, 2012, 186). With later DYCI2 agents, Nika et al. likewise explore "the creative dynamics of improvised interactions between human and artificial agents, featuring an informed artificial listening scheme, a musical structure discovery and learning scheme, and a generalized interaction / knowledge / decision dynamics scheme" (Nika et al., 2017, 1). Albert also talks about an interactive musical *partner* while exploring human-machine collaboration in duo improvisations (Albert, 2013).

**Collective Creation Systems (Renwick) and Multi-user Instruments (Jordà 2005)** Renwick uses the term *collective creation systems* as "an extension of [the] composition support systems" that are used by communities to collectively compose and anonymously assume "responsibility for small segments that form part of a larger piece of music" (Renwick, 2017, 40-41). Jordà uses the term *distributed instruments* and *multi-user instruments* in this regard to capture the shared collective control that he uses in his models focused on the limitations and dependencies of that control (Jordà, 2005).

**Computational Systems for Music Improvisation (Gifford et al. 2018)** Gifford et al. also refer to computational *systems* for music improvisation when they suggest a taxonomy for improvisational partners that "organizes the field along a number of descriptive axes, relating to the level of creative agency, incorporation of musical analyses and aesthetic tactics, aspects of the interaction design and the underlying algorithms used." They conclude: "system design complexity isn't necessary to achieve some degree of creative agency in a system" (Gifford et al., 2018, 33).

**Music Response Systems (Ravikumar and Wyse 2019)** Ravikumar and Wyse use the term *music response systems* for "systems that interact using musical actions, alternate between lead and follower roles, and make decisions to negotiate transitions in the music" (Ravikumar and Wyse, 2019, 1).

**Computer Systems for Expressive Music Performance (CSEMP)** Kirke and Miranda use the term *computer systems for expressive music performance* (CSEMP) to identify "a computer system able to generate expressive performances of music" (Kirke and Miranda, 2021, 522). Performance creativity, they note, "refers to the ability of the system to generate novel and original performances, as opposed to simulating previous human strategies" (Kirke and Miranda, 2021, 525). Fig. 4.4 depicts how they understand systems that involve this kind of performance creativity.

Reviewing the terminology surrounding intelligent music systems reveals common trends connected to their era and approach. As stated in chapter 3, interactive technology was "one of the hot concepts of the 1990s" (Saltz, 2003, 395), and its close relation to computer technology encouraged designers and composers to conceive of

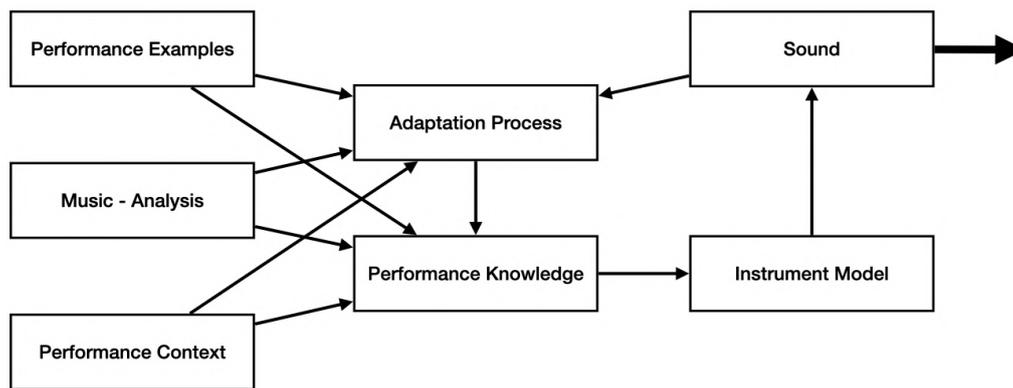


Figure 4.4: Kirke and Miranda's generic model for most current CSEMPs, adapted from (Kirke and Miranda, 2021, 526)

interactive systems; see (Rowe, 1993). In the 2000s more concepts for collaboration with the computer arise, and we come across terms such as interactive improviser (Blackwell and Bentley, 2002, 1467) or companion (Thom, 2000, 309)—an approach that can still be seen in many current systems today (e.g., (Nika et al., 2017, 1)) that focus on the model of one human interacting with one machine, or a back-and-forth creative dialogue. Another line of thought favors collective- or community-focused creativity (sometimes also called distributed creativity), which centers upon the collaboration of more than one human and/or machine, as in a network (see, for example, Weinberg's interactive musical networks (Weinberg, 2005) or Jorda's multi-user instruments (Jordà, 2005).) These systems acknowledge that "macro-structures may emerge as a consequence of the micro-interactions between performers (...) and they can do so without the agents being aware of the developing structure, or even needing to be aware" (Blackwell et al., 2012, 149). Ultimately, we can discern three respective areas of focus in interactive intelligent music systems: on the structure (system, architecture, topology, etc.), on the interactive (the exchange between agents), and on the collective (network, community, etc.). Today, however, we can also discern a return to the systems themselves (as it was in the 1990s), in the wake of wave after wave of new computational methodologies and theories.

## 4.2.2 Practices: Musical Interaction with Generative Systems

Having reviewed the different ways practitioners and researchers approach the design and conception of intelligent performance systems, we will next turn to concepts and models of interaction with technology. The relationships and interactions between people and machines are usually collected under the heading "human-computer interaction" (HCI). In the field of music, HCI is also called simply "music interaction," or "the design, refinement, evaluation, analysis and use of interactive systems that involve computer technology for any kind of musical activity, and in particular, scientific research on any aspect of this topic" (Holland et al., 2013, 5). I will review three approaches to musical interaction below.

Lubart suggests four categories of HCI dedicated to promoting creativity: "computers may facilitate (a) the management of creative work, (b) communication between individuals collaborating on creative projects, (c) the use of creativity enhancement techniques, (d) the creative act through integrated human-computer cooperation during idea production" (Lubart, 2005, 365). He then recognizes four roles that computers can take on as partners with humans in creative engagements. The first is the computer "as nanny," intended to "encourage creativity by monitoring the working process and supporting the potentially creative person as he or she proceeds" (Lubart, 2005, 366). He laments those "periods of work in which doubt, ambiguity and a lack of perseverance can lead people to abandon the creative process" and proposes that rather than "being slowed down by technology," people should take advantage of a "computer system [...] designed in terms of its affordances (in the Gibsonian sense) as a support for creative thought" (Lubart, 2005, 366). Lubart's second role is the computer as "pen-pal" to enhance the communicative, collaborative, and evaluative aspects of the creative process. Along the lines of other mediation and communication tools, he proposes that "the development of electronic brainstorming software that allows individuals to exchange ideas electronically, yielding a common pool of ideas that itself is submitted to group discussion, illustrates well the computer as a networker or facilitator" (Lubart, 2005, 367). The third role is the computer as coach: "the computer as an expert system, knowledgeable in creativity-relevant techniques, can help the user to go as far as possible" by stimulating creativity or suggesting certain cognitive processes "such as divergent thinking, metaphorical thinking, or free association" (Lubart, 2005, 367). The fourth of Lubart's "possible futures for computers in the field of creativity" (Lubart, 2005, 366) is the computer as colleague, the "most ambitious vision of human-computer interaction for creativity" in that it "involves a real partnership, in which humans and computers work hand in hand" (Lubart, 2005, 368). He asks us to "imagine a creative team composed of a human-computer duo in which the user proposes an initial idea, that the computer modifies in a random or heuristic way, which in turn the human modifies in a cycle that continues until the user or outside judges decides that the production is satisfactory" (Lubart, 2005, 368). Ultimately, Lubart argues that these four types of creative partnership will vary greatly depending on the nature of the creative task as well as the user's personality type, creative skillset, and interactional preferences (Lubart, 2005, 369).

Another approach to musical HCI involves Bown and Brown's three paradigms for metacreative design: operation-based, request-based, and ambient-based. In operation-based scenarios, users operate a system by controlling its parameters, training networks, or manipulating search algorithms while the "system may have degrees of autonomy and computational intelligence but is presented as a production tool" (Bown and Brown, 2018, 77). In request-based scenarios, on the other hand, users give prompts to systems that offer responses so that the machine takes on a "clear role that resembles the activities of a person, such as producing something or evaluating something, and would give a clear sense of autonomy" (Bown and Brown, 2018, 77). This is what we see today in the remotely accessible generative models

such as Midjourney or OpenAI's [ChatGPT](#). Lastly, the ambient-based paradigm describes systems operating ambiently and proactively without being directly manipulated or prompted by the user. Such scenarios ask the machine to take on the role of assistant, background analyst, or system that adapts and generates based on "an individual user's needs" ([Bown and Brown, 2018](#), 78).

In terms of the relation and distribution of agency between musical instruments and instrumentalists, Auslander sees two approaches that diverge according to the "basic distinction between those who posit the instrument as an extension of the performer and those who see the two entities as separate" ([Auslander, 2021](#), 65). Transcending the extensionalist view of this relation—in the McLuhanesque sense of a "technological prosthetic that extends the capacities of the human body"—Auslander turns to the concept of *ventriloquism*, which Burrow describes as "not simply [an act] of concealment and transformation but [one that] involves splitting the performer's personality and displacing part of it onto an alter ego that acts as a foil, not a clone" ([Burrows, 1987](#)). Understood in this way, a performer's musical instrument—aside from its autonomy—is "an entity perceived as distinct from, and in tension with, the musician" ([Auslander, 2021](#), 66). Auslander presents two related case studies involving B. B. King's guitar "Lucille" and *GuitarBotana*, a work performed with a music robot designed by Mari Kimura ([Auslander, 2021](#), 68), to argue that the relationship between instrument and performer "is always enacted, though not usually foregrounded, in conventional musical performance" ([Auslander, 2021](#), 69). B. B. King, who "describes Lucille as a distinct individual, with her own sensibility," invites the audience to take part in his relationship. Like B. B. King, Mari Kimura collaborates with her GuitarBot as a fellow musician rather than an instrument and, even though she does "not touch GuitarBot, she nevertheless determines what it plays during the scored sections of the piece through her programming of the computer that controls it" ([Auslander, 2021](#), 77). Auslander's analysis equates the performance with the enactment of the autonomy of a system rather than the autonomy of the system independent of its perceived version in the performance context. He insists, "Kimura's programming of GuitarBot does not grant it true autonomy; its apparent autonomy is an effect created through the way Kimura performs with it, just as Lucille's autonomy is an effect of King's performance" ([Auslander, 2021](#), 77). Ultimately, autonomy in this context is better described as apparent than as actual, in the sense that it derives from "creating the impression that an instrument possesses an identity and agency" ([Auslander, 2021](#), 82).

### 4.2.3 Touch and Immediacy: Embodiment and Human-Machine Integration

Thoughts about the likely future of the relations between instruments, performers, and listeners inevitably lead us to the areas of embodied instruments, musical robotics, and implants. Auslander's distinction between "instruments as extensions" and "instruments as partners" quickly blurs when we begin to consider brain-music interfaces (BMI), brain implants, or even "mere" prosthetic robotic arms as sug-

gested by (Yang et al., 2021). Mueller et al. have called this the paradigm of *integration* and position it as a fourth wave in HCI<sup>5</sup> as depicted in Fig. 4.5 (Mueller et al., 2020). Lopes also uses *integrated* to describe the potential overlap between humans and technological devices that are neither separate entities nor implants in the body even as they connect to our muscles, tissue, and senses (Lopes, 2017), as illustrated in Fig 4.6. Returning to musical applications, it might be argued that the close connection to technological devices that embed levels of autonomy creates a category that falls somewhere between an extension and a co-performing partner.

<b>Era / Paradigm</b>	<b>Users : Machines</b>
Mainframe	many : 1
PC	1 : 1
Mobile	1 : many
Ubiquitous	many : many
<b>Integrated</b>	<b>blurred boundary</b>

Figure 4.5: Paradigms in HCI suggested by Mueller et al., adapted from (Mueller et al., 2020)

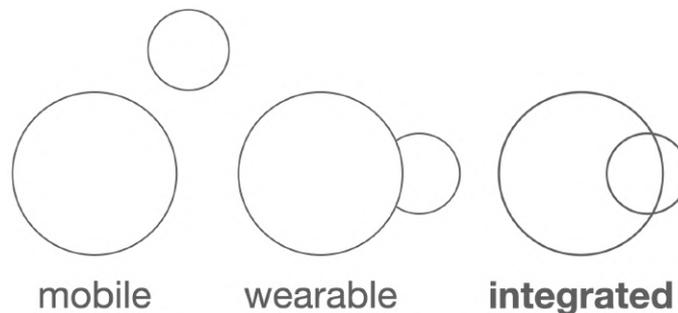


Figure 4.6: Levels of human-machine embodiment as suggested by Lopes, adapted from (Lopes, 2017)

Magnusson considers music-technological artifacts to be *epistemic tools*, knowledge tools that extend human thinking with the help of computing technology (Magnusson, 2009). Because the sense of touch is so critical to musical instruments, they are especially suited to explorations of how embedded AI devices feel - not only what they create and how they respond to our prompts but also how their tactility reacts to our musical gestures and intentions. Integrated musical instruments could even be seen as cyborgian symbioses between humans and machines. Whether these beings are mergings of intentionalities depends on the situation surrounding non-human agency, as discussed earlier. A posthumanist view would suggest that integrated technology is under our final (human) control, while a transhumanist

<sup>5</sup>More about waves in HCI can be found in (Bødker, 2006) and (Bødker, 2015)

view would suggest that integrated technology is on an equal level when interacting with humans. From the perspective of new materialism, we can say that sound changes the bodies with which it is entangled and builds entities and leaves marks, as Fairbairn suggests (Toksöz Fairbairn, 2022). Relatedly, Barad proposes the term intra-action as an alternative to interaction to acknowledge the existence of previously unconnected entities (such as cyborgs, for example) (Barad, 2003, 815).

#### 4.2.4 Models: Conceiving Interaction with Creative Agents

Many HCI models for intelligent musical instruments revolve around the idea that a human delivers some kind of input into a system that generates its output based on a search algorithms, learned behavior, or a combination of both. Examples of such models appear in Figs. 4.7, 4.8, 4.9 and 4.10. In contrast to these, Jon McCormack suggests a model that sees the musical agent as a somewhat separate entity that is able to navigate the environment independently of the human counterpart. This non-human agent is a collaborative partner that interacts with humans at times but does not depend on them, meaning that at times it can also learn, navigate, and develop ideas without any human input and at other times (such as performances or productions) it can interact creatively with its enhanced understanding of its environment. McCormack’s visualization of this agent is given in Fig. 4.11.

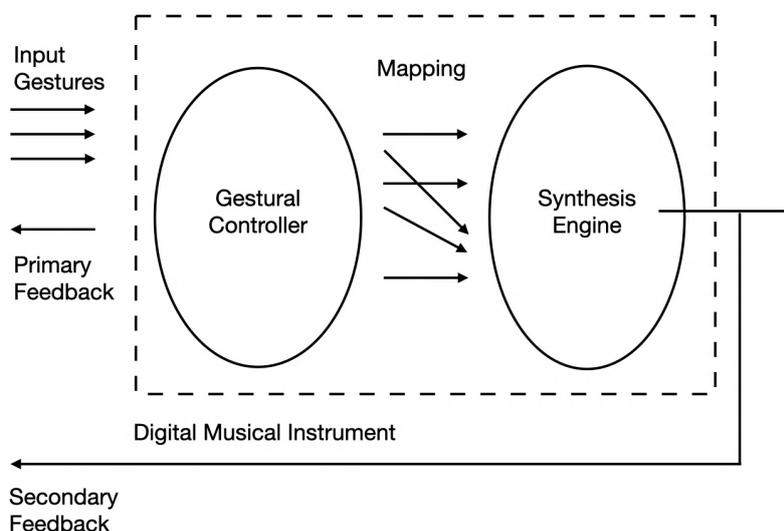


Figure 4.7: Interaction with digital musical instruments, adapted from (Miranda and Wanderley, 2006)

A similarly interconnected smart environment informs Turchet’s term *internet of musical things* (based on the internet of things), which envisions many smart devices organized in a network structure. Turchet’s experiments with musical wearables such as wristbands, backpacks, and other actuators explore the ability to display haptic feedback to performers and audience members to add layers of connectedness to the audio-visual experience (Turchet and Barthet, 2017) (Turchet et al., 2020). The inconsistency of his results suggests a need for a more refined understanding of

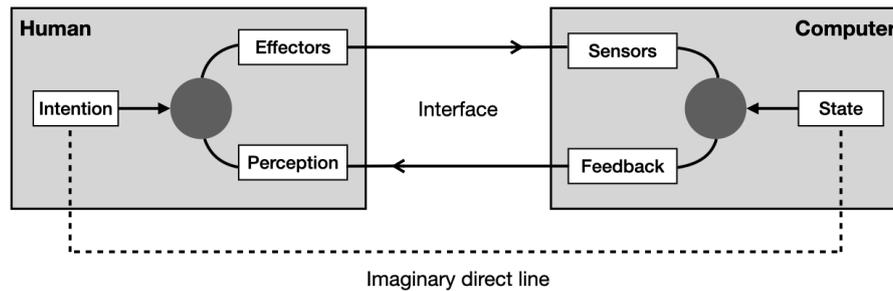


Figure 4.8: Human-machine interaction loop in sonic interaction design, adapted from (Franinović and Serafin, 2013, 86)

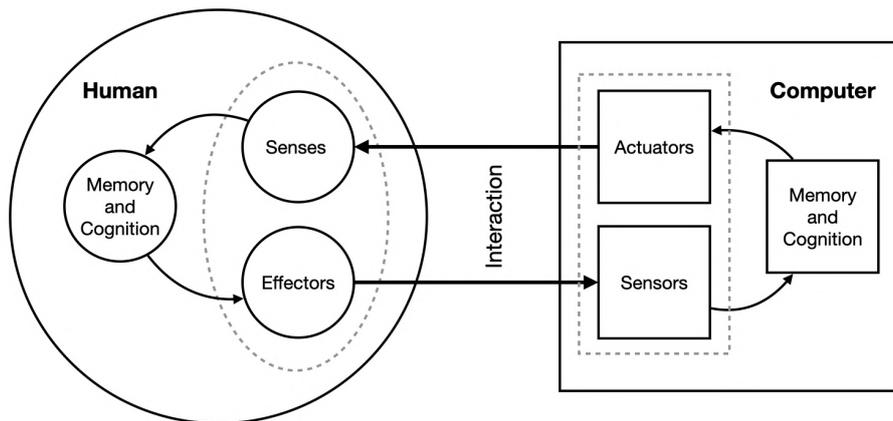


Figure 4.9: Human-machine interaction loop in sonic interaction design, adapted from (Bongers, 2007)

how to deploy such devices and integrate those paradigms into aesthetic experiences when a basic coding-decoding approach is inadequate. Kantosalo et al.’s concept of human-computer co-creativity includes an interaction-based framework for creative tasks, as depicted in Fig. 4.12 (Kantosalo et al., 2020).

To unpack the autonomy of creative agents, Arne Eigenfeldt et al. propose an eight-step model that describes the degree of independence of a system from its user and thus its relative intelligence within its environment. This autonomy, in turn, impacts five different areas of music making:

1. Composition – being the process of creating a series of performance instructions for musical performers (i.e., a score);
2. Interpretation – being the process of performing a musical composition and producing an audio rendering;
3. Improvisation – which combines (1) and (2) in real-time performance;
4. Accompaniment – being the process of following a live performer in an accompanying role, possibly performing pre-composed music; and
5. Continuation – being the process of continuing a given musical input in the same style.

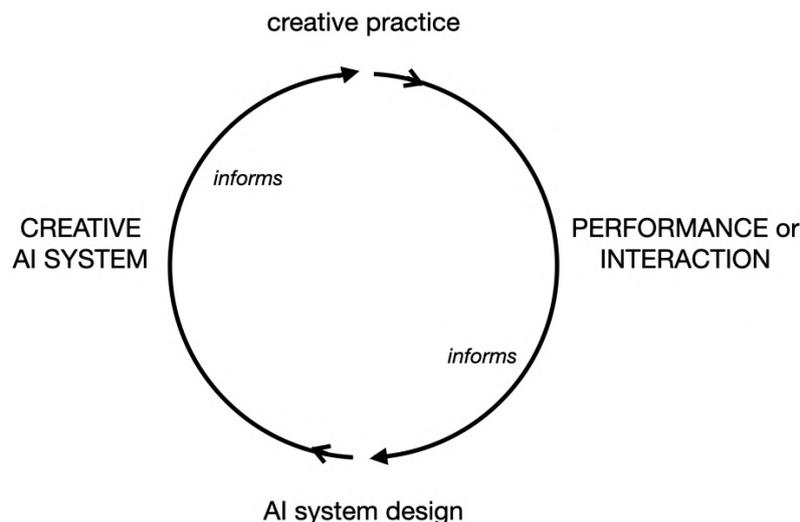


Figure 4.10: Creative practice and AI system design informing each other in human–machine collaborations, adapted from (McCormack et al., 2020, 42)

The eight steps of autonomy range from 0 (compositions that contain no metacreativity) to 7 (complete volition or a "free-standing creative system, independently capable of deciding whether it wants to create" (Eigenfeldt et al., 2013, 6)). The full model with its intermediary steps is discussed in [article 4](#), so it does not appear here.

Models such as the ones described above often have human creativity in mind as a blueprint for computational creativity. However, this approach might not be as helpful as many creativity researchers think. Gibson says that our senses evolved through interaction with the environment. Likewise, autonomous technology should do the same. In line with Auslander, who defines what he calls the *interactionist turn* as "Music is what musicians do" (Auslander, 2021, 83), it could be said that AI music is what AI musicians do. Those AI musicians, as we have seen, take a wide variety of forms that are more or less connected, and more or less dependent on or autonomous from their human collaborators. But they rarely do what human musicians do, and in that sense their creative skills and outputs are quite different. It is the embrace of these new forms of creativity—rather than the simulation of our own human ones—that holds the biggest potential (and most fun) for future music making with machines.

### 4.3 How Can We Use Autonomy for Participatory Performance Systems?

The discussion in the preceding chapters reflects the general assumptions that technology is always embodied in relation to our human (emotional) intelligence, and that the intelligence of systems lies instead in their interaction with us (leaving us as products of the interaction, according to Verbeek). If we accept those assumptions, participatory compositional design becomes the task of creating systems of relations

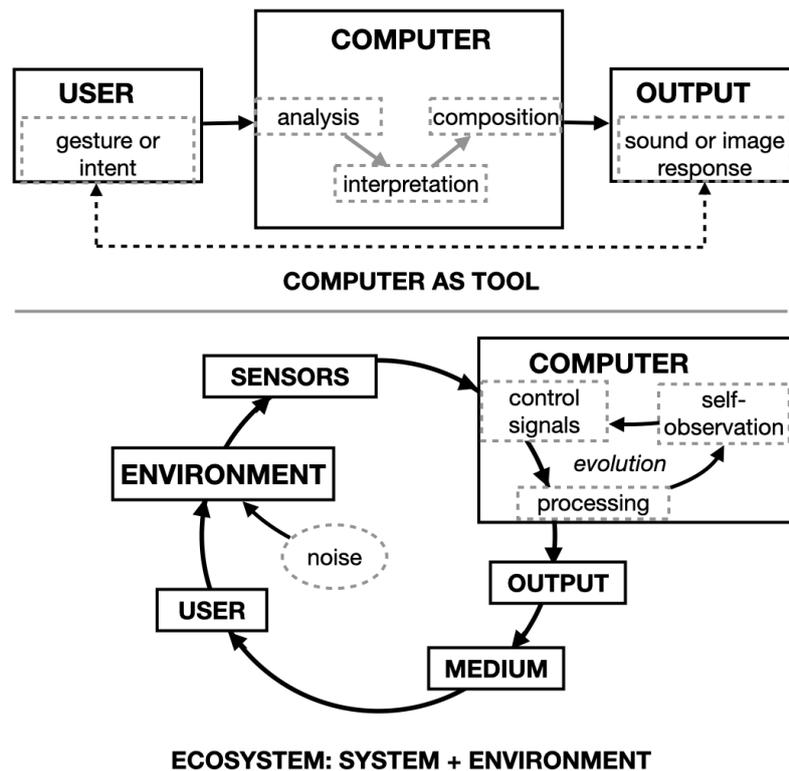


Figure 4.11: McCormack's environmental approach to musical interaction with AI, adapted from (McCormack, 2021)

and possibilities rather than a set of single affordances for play. In turn, the research into those systems should focus on that connection and interplay rather than their underlying processes. In the case of participatory systems for audiences, this brings up the question of how we can afford autonomy to audience members so they may contribute to a musical performance. There are three main possibilities here.

## Individual interface level

First, autonomy can play a role in the interface used by the single audience members. At this level, we must respect the aforementioned general principles of HCI, such as considering immediate levels of multisensory feedback for individual users. But when we are using movement or biosensors as the input, the interactive system must develop an understanding of the creative implications of those data and be trained according to the (individual) creative goals. Audience interfaces should not only be able to sense the emotional stages of the participants but also afford the communication of their creative intentions. In that regard, autonomous systems can help translate emotional states into creative contributions and learn about the personal, embodied input of individual users to the collective state, which can in turn be of creative value to the performance.

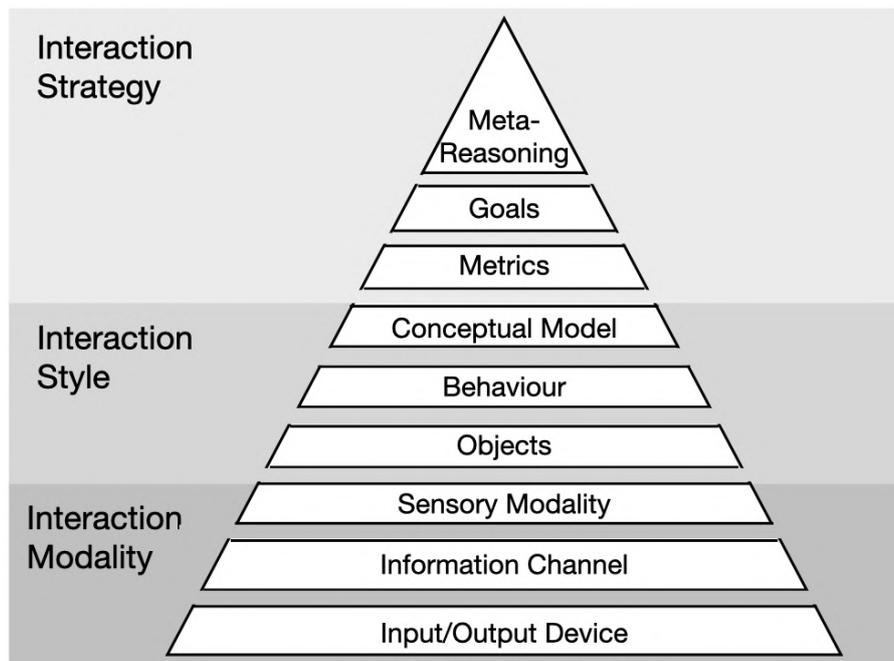


Figure 4.12: Interaction layers of Kantosalo et al.'s framework for human-computer co-creativity, adapted from (Kantosalo et al., 2020, 59)

## Audience gestures

Second, autonomous agents can help recognize audience gestures based on the input data of individual interfaces or sensors and be used to represent different states of collective agency. Those states, in turn, should become more than mere categorical analyses or trivial representations; they should be developed into explorable spaces for both human and non-human agents. Further, models should be trained in relation to their cultural contexts and be able to accommodate future scenarios as well. Ideally, they should be able to recognize and translate audience members' intentions for performers and lead to dialogic scenarios that stay accessible and debatable over time to produce still more refined and valuable creative interactions.

## Feed into performance

Third, autonomous systems can help in aggregating the individual contributions of many participants. Rather than creating autonomous agents in a closed system, single inputs should be understood as nodes of an interface space that has been offered for creative input into performance. Helpful metaphors such as swarms or other biologically inspired approaches might be useful for feeding individual states into a multitude of agents that are then used for creative participation. Feeding those states into agents that are more directly represented as part of the performance is also possible because they are recognizable as such. While the audio domain is quite limited, we could imagine representations of single entities in the set design of a performance, the lighting, the visuals, or other perceivable channels such as haptics, robotics, or virtual avatars. Finally, those systems could be connected to

other personal data sources of the participants, which, with their consent and in anonymized form, could contribute to the live performance based on historical data from previous performances or external data about the concert from social media.

## 4.4 Conclusion

This chapter has discussed concepts of distributed and artificial creativity, including non-human agency and intentionality. It briefly introduced the underlying algorithmic principles currently used to generate autonomous behavior and addressed questions surrounding the design of autonomous technology for live performance. After a review of the terminology used in existing systems, the chapter looked at interaction with non-human agents and the integration of autonomous technologies into participatory performance works. It became clear that audience (or listener) agency is underrepresented in current interactive music systems that involve levels of autonomy<sup>6</sup>.

While some scholars remain very critical of artificial creativity (Cole, 2020), it is generally accepted today that autonomous technology contributes to collective intelligence and distributes creativity across multiple agents. While some approaches still try to replicate and simulate human creativity with closed systems, it is also generally accepted that systems need a connection to human agents in order to stay relevant. These connections can be structured in many ways and offer access to individuals who are able to contribute to collective states in different manners as well. Often, a mix of underlying principles helps designers most in their development of scenarios in which humans become entangled with machines, such as different learning and search algorithms. Training systems with audience data will make the relations and transitions between felt emotions and creative intentions more tangible and possibly allow them to be explored in latent social spaces.

Autonomous agents understood in this way become social agents, and AI becomes an interactive and emotional knowledge shared among the participating individuals. The term techno-cognition—initially suggested by Jensenius for the use of instruments (Jensenius, 2022, XV)—could be extended into a social dimension, bearing in mind the audience’s awareness of its ability to interact with parts of the performance. Techno-cognition also leads back to the twofold experience suggested by Ihde of a first and second phenomenology relating to the direct, physical experience and the subsequent, considered impression of it (Ihde, 2009). In addition, those experiences could be fed back as recognized gestures into a participatory system that was visible to all audience members with the help of autonomous agents.

---

<sup>6</sup>While there are many projects that make use of participatory technology for audiences (see, for example, (Matuszewski et al., 2019) and (Clément et al., 2021)), very few of the musical agents (including levels of autonomy) developed for live performance afford audience participation in their design, as Tatar et al.’s review suggests (Tatar and Pasquier, 2019).

# Chapter 5

## Conceptual Frameworks Revisited

*This chapter will discuss the main theoretical points presented in chapters 2-4, then summarize the concepts developed for the thesis. Finally, it will bridge the theoretical framework developed in part I with the empirical studies in part II by connecting the conceptual and methodological approaches used for this thesis.*

### 5.1 Summary of Theoretical Perspectives

By referring to the title of this thesis, this section will review the main points of the theoretical framework as groundwork for a discussion of the main conceptual approaches in the succeeding section.

#### 5.1.1 Performative Co-Creation

The concept of musical performance, understood as the experience surrounding and emerging from a staged, performative center, is the jumping-off point of any exploration of participatory concert practice. However, this thesis tweaks Frith's well-known (if very broad) understanding of performance in stating that audience members should also be thought of as performers. Musicking recognizes the generative activities of the performance and clarifies their co-creative aspect; it also acknowledges certain assumptions from the field of Embodied Music Cognition such as the central role of bodily gestures in performances involving musical instruments. Those instruments, now often advanced and augmented to become technological mediators, have novel performative applications in popular music concerts. Technology, that is, not only creates new possibilities within established forms of musicking but also pervades the concept of musicking on a formal level. This aspect is highlighted in the thesis title's reference to Ihde's postphenomenological view of technology—in short, that an amplification of one aspect of the experience must mean a reduction in another aspect of it. For example, the ability to design a musical work using an artificial agent based on a large language model may amplify this experience for musical amateurs. But it may reduce the experience as well, in that AI replaces the actual work of generating music directly, robbing amateurs of this work and its rewards.

Ihde's view appears valid to musical composition and performance, but it also appears to imply that human perception is limited in the first place. Stage performers refer to this limitation in terms of bandwidth in the interviews for one of this thesis' studies—there is only so much that a performer can perceive, focus on, and interact with while working in front of an audience. While this is true for a given concrete, momentary scenario, I would argue that a longer view would find human cognition evolving along with and as a part of our musical instruments. As tools that shape our ontology, then, musical instruments and their affordances might not only amplify or reduce experience but also extend it. In essence, the ontological changes that technology is bringing to the space of performed musics are informing the design of new interactive instrumental paradigms and the ways in which we are using them.

At the same time, audiences likewise adjust to the use of novel instruments and approaches to performance. Gibson's ecological approach to perception suggests that listeners determine their individual ways of listening based upon their backgrounds and previous experiences. While this is clearly an active, exploratory process, not everyone will be interested in either sharing it with others or contributing creatively otherwise while attending a given performance. This interest, in turn, seems to depend on both personal background and experience as well as current mood. As section 2.2 demonstrated, models of emotional response help us to understand the effect of music on the emotional constitution of the audience members<sup>1</sup>. Furthermore, the group effects of contagion and emergence are useful theorems for the study and design of participatory music systems, as we will see. An interactive concert, then, could be framed as a co-creative process triggered by the musical performers, accompanied by a subsequent period of attuning (Leman, 2008, 115), and concluding with an improvisation within a staged composition that accommodates a performative agency of many individuals. In this way, composition and performance are intertwined and at times hard to separate.

The idea of performance in popular music remains stubbornly dependent upon the perceived divides between musicians and audiences, performers and listeners, artists and fans. The musical concert is the epicenter where these performer and audience agencies actually meet and, together with the instrumental affordances of technologies, create novel scenarios. Expanding the rigid notion of performance as one-directional rather than interactive is among this thesis's goals, as we will continue to see below.

### 5.1.2 Participation and Distributed Creativity

The term *Participation* is used in conjunction with the terms *Distributed Music* and *Distributed Creativity* because a participatory performance allows for access to and distribution of musical creativity whatever one's group or position in the concert

---

<sup>1</sup>Cognitive approaches and the notion of affect rather than emotion also hold great promise for studies of audience participation systems, given that audiences are actively contributing to the musical performance and affect, as a concept, is more immediate than emotion.

setting as such. Mazzola also speaks of a distributed identity for the product of a participatory performance—"the performed piece of music that is distributed among a number of agents" (Mazzola et al., 2011, 242). This distributed music scenario makes a number of assumptions, however, that merit more nuanced discussion.

First, the concept of distributed music assumes creativity to be distributable or shareable. But what does that mean? If creativity is traditionally viewed as a set of cognitive skills, distributing it could mean sharing the tasks to be solved via the creative agency of a given number of human (and/or nonhuman) agents—that is, creativity as problem solving. On the other hand, creativity as problem creating transcends a specified number of cognitive skills to become a process that is unpredictable, so distributing it would instead mean inviting more agents into this open-ended process. This, in turn, links up with Sawyer's notion of group creativity, an emergent process wherein multiple agents produce a result that is "greater than the sum of its parts."

Second, the traditional navigators of this particular creativity, the composers and performers, would have to be open to the distribution of "their creativity" via an invitation to the audience to co-create. This co-creative process can take many forms, from back-and-forth interactive co-creation to meta-creative scenarios where audiences are given a prepared and limited area for their participation (Rebelo, 2009). Still, it always delivers some level of agency to individuals who would otherwise only attend a performance.

Third, distributed music scenarios presuppose a certain level of engagement on the part of the audience members. Sometimes this lean-in disposition will be readily available; other times, it might not be. For example, in the case of chip-tune music (which was part of this thesis's project), audience members tend to bring with them a high level of familiarity that can translate to a greater interest in involvement. Classical concertgoers, on the other hand, are accustomed to attend highly structured presentations of rigidly notated compositions that accommodate little in the way of audience engagement.

Once a general openness and motivation is shared by both performers and audience members, the qualities of the interactive process itself come into play. King and Gritten's differentiation between communication and interaction (King and Gritten, 2017) offers a useful starting point, and the notion of interactivity as a two-way process was particularly useful to this thesis project. Similarly, Leman's ideas about social musical cognition that includes technological mediators, (Leman, 2008) as depicted in figures 3.1 and 3.2, will serve as the basis for an extended model later in this chapter. While these various frameworks reflect an intuitive understanding of the multi-modal streams of content underlying participatory performances, they offer little insight into the time structures of those interaction loops. Which of these interactive processes rely upon a back-and-forth paradigm wherein musical elements are communicated to the other party, then processed and returned, and which of them work in parallel at the same time? In truth, it depends on the musical process in question and especially which parts of it are shared among the co-creators and

which ones remain hidden to them<sup>2</sup>. In addition, scenarios in which different agents influence different and unconnected parts of the performance may not involve interactive co-creation at all but rather the simple juxtaposition of musical elements that then converge into a co-creative result. Finally, chapter 3 discussed approaches to the structuring of participatory works on a compositional level, and this thesis relied upon Rebelo's *directive approach*, as presented in figure 3.8 (Rebelo, 2009), for the concert studies because those concerts involved an established musical act directing the participatory parts rather than just projecting it, or resolving the performative center all together.

In either case, from the moment audience members make even a small contribution to a concert, questions arise concerning the ownership of the musical content. This ownership encompasses both the perception of the co-creator regarding their part in the co-creation and the legal understanding of what those individual contributions imply for distributive processes, especially with regard to copyright and performance rights. While the second aspect is outside the scope of this thesis, the first one appears in later study questionnaires, when participants are asked about the degree to which they find that their contributions have impacted their sense of ownership and community during the experiments. Returning to the notion of distributed creativity, a form of extended ownership among many collaborators (possibly outliving the musical performance itself) might even require a new notion of "the artist" in this context. Should performers who incorporate participation be referred to as collectives to acknowledge this shared ownership? Is this "distributed creativity" in fact better labeled "additive," "enhanced," or "participatory creativity" that is owned by a collective of human (and possibly also nonhuman) agents rather than a limited group of dedicated composers? This model of ownership could then be tied to Rebelo's continuum (Rebelo, 2009), though the relation is not straightforward, as later studies will show.

### 5.1.3 Accessible Music Creation Technologies

Today, Leman's 2008 vision of "mediators that can deal with effects of social music cognition" while creating "the illusion of non-mediation" (Leman, 2008) appears at once preliminary and optimistic. Even the most accessible and user-friendly music technologies involve tools that interact with human physiology and use interactive paradigms that are impossible to mitigate as mediatory boundaries. (Leman appears to have forgotten about embodiment in proposing this vision). Still, Leman's insistence upon integrating the effects of social music cognition into the design of mediators is a useful prompt for our work on Intelligent Music Performance Systems, and it suggests a few follow-up questions as well: What does it mean to deal with the effects of social music cognition, and how can those effects be integrated into existing mediators and tools? How can a sense of both individual and collective ownership be addressed in their design, and how can individuals be provided with transparency

---

<sup>2</sup>As one example of this, Weeks points out that formative group interactions among performers often happen in rehearsal and are therefore outside the purview of the audience (Weeks, 1996, 200).

regarding their own versus the collective's contributions? Specifically in terms of Leman's requirement to "incorporate mirror processes," how can these reactions be gathered and made accessible in an ethical way during co-creative scenarios? And what does "accessibility" with regard to those mediators actually designate?

The goal seems obvious: intelligent instruments will enable musical amateurs to collaborate with professionals. But what parts of the musical process are the amateurs given access to, and what parts can the artificial agents do for them? The concept of "levels of autonomy" (Eigenfeldt et al., 2013) applies here because it captures the relative independence of a given agent in a given musical context. In addition, Dahlstedt's centering of human agency throughout the multiple stages of composition and performance with AI tools is a useful reminder that AI autonomy is not only inspired and generated by human creativity but also remains dependent upon continuous human interaction (Dahlstedt, 2021).

Chapter 4 advanced a number of concepts of creativity that encompassed non-human agency as a starting point for a discussion of Accessible Music Creation Technologies in general and Intelligent Music Performance Systems in particular. Those included Boden's P-, H-, and transformative notions of creativity (Boden, 2004), Glăveanu's 5 A's (Glăveanu, 2013), Kantosalo and Takala's 5 C's (Kantosalo and Takala, 2020), and the socially grounded theories of creativity proposed by Csikszentmihalyi (Csikszentmihalyi, 1999) and Sawyer (Sawyer, 2006), among others. Participatory compositions arguably include both P- and H-creativity: they invite individual, personal creativities that might not be stageable but that can be included in works of historical value (H-creativity). Furthermore, Boden's transformative creativity relates to participatory work in that they impact a performance at a level beyond the stage content and change the paradigm through which music is both created and perceived.

Considering more recent approaches to musical creativity, Glăveanu's segmentation of *Press* (in Rhodes 4 P's of creativity (Rhodes, 1961)) into Audience and Affordances fails to reframe the audience as a co-creative performer; in relation to this thesis, then, it needs to be reframed. Kantosalo and Takala's 5 C's propose a more applicable theory that focuses on the environmental aspects *Collective*, *Collaboration*, and *Community* because such social factors exist before a *Context* around the epicenter of *Contributions*. The critical role of active contributions in the co-creative process is ultimately underlined by both Csikszentmihalyi and Sawyer, who position the effects of group flow and emergence as focal points of creativity. Returning once more to postphenomenology, agents in the creative process might also be regarded not as the initiators of creative artifacts but rather as products of a co-creative process that revolves around their contributions (Verbeek, 2015). On a more practical level, interactive music systems must afford participants ways to contribute, and the creativity involved in designing them is compositional rather than performative, as mentioned earlier.

The equation "composition plus improvisation equals sounding music" (Mazzola et al., 2011, 234), as suggested by Mazzola, could be reframed in two ways. First, in the context of participatory practice, it could be changed to "composition and

participatory performance (which includes improvisation) equals co-creation." Improvisation is subsumed into the performance, and other composed parts of it are acknowledged as less improvisatory in nature. Second, taking into account the underlying principles of Artificial Creativity as described in section 4.1.3, it could be changed (and inverted) to co-creation = composition + (learning/planning/search) + performance.

Returning to the continuum of compositional and performative creativity, we might decide that the practices of learning and planning lean toward the compositional stage of music creation, as they require careful consideration, the selection of datasets, and the training of neural networks, for example. On the other hand, "search" (such as, for example, genetic algorithms) is a concept that fits best with live performance contexts, given that it takes little time to perform in a real-time scenario. Either way, some understanding of creativity resides at the center of the design of Intelligent Music Performance Systems (IMPS), which is why we will next review the phenomenological aspects that are key variables of that concept.

When we work with musical agents on a musical performance, how we perceive, relate to, and feel about them matters. I already highlighted Lubart's ideas about how to relate to computers in general (Lubart, 2005), together with Auslander's work on performing with musical agents (Auslander, 2021), in the context of designing IMPS. Their insights concern both the bodily aspect of this work—for example, the sense of touch, as discussed in section 4.2.3—as well as the perceived closeness of other agents, as is discussed with the help of musical proxemics in article 3. Perceived closeness during a musical performance has implications for both the physical and the timely domains. Tanimoto's "levels of liveness" (Tanimoto, 2013) offer a way to conceive of felt closeness to real time; high levels of liveness, that is, might characterize real-time systems that can act upon or predict user actions while remaining receptive to surprise contributions. This could in turn further the entanglement of the agents, which might be received as greater perceived closeness. In that sense, autonomy in IMPS might impact not only the five areas of music making (Composition, Interpretation, Improvisation, Accompaniment, and Continuation), as suggested by Eigenfeldt et al. (Eigenfeldt et al., 2013), but also, with the help of recognizing gestures, the meta-level of musical activity. This extended autonomy might create new scenarios that could be conceived with the help of Kantosalo et al.'s idea of meta-reasoning in HCI, which allows for a high-level interaction with IMPS rather than the simulation of human creativity with the idea of a fully autonomous agent in mind (Kantosalo and Takala, 2020). The IMPS model as elaborated in article 4 centers all these aspects to suggest an essential relation between the perceived closeness of musical agents and their autonomy. Rather than the re-creation of human creativity as such, human *co*-creativity revolves around feeling at once autonomous *and* close. How best to include artificial agents in such a dynamic is part of the discussion below.

## 5.2 Summary of Conceptual Approaches

This section integrates the previously discussed theoretical inputs into an inclusive framework of co-creativity. It will present the ways in which these concepts inform the thesis's overall methodological approach, which will be presented in more detail in part II.

### 5.2.1 Audience Gestures, Audience Instruments, and Intelligent Music Performance Systems

Three central concepts informing co-creation were developed above. Section 2.2.5 proposed *audience gestures* as a descriptor for how audiences react, express, and interact during a staged, real-time performance. It then framed them according to the aspects of 4E cognition as embodied, embedded, enacted, and extended in the context of a musical event. This framework was found to be helpful to further the use of the concept of audience gestures, while it might not be exhaustive for all social processes active when audience gestures occur.

An acknowledgment of the communicative dimension of audience gestures (and their musical integration) brings us to the instrumental level of the way in which audiences can co-create with stage performers—that is, audience instruments. Three aspects of audience instruments will impact later empirical work. First, they are in the hands of individual audience members, and as such can be situated in the tradition of single-user instruments, whether acoustic examples or DMIs. Second, they work as multi-user instruments and must accommodate group effects in their design. Third, they must allow for their integration into the context of the staged performance and afford interactive scenarios with the musicians on stage. (Some audience instruments could also work in a non-interactive manner by remaining exclusive to the audience rather than shared with the stage.)

Finally, Intelligent Music Performance Systems (IMPS) are positioned more inclusively as audience instruments that add artificial agents to their design. "Intelligence" here labels the system's ability to understand, interpret, and interact with the musical gestures of the stage performers, the audience members, and the artificial agents. An IMPS can connect those gestural musical intentions, at times serving as a "creative translator" among agents. Still, IMPS do not dominate the creative activity, and group effects can arise outside of the system in the performance space. In addition, we saw musical proxemics, or the study of perceived distance (to other agents and instrumental tools), as central to the concept of IMPS because it connects its individual and collective interactive affordances. Finally, IMPS are systems that connect performer and audience instruments via a more- or less-autonomous paradigm that regulates the proximities between them, and, as a consequence, their perceived distances in the co-creative process. I will return to this in the following section.

## 5.2.2 Co-Creation as a Working Concept

Now that we have reviewed the theoretical inputs and conceptual approaches, we will bring them together in an overarching approach to co-creation as used in this thesis.

Given a performative space within which musical agents and systems can act<sup>3</sup>, "co-creation" is the result of the contributed, embodied enactments of those agents afforded by IMPS as part of that environment. It is also the collective exploration of IMPS as a multi-user instrument and is in that sense process rather than result oriented. The following discussion recognizes both of these aspects of co-creation: while the active process is seen as its centerpiece, the result remains relevant as well. Ultimately, co-creation is assigned a gravitational pull, in the sense that it brings performers and audiences together and, with the help of IMPS, affords them the tools to develop their perceived connections. How this process works and where its limitations lie are central concerns of this thesis.

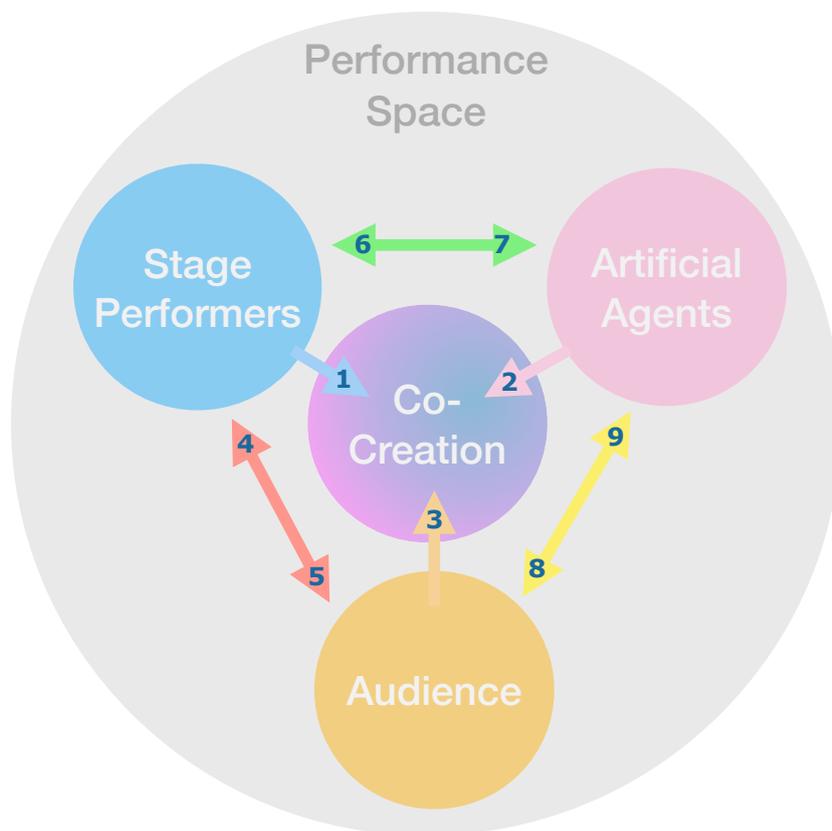


Figure 5.1: Working concept of performative co-creation

Figure 5.1 summarizes the thesis's conceptual approach to performative co-creation graphically. It is based on Leman's model for performer-audience interaction as presented in figure 3.2 (Leman, 2008). Rather than positioning instruments and interfaces as mediators, however, it offers them levels of autonomy that are specific to the performative, real-time context. Furthermore, Kantosalo and Takala's

<sup>3</sup>Note that Iazzetta already suggested in the late 1990s that interactive music systems depend on a shared, dynamic space for their integration into the musical performance (Iazzetta, 1996).

5C's, as referred to in figure 4.3 (Kantosalo and Takala, 2020), are the co-creative contributions at the center of performative co-creation. A detailed consideration of the figure follows below.

- **Arrows 1-3:** These arrows represent the creative contributions of the musical agents: the stage performers, artificial agents, and audience members. Note that the composition can influence all these contributions—it might include directions for the audience members, for example, or accommodate artificial agents via system pretraining.
- **Arrows 4,5:** Arrows 4 and 5 depict the interactive channels between stage performers and audience members, which are comparable to those suggested by Leman (Leman, 2008). These channels can carry multi-modal content but do not include artificial agents in an intermediary stage. For example, such systems can enable audience members to display certain perceptive processes to stage performers without contributing to the auditory domain of the performance directly. Other direct performer-audience interactions are captured by these channels as well, such as applause, for example.
- **Arrows 6,7:** Arrows 6 and 7 depict the interactions between stage performers and artificial agents via, for example, instruments that take their input exclusively from the group of stage musicians (or even a single stage musician) and are not connected to audience members. These interactions, again, can encompass various gestural and modal levels from solely auditory ones to multi-modal gestural ones with the IMPS.
- **Arrows 8,9:** Lastly, arrows 8 and 9 depict the interactions between audience members and artificial agents, which are then communicated directly into the co-creative result or to the stage performers, who use this input for their own contributions. These interactions can also be multi-modal, and some IMPS offer tactile content, as discussed earlier.

The model suggests three aspects as central to the thesis's studies: time, space, and intelligence. The aspect of time returns us to the earlier comparison of performance in real time and composition in what Mazzola calls "logical time" (Mazzola et al., 2011, 246). Co-creation is here understood as a performative process that at once departs from and inspires composition. Mazzola points to Schoenberg's extreme characterization of composition during performance as "slow improvisation," noting that in "every composition there is an improvisational genealogy, and vice versa" (Mazzola et al., 2011, 246). It could likewise be said that in every composition reside its anticipated performative aspects. Participatory performance, in turn, encompasses aspects of composition—for example, planning a parameter change before executing it, or listening for a while before jumping in. In this sense, composition can also be considered participatory in conjunction with the design of the IMPS and its affordance of improvisational freedom. In addition, the audience can

be engaged before the actual performance of a work via workshops or through on-line composition and design. While the continuum of composition and performance does contribute to this thesis's understanding of co-creation, it remains focused on performative co-creation during a real-time event.

As the model suggests, the interactions among the co-creative agents and their contributions to the performative, co-creative result all happen within a performance *space*. This space can be thought of as the physical space of a single concert venue or the much broader space(s) of co-creative processes across multiple venues over longer periods of time and/or the online or hybrid spaces of digital tools. While this thesis focuses on physical concert venues as live performance spaces, co-creative agents experience both actual and perceived distances as they work. As explained earlier, musical proxemics helps us conceive of the perceived distances in particular and integrate them into the design of IMPS.

Let us lastly consider aspects of creative intelligence and artificial creativity in relation to the graph. First, the intelligence in the co-creative scenario is distributed among the different agents: it is embodied in the instrumental play of the performers and connected to audience contributions through the IMPS in the form of distributed, participatory creativity. Distributed creativity, in this sense, is always interactive, meaning that the contributions are to some extent based on interactions between the different agents. Those interactions, in turn, are communicated by musical, multi-modal gestures, or hyper-gestures, as mentioned earlier. The intelligence encoded in the IMPS understands the audience gestures and communicates them to the other agents, serving as a fluid capacity between the system and the agents rather than a limited resource retained by an IMPS alone. Ultimately, this intelligence develops over multiple performances in conjunction with the agents using the IMPS over time.

In conclusion, performative co-creation depends on a number of aspects related to time, space, and different levels of personal, collective, and artificial intelligence. IMPS as technological designs for co-creative performance define and limit the interactive freedom of the actors, and at the same time influence their relations (that is, their perceived distances from one another). Audience instruments, as part of the IMPS, afford not only novel means of interaction between stage performers and audiences but also artificial agents that transcend the simple mediation of musical intentions among human agents. IMPS can afford those perceptions for co-creative use in a performance, or, to cite Heidegger once again, bring forth the pre-existing truth (the felt distances and intentions that are there). IMPS are simplifiers and translators, and they can interact with both amateurs and professionals at the same time. They become part of the musical composition and performative gestures, but they also change the performative co-creation on an ontological level.

### 5.3 Methodological Approaches

This section finalizes the discussion of the theoretical inputs and suggested concepts and connects them to the empirical work via the quantitative and qualitative methodological tools used for the experimental and analytical studies.

Figure 5.2 shows an overview of these connections. The theoretical inputs and related concepts appear at the top of the figure, and the methods and studies that correspond to the four published articles appear at the bottom. The methods will be discussed below along with the concepts audience gestures, audience instruments, and IMPS below. At the beginning of the empirical part II, the qualitative and quantitative methods are presented individually in more detail.

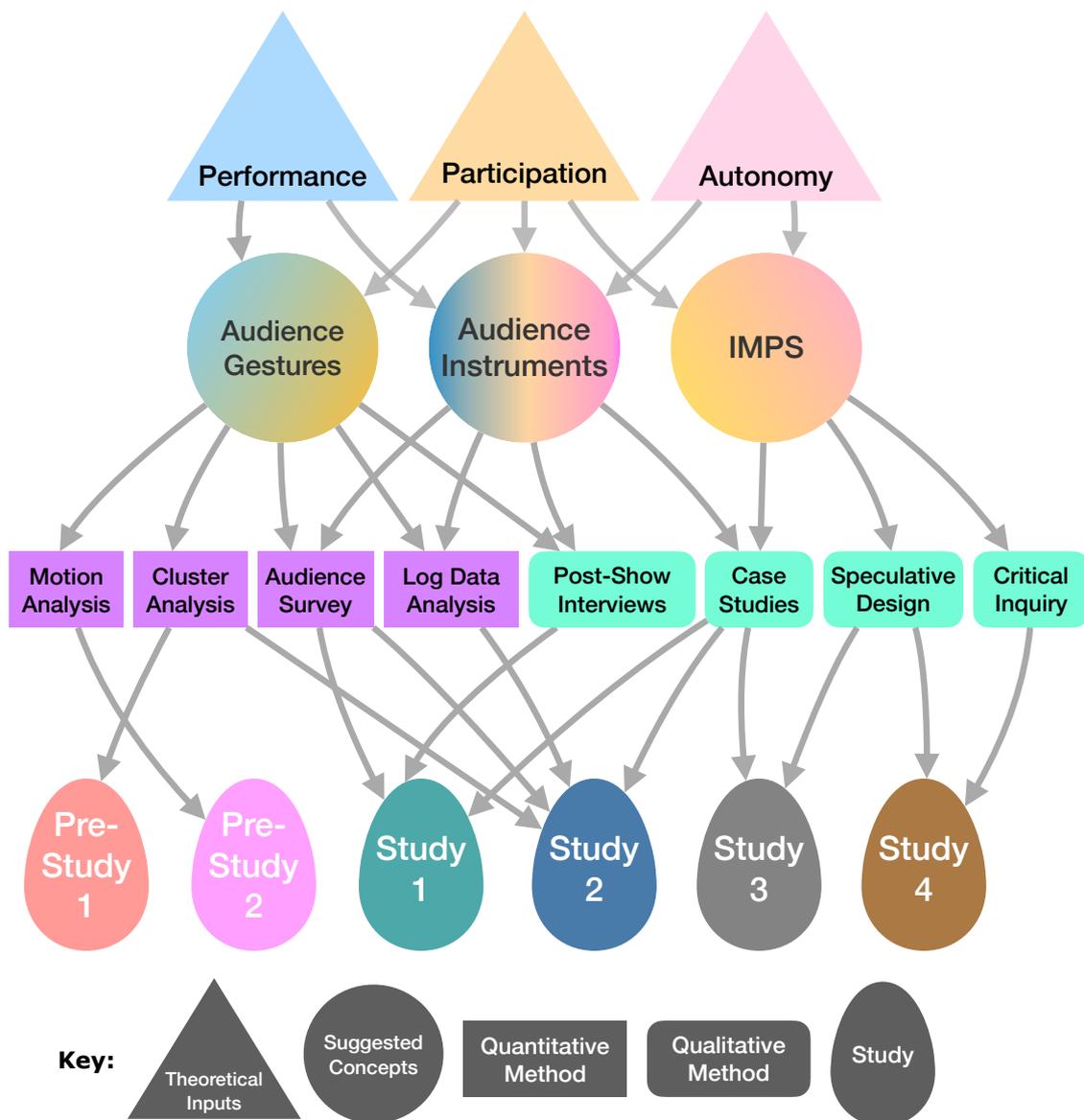


Figure 5.2: Overview of theories, concepts, methods, and studies of the thesis project

The concept of audience gestures discussed above was initially prompted by a review of audience-reaction audio during preliminary study 1. Cluster analysis

uncovered collective patterns in the data that could then be framed as collective gestural feedback. In terms of the theoretical inputs, effects of group cognition such as contagion and attuning were among the outcomes of that study. Against the background of the emotional and social states that exist outside of the concert as well, audience gestures were analyzed at both individual and collective levels to better understand the interaction between stage performers and their audiences. Both emotional response and social music cognition play a part in those reactions, as discussed in section 2.2, as will be clear in the empirical part of the thesis via questionnaires investigating backgrounds, preconditions, and demographics (the individual level) as well as clustering techniques (the collective level).

Next, those audience gestures needed to be made accessible to the performance using the concepts of distributed creativity and composition. This gave rise to an exploratory study that involved building a participatory system that was musically integrated—that is, the compositional design of the live performance was created together with a group of performers. To unpack the experience of the first concert study, we logged the interaction data, created a self-report survey, and held interviews with both audience members and performers after the concert. As a conceptual backdrop, we established "audience instrument" as a descriptor for the instrumental level of the audience's contributions (alongside the direct audience gestures expressed in a traditional manner without participatory technology). In addition, motivated by the concepts of musical embodiment, we conducted a motion analysis of a performer during one of the interactive modes and compared this motion data to the participatory (collective) log data from the audience instrument in order to study the audience-performer interaction loop. This analysis resonated with the concepts of ecological listening as well as emergence effects, both of which were clear results of preliminary study 2. In a second concert study, we refined the analysis of the individual level of the audience instrument (in contrast to the analysis of the quantities of the collective contributions) to analyze the timely dimension of single contributors over time. Again, cluster analysis helped to structure the data, which supported the discussion of various participatory audience types.

Several case studies of performance systems ultimately extended the concept of audience instruments into the realm of the Intelligent Music Performance System (IMPS). Theoretical inputs from the study of artificial creativity helped to describe the design of those systems, together with aspects of embodiment and participatory art. The positioning of the single cases of IMPS was done hermeneutically by the author, and the implications derived from that study remain limited to this perspective, particularly with regard to the tendencies found and the speculations about future designs. The respective inquiries included design recommendations, and methods from Speculative Design Research offered insight and direction for further development of these topics of study. Lastly, the concept of autonomous agents gave way to entanglement theories that produced a fourth critical study of genetic algorithms as a way to implement artificial agents. This thesis concludes by suggesting concepts for implementation and is intended to be a first step toward future work with and studies of IMPS.

Part II  
Empirical Studies



# Chapter 6

## Research Design and Applied Methods

### 6.1 Research Design

The PhD project used an abductive mixed-methods approach built upon multiple **exploratory** studies that were reshaped and tailored to sequentially incorporate relevant theories. This is especially true for the work on participatory performance. On the other hand, the third article on Intelligent Music Performance Systems, took a **confirmatory** approach by testing a model based on existing theories with the help of some real-world cases. On a more general level, the project belongs to a post-positivistic and phenomenological tradition in its reliance upon measures taken from experiments in tandem with experiences taken from interviews and from designing systems as data. Its methodological approach is also connected to postphenomenology (as presented in the introduction) in terms of its understanding of technology as a way to create, communicate, and analyze musical processes at live concerts.

The research design changed along the way and some of the methods presented early on were later abandoned. In this sense, the methodological approach was itself exploratory as well as abductive in nature. While answering the research questions and testing theoretical concepts against the empirical evidence, some of the methods were found to be ill suited or too demanding of either time or resources. For example, it would have been very useful in the second concert experiment to simultaneously record motion data and do interviews, but eventually we decided to limit ourselves to a focus upon the user behavior with the participation interface. After the second concert study, as well, we decided to fall back upon theories of participatory compositions to better understand autonomous concepts within interactive musicking. The next step here would be to return to the empirical studies and try to validate the applicability of the model suggested in article 3. It would also be interesting to apply the recommendations for IMPS design to the work done on the participatory concerts and conduct a further empirical study of a confirmatory nature.

## 6.2 Quantitative Methods

The project's first method uses both audio and log data for various quantitative analyses. Method 2 informs the analysis of motion data stemming from accelerometers placed on the performer's limbs. Method 3 uses interaction logs and suggests a novel measurement unit (Quantity of Participation or QoP), and method 4 presents the analysis of the self-report questionnaire data gathered after the concert experiments.

### 6.2.1 Machine-Learning Methods

Both supervised and unsupervised machine-learning techniques were used in the research. The first preliminary study drew upon clustering as an unsupervised technique for uncovering similarities among the audience-reaction audio files. Specifically, k-means clustering analyzed the data with the aim of finding groups of individual samples and optimized the distances between them. The silhouette score then shows the summed errors for different numbers of clusters ([Menardi, 2011](#)), which makes it possible to find the ideal number of clusters for a dataset. Clustering was also used in the later concert studies to uncover similarities in the interaction log data from the audience members in the interests of discerning common user groups or personas. In addition, after the annotation of audio samples in the first preliminary study, regression was used as a supervised technique to create a model that can predict the scores within four categories attributed to the audience gestures.

### 6.2.2 Motion-Capture Methods

Motion analysis was used in the second preliminary study to better understand the effect that audience participation has on the behavior of the stage performers and especially their bodily motion. I learned about motion analysis in an online course and a physical course at the University of Oslo, including infrared motion capture, the use of accelerometers, and qualitative methods of analysing motion. For the concert studies, which were conducted in an ecological environment (in this case, a small concert venue), the use of accelerometers proved to be the most technically practical alternative. Infrared systems do supply more refined results, but installing such a system in the concert venue was too challenging given the existing technical infrastructure. We used video recordings in tandem with the AX3 accelerometers, but the latter gave us the most usable data for the research design and we ended up placing them on both wrists and the right ankle of the drummer for the experiment. Since we did not return to this method in later studies, we do not discuss it in depth here, but we refer the reader to ([Nymoene, 2013](#)) for a review of the methods of motion capture and motion data analysis for musical applications.

### 6.2.3 Log Data Analysis

To accommodate the analysis of the use of the participation interface, the interaction data was logged in anonymous form. As a measurement unit, we suggested the

Quantity of Participation (QoP) so as to observe how many contributions were made at a given time. They were mostly analyzed as cumulative plots that continuously totaled the total click counts. The log data was also analyzed with the before-mentioned cluster technique, this time looking at variables connected to the dwell time, total click numbers, and button repeats performed by the audience members on their personal smartphones. It is important to note that this method was somewhat tangled up with the participatory concept itself. On the one hand, the participants' contributions were intended to be creative collaborations to the sound and lights of the concert; on the other hand we used this data as a research method. The participants awareness that their creative acts were also an analytical tool might have resulted in some bias on their part, but this was not mentioned during the interviews and did not show up for us during the process.

### 6.2.4 Questionnaire Analysis

The questionnaire used for the studies was developed specifically for the concert experiments and included a first part with questions about the participants' backgrounds (demographics, personality traits, musical experience, genre familiarity, etc.) and a second part with questions about the participatory concert experience. A detailed description of the categories of the questionnaire can be found in [article 1](#) and [article 2](#). The full questionnaire is included in [the appendix](#) at the end of the thesis. Most of the questions were 7-point Likert-type questions analyzed as continuous data. Some of the questions were multiple choice, such as the ones about preferences among the three different interaction modes. The data was reported with arithmetical means and standard deviations using percentages for multiple-choice questions. There was an attempt to find correlations between the first part (personal background) and the second part (the experience) via cross-correlation matrices, but we eventually decided that reporting the general demographics and the results of the experience was enough for the purpose of the research projects. However, productive work remains on additional analysis of those relations, as well as correlations between the backgrounds/experiences and the user behavior as an outcome of the interaction log analysis.

## 6.3 Qualitative Methods

Qualitative methods were used for exploring the field, gaining insight into the experiences from participants, and studying IMPS cases and the culture surrounding them. They are explained briefly below.

### 6.3.1 Practice-based Research Design

An early method in the thesis process involved creating a system for analysis and using it to design participatory sections of a concert. This is an action-based method on its own, even though it was not logged or studied auto-ethnographically in itself.

Instead, those experiences are reported in the articles and inform the model-building done in the third article. In addition, the process of creating the interactions and selecting the musical parts was described by Vegard Kummen, who acted as both researcher and performer for the concerts. His was an interesting perspective upon designing the concert experience, and some of the challenges of the compository design process also came up during the interviews after the first concert experiment. On a more general level, the method also raised the issue of explorability: What can actually be explored with a practice-based approach? How do the decisions made in the approach in turn determine what it is possible to explore? For a more general discussion of art-based methods, see (Orning, 2017) and (Leavy, 2017).

### 6.3.2 Interviews

Interviews were given after the first concert experiment to both performer and audience groups to acquire a refined picture of their experiences. The interviews were semi-structured and used the same categories as the survey. However, participants were also invited to introduce their own topics, and the questions were often posed in open form. The related interview guides can be found in [the appendix](#). The audience sample for the interview was voluntary, and audience members were asked to stay after the concert to take part; about a third of them did. The interview was conducted by Vegard Kummen and Matthias Jung, who were developing the musical and interactive parts of the concert, respectively. This might have created interviewee bias in the sense that they were clearly involved with the project as designers, composers, and performers. Nevertheless, the interview supplied critical first impressions and descriptions of the experience of the interface and the listening experience in general. It offered interesting inputs for the further development of the software and helped us refine the research for subsequent projects. The performer interview also brought up additional aspects of the experience, and since the performers did not fill out the surveys, it was the only data source regarding perspectives from the stage. The details and results of that interview are discussed in [article 1](#).

### 6.3.3 Case Studies

A number of IMPS cases were used in article 3 to validate the theoretical model proposed in that article. The case studies included an analysis of the cultural context of the systems, their musical applications, and their accessibility, and it focused on the main categories suggested by the model - the actual and perceived physical distances connected to the systems (the proxemic plane), as well as the level of autonomous technology used for the systems. The case-study approach then led into a discussion of design processes and desired developments in IMPS and connected to the speculative design research. The case studies were therefore critical to understanding trends and trajectories of future development as well as past, present, and future processes within this development.

### 6.3.4 Speculative Design

Speculative design as described by (Dunne and Raby, 2013) and discussed by (Grunwald, 2015) and (Elsden et al., 2017) is an attempt to theorize the speculation concerning future development in design processes. It implies a space of (technologically) possible future developments that are then separated into plausible, probable, and preferable developments. Given that the development of IMPS is at a highly dynamic (and thus speculative) stage at this point in time, this method was a beneficial means of discussing future developments and theorizing them from multiple perspectives in a refined manner. It offered a more holistic view that incorporated not only historic and current cultural developments but also technological limitations and the culture in the design space itself.



# Chapter 7

## Preliminary Studies

The following studies were both conducted as parts of courses taken during the PhD period. The first one looks into audience gestures in the audio domain and tries to cluster those and find main categories. The second one attempts to open up the performers' instrumental sounds through performer-audience improvisation and takes an analytical approach to understanding how audience contributions affect the performer's movements on stage. This study led to the article 1, which is summarized in chapter 8.

### 7.1 Preliminary Study 1: Clustering Audience Reaction Audio

*This study was done as part of the course "Music and Machine Learning" at the University of Oslo in autumn 2021 and is part of the project report. Since it was offered as a Master seminar, it is not part of the coursework of the PhD.*

#### 7.1.1 Application and Relevance

Audience reactions such as clapping and cheering are critical forms of feedback for performers. In live recordings, they clearly contribute to performer-audience interactions and represent highlights in the environmental or "atmospheric" dimensions of a show. Detection of the audio signals of these reactions has been studied to enable auto-annotation for live videos or concert analysis (([Gemmeke et al., 2017](#)), ([Hershey et al., 2021](#)), ([Gillick et al., 2021](#))). Previous work includes discriminative approaches to applause signals as part of media content and is usually limited to an applause/non-applause classification ([Jarina and Olajec, 2007](#)). On a more general level, applause can be seen as an emotional response and as such can be classified in emotional response models as a reaction of both high valence and high arousal as discussed in ([Schuller, 2013](#), 308ff). With this study, I sought to look at a variety of audience reaction sounds and explore their properties with clustering techniques. This could be helpful to an understanding of the dynamics of audience interaction from the perspective of performers and their content, but the refined detection of

audience-reaction audio would also be useful for estimating audience size, venue characteristics, agitation level, and the collective emotional status of the audience as well as other specifics concerned with reacting to staged content, such as cheering, laughing, whistling, or bawling. I limited the dataset to audio only for reasons of access and complexity, though video is a promising extension of the audio-only approach. In addition, with a bigger visually supported database, one could better correlate musical signals and audience reactions in order to extend existing work on music-induced collective behavior such as (O’Neill and Sloboda, 2017) (Matuszewski et al., 2019) (Hödl et al., 2020).

### 7.1.2 Machine-Learning Algorithm and Feature Extraction

As a basis for feature extraction, two-second clips of the audio samples are loaded following a 100 ms offset (to exclude fade-ins in the audio files). The limitation of two seconds is chosen because some of the samples are not much longer than this and it is adequate for feature extraction because the applause audio does not change significantly in the longer term (multiple seconds). The audio files are then normalized (transformed into unitary scale) and slightly expanded dynamically, since this leads to better clustering results later on (the dynamic variance in the audio becomes more separable that way).

#### Extracted Features

In a first approach, three low-level features are extracted from the audio files: the mean of the zero-crossing-rate, the variance of the rms, and the mean of the spectral flatness. The first two are supposed to point to differences in frequency and intensity of the applause and other reaction sounds, while the third is supposed to distinguish between different forms of reactions such as cheering or other vocally induced sounds. This results in the extraction of three features, which will be referred to as features-low.

In a second approach inspired by (Manoj et al., 2011) and (Uhle, 2011), MFCCs (2nd order derivatives) with a spectral resolution of 20 bands are extracted from the same audio slices. The result are 1740 features, which will be referred to as features-mfcc.

#### Scaling

Three different scaling methods from the Scikit-learn package were used: StandardScaler, MinMaxScaler, and RobustScaler.

#### Machine-Learning Algorithm

In order to cluster the 80 audio samples, we compared the following four approaches:

1. Low-level features -> Scaling -> Kmeans clustering  
The first approach uses the scaled three low-level features directly for Kmeans clustering without dimensionality reduction.
2. Low-level features -> Scaling -> Isomap -> Kmeans clustering  
The second approach is identical to the first but uses Isomap as a nonlinear dimensionality reduction technique before Kmeans clustering.
3. MFCC features -> Scaling -> PCA -> MeanShift clustering  
The third approach uses the MFCCs and PCA to reduce dimensionality to two.
4. MFCC features -> Scaling -> Isomap -> DBSCAN clustering  
The fourth approach uses Isomap to reduce dimensions to three and DBSCAN for clustering.

After performing these clustering approaches, the resulting clusters are analyzed by listening to the most representative examples that are closest to the cluster centers. This is used as the basis for a second, supervised approach, for which all 80 samples were manually annotated using the four categories *crowd size*, *agitation*, *whistling*, and *vocal component* - assigning to each category decimal values within the unit scale [0.0, 1.0] as shown in 7.1, which are then imported as targets. The MFCCs are used as features since they have better temporal and spectral resolution. Both a Linear Regressor and an MLP Regressor are tested after Isomapping and standard-scaling the features.

filename	crowd_size	agitation	whistling	vocal_component
ES_Crowd Applause 7 - SFX Producer.mp3	0.2	0.1	0.0	0.0
ES_Crowd Teens Cheer 5 - SFX Producer.mp3	0.7	0.6	0.3	0.2
ES_Crowd Teens Cheer 8 - SFX Producer.mp3	0.6	0.5	0.2	0.2
ES_Crowd Applause 16 - SFX Producer.mp3	0.2	0.2	0.0	0.0
ES_Crowd Cheer 3 - SFX Producer.mp3	0.8	0.5	0.0	0.8
ES_Human Crowd Studio 55 - SFX Producer.mp3	0.6	0.9	0.0	0.8

Figure 7.1: Manual annotation of the dataset using four categories

### 7.1.3 Dataset

The dataset consists of 80 audio files of audience reactions (applause, cheering, whistling, bawling) downloaded from the royalty-free audio platform [Epidemic Sound](#). The files differ broadly in terms of audience size and demographics, recording method, loudness, and certain venue characteristics. Initially, there were 120 files, but some files turned out to be very bad quality and were manually excluded .

### 7.1.4 Implementation

The implementation uses the following packages: Jupyter Notebooks and various Python packages, in particular Librosa, Scikit-learn, and some other functional packages. For feature extraction, the Librosa functions for normalization and expansion are used before extraction of the zero crossing rate (mean), the rms (variance), and spectral flatness (mean). Librosa is also used for extracting the MFCCs. The Scikit-learn StandardScaler, MinMaxScaler, and RobustScaler are used as scalers. Scikit-learn is also used for calculating KMeans clustering, Isomap dimensionality reduction, PCA, MeanShift, and DBSCAN in the Clustering notebook, as well as for Linear and MLP Regression, Pipeline, RepeatedKFold, and GridSearchCV in the Regression notebook. Matplotlib is used for plotting, and IPython.display is used for listening to the clustered audio examples.

### 7.1.5 Evaluation

As described above, the results are twofold (clustering and regression).

#### Clustering

The Silhouette Score is taken as the main metric to compare the four clustering approaches, since it appears to be the best indicator of clustering as it encompasses both intra-cluster distance and the mean nearest-cluster distance (for a detailed discussion, see (Menardi, 2011)). This leads to the following results:

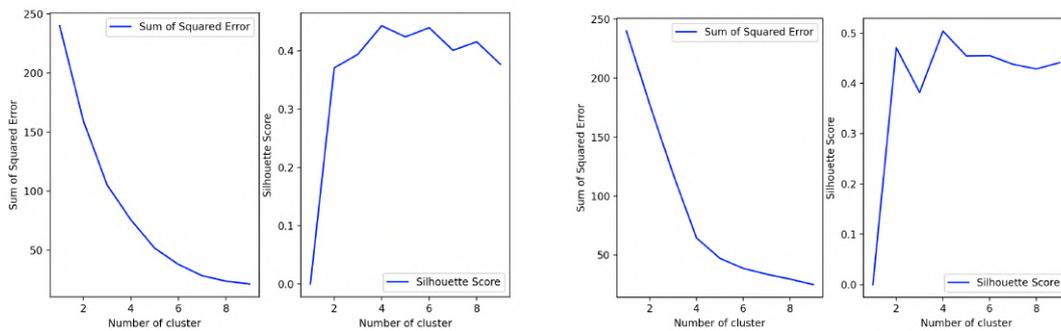
- Approach 1: Silhouette Score = 0.44 (4 clusters using StandardScaler); 0.81 (2 clusters using RobustScaler)

It seems that a small amount of low-level features is enough for a basic level of clustering. Using RobustScaler instead of StandardScaler raises the silhouette score to 0.81 at 2 clusters. This is not surprising, since RobustScaling lowers the significance of outliers and the dataset is known to have only a small number of examples of very sparse applause samples. Despite this higher Silhouette Score, two clusters do not seem to cover the variety present in the data and are considered too limited for the the expected clustering task.

- Approach 2: Silhouette Score = 0.50 (4 clusters); 0.63 (2 clusters)  
Performing Isomap on the data while keeping the dimensions at three leads to a small improvement (0.50). Scaling the features before Isomap even raises the score to 0.55. Reducing the dimensions during Isomap to two raises the Silhouette Score to 0.63, but this is with two clusters, which does not seem to be enough (see figure 7.2).
- Approach 3: Silhouette Score = 0.33 (3 clusters found)  
Using a higher spectral and temporal resolution with MFCCs does not lead to better results. This approach involves PCAing the array down to two dimensions and finds three clusters. Listening to the audio samples in the cluster

centers suggests that this clustering fails to generate meaningful findings. PCA as a linear technique therefore seems to be too limited here in comparison to Isomap.

- Approach 4: Silhouette Score = 0.10 (2 clusters found)  
A reduction to three dimensions before DBSCAN clustering leads to a Silhouette Score of 0.10 and two clusters. Both MeanShift and DBSCAN start with random points, so different random states are tested but do not produce significantly better results. Also, changing the order of scaling and reducing dimensions does not improve these results compared to approaches 1 and 2.



(a) Elbow and silhouette score for approach 1

(b) Elbow and silhouette score for approach 2

Figure 7.2: Plots for use of 3 low level features

Figure 7.3 shows scatter plots of the four clustering approaches. All of them point to two relatively big clusters and two (or one) smaller cluster(s). Listening to the audio files of the different clusters reveals that the two predominant groups contain most of the examples with whistling and/or sounds that involve vocal contributions such as shouting or bawling. The best possible Silhouette Score that generated four clusters was reached with approach 2 (0.50). Listening to the most representative examples in comparison to the other approaches also confirms this result. The first found cluster seems to assemble examples in which applause (clapping) is present. Clusters 2 and 3 constitute the biggest clusters composed of examples with more vocally induced sounds. Cluster 4 contains a small number of examples of small groups with not much reaction at all - one might imagine a small, less enthusiastic group. Following these clustering results, all the examples were annotated and tested with two regression techniques, as discussed in the next section.

## Regression

Linear Regression, as described above, works surprisingly well, with an R2 score of 0.54. The individual R2 scores for the four categories are 0.45 for crowd size, 0.58 for agitation, 0.52 for whistling, and 0.60 for vocal component. Performing 10-times repeated 4-fold cross-validation produces an MSE mean error of -0.042 (variance

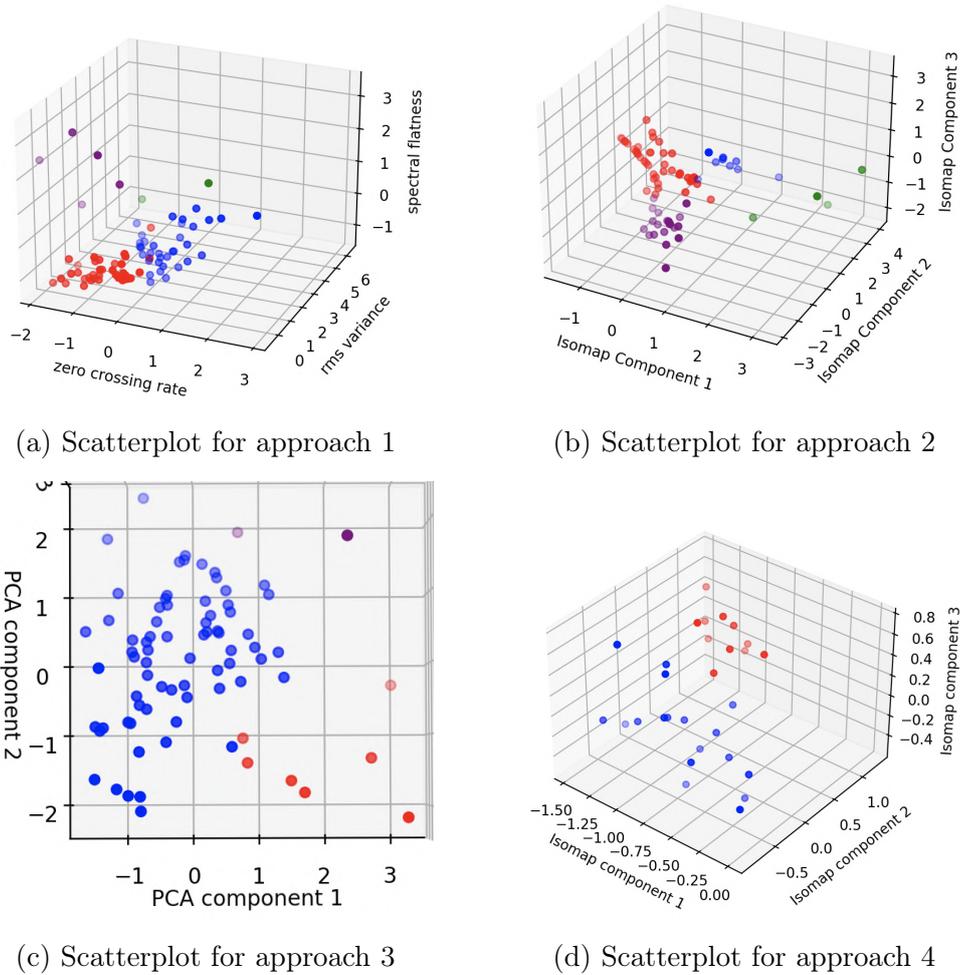


Figure 7.3: Scatterplots for different clustering approaches

0.00004) and an R2 mean score of 0.36 (variance 0.014).

Using an MLP Regressor with the same features and targets generates an R2 score of -1.35. The individual R2 scores for the four categories are -0.46 for crowd size, -4.56 for agitation, -0.29 for whistling, and -0.68 for vocal component. The loss function is shown in figure 7.4. The MSE is 0.13 and is considered more meaningful than the median absolute error, since the MSE is more sensitive to outliers, which seems important in the present dataset. Performing 10-times repeated 4-fold cross-validation produces an MSE mean error of -0.08 (variance 0.0006) and an R2 mean score of -0.38 (variance 0.41).

In order to improve the limited results of the MLP Regressor, we perform a grid search with higher number of hidden layers, different activation functions, and learning rates. The best R2 score of 0.32 is achieved with an Isomap dimensionality reduction to 10, hidden layer sizes of (80, 50, 20), a logistic activation function (sigmoid), and a constant learning rate.

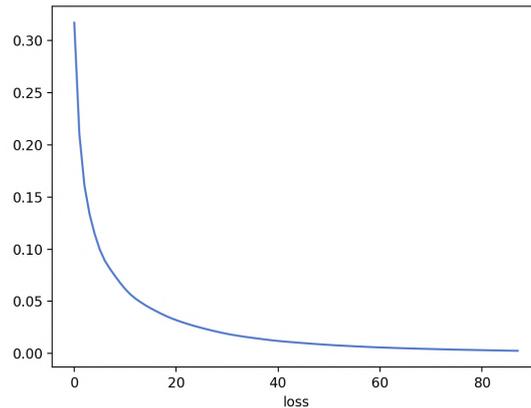


Figure 7.4: Loss function of MLP Regressor

### 7.1.6 Discussion, Conclusion, and Further Work

A database of 80 audio samples containing audience reaction sounds was analyzed with clustering techniques, and an initial attempt was made to build a model for clustering based on regression techniques. The results of the clustering pointed in some interesting directions but remained rather limited given the variety present in the audio data. It was especially challenging to abstract recording techniques of the initial audio from parameters such as the size of the audience or the agitation level present with such a small dataset. However, our exploration of the dataset via various clustering techniques generated more ideas about what is relevant in these signals and how we might learn from annotated data in meaningful ways. Future work should include testing more clustering techniques, such as using Kmedians instead of Kmeans, as proposed by (Seif, 2018), Self-Organizing Maps, or the Gaussian Mixture Model (GMM). The use of a Linear Regressor and an MLP Regressor to validate the found clustering results was preliminary and should be explored in more depth via other methods for the regression task with more annotated data. Since this experiment was limited to the audio domain but audience gestures are multimodal, the study should be extended to those other dimensions as well, incorporating, for example, reaction videos or movement studies. Further, the results of the clusters could be integrated into established emotional response frameworks as described in chapter 3. For example, it might be helpful to trace gestural aspects of audience reactions such as applause and vocal components onto valence and arousal measures. After all, it seems clear that the recognition of audience reactions is a helpful tool for both performers and media in their professional practices.

## 7.2 Preliminary Study 2: How Does Audience Participation Affect Performer Motion?

*This study was done as part of the course "Motion Capture" at the University of Oslo in spring 2022, and its project report was part of the coursework of the PhD program.*

### 7.2.1 Motivation, Background, and Previous Work

This experiment investigates the question of how participating audiences affect the movement of performing musicians at live concerts. Audience participation has been explored in recent years in tandem with the availability of accessible music-creation technologies, often using the personal smartphones of audience members as input devices. Previous work ranges from systems that let the audience directly control or remix publicly displayed sound, such as (Lee and Freeman, 2013), to voting systems with social-media-like features for bigger audiences, such as (Lee et al., 2016). For an in-depth discussion of participatory concert typologies, including their technical implications, see, for example, (Weinberg, 2005), (Renwick, 2017) and (Matuszewski et al., 2019). For a more qualitative discussion of these systems from both a performer and an audience point of view, including communicative theory, please refer to (O'Neill and Sloboda, 2017) and (Hödl et al., 2020).

Whereas much previous work on audience participation covers the aspects of interactive paradigms, system design, and user experience, the aspect of performer and audience movement is rarely considered. Since this experiment includes the performer's body motion, other relevant background includes work on musical gesture (Godøy and Leman, 2010), the relationships between music and body motion (Godøy et al., 2016), and the methods and technologies for analysing these (Nymoen, 2013).

### 7.2.2 Experimental Setup

This experiment took place as part of a concert study with the chipmusic duo Kubbi on April 10, 2022, at Vaktbua in Kristiansand, a small but established concert venue for popular and especially electronic music. This ecological environment served as the research setting for a pilot study on audience participation and included three participatory experiments. The first invited the audience to change the backgrounds of the music; the second let them change the melodic center in the foreground; and the third gave them the ability to change the parameters of the digital instruments of the performers. It is the third experiment that is used for analysis here.

The concert was attended by 25 people who were using their personal smartphones to access a [Max patch](#) via a Node.JS server connected to the browser-based [Miraweb package](#). Figure 7.5 shows the visual interface that I created for this ex-

periment and that was accessible to the audience members during the concert. The audience members could change two parameters of the following four instruments:

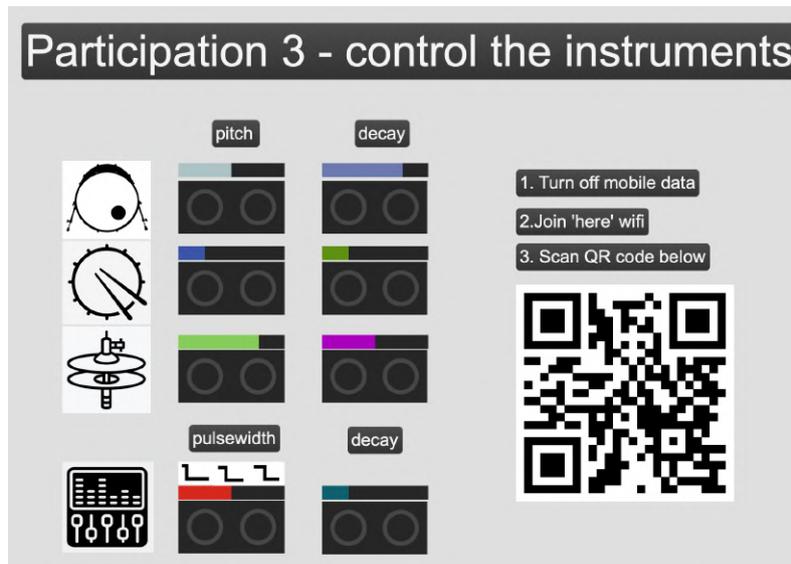


Figure 7.5: Interaction interface for audience participation (experiment 3)

- The pitch and decay of the bass drum.
- The pitch and decay of the snare drum.
- The pitch and decay of the hi-hat.
- The pulse width and decay of the Gameboy chip emulation.

All of these parameters could be changed via counter buttons (up and down) and were fixed to four states from 0 to 100 percent for keeping the interaction simple. Instruments 1-3 were played live by the drummer on an electronic drum set that generated the audio on a computer that was separate from the one used for the web applications. Instrument 4 was an arpeggio inside of a Gameboy chip emulation that reflected the parameters, that are commonly controlled in a chipmusic aesthetic. All instrument parameters were changed via MIDI in real time and played back through the public loudspeakers for the audience as well as the in-ear systems of the performers, so that they were perceivable to everyone present at the concert.

Concerning the movement data, 3 [AX3 devices](#) were used on both wrists and the right ankle of the drummer. Figure 7.6 shows a video still of the performing drummer during the experiment. Figure 7.7 shows a picture taken just before the concert on site.

### 7.2.3 Data Collection

For the experiment, the following data was collected:



Figure 7.6: Performing drummer during the participatory concert (experiment 3)



Figure 7.7: Audience and stage at the participatory concert

- Video recording of the drummer with a GoPro camera (frontal view).
- Audio in the form of stems, mixed down roughly by the audio engineer after the concert.
- Interaction data from the audience members, collected as a log in the Max console.
- Sensor data from the accelerometers (AX3 devices) on wrists and right ankle of the drummer.

All data types were synced using five fast and large movements of the drummer's arms, and at the end of the fifth one, the drummer hit a MIDI pad that triggered a clock object within the Max patch. All interaction data were then logged with timestamps after this initial point in time, and video, audio, and acceleration data were cut down to that point accordingly. The total length of the experiment had a duration of 105760ms (i.e., close to 2 minutes) from the first to the last pressed button by audience members.

### Video and Audio Data

Video of the drummer was recorded with a GoPro camera mounted on a truss roughly two meters away from him. It mainly shows the movement of his arms

and hands on the electronic drum kit that is integrated into his acoustic drum kit. During this section of the concert, he only plays the electronic drums, which consist of a MIDI pad for the hi-hat (right hand), an electronic snare drum (left hand), and a bass drum pedal (right foot). Audio was gathered as a multitrack recording through the mixing desk at the venue. The audio engineer delivered stems for most instruments and provided a rough mix that is used here for analysis and synced to the video recording. The two-minute video of the third participation experiment can be accessed [here](#).

For better audibility of the changing parameters during the audience participation, a slightly slower version of the third participation part was created and can be accessed [here](#).

### Audience Interaction Data

Audience interaction was gathered in the form of a log within the Max environment. It provides a timestamp for each parameter change, the button that was used for changing the parameter, and a counter that accumulates the clicks on each of the buttons. See table 7.8 for a table view of the data.

	<b>button</b>	<b>click_counter</b>	<b>single_count</b>
<b>timestamp</b>			
<b>0.0</b>	hh_decay_up	1.0	1
<b>514.0</b>	hh_decay_up	2.0	1
<b>2540.0</b>	sn_decay_up	1.0	1
<b>2711.0</b>	sn_decay_up	2.0	1
<b>3784.0</b>	sn_decay_down	1.0	1
...	...	...	...
<b>105313.0</b>	synth_decay_down	52.0	1
<b>105567.0</b>	synth_decay_down	53.0	1
<b>105585.0</b>	bd_pitch_down	46.0	1
<b>105757.0</b>	synth_decay_down	54.0	1
<b>105770.0</b>	sn_pitch_down	77.0	1

844 rows × 3 columns

Figure 7.8: Audience interaction data

## Accelerometer Data

Accelerometer data was collected from the AX3 devices and imported into dataframes. Figure 7.9 shows the accelerometer data from the right wrist of the drummer at a sample frequency of 100Hz.

	<b>X</b>	<b>Y</b>	<b>Z</b>
<b>time</b>			
<b>0</b>	-0.921875	-0.031250	0.203125
<b>10</b>	-0.921875	-0.031250	0.203125
<b>20</b>	-0.906250	-0.031250	0.187500
<b>30</b>	-0.890625	-0.031250	0.171875
<b>40</b>	-0.890625	-0.015625	0.156250
...	...	...	...
<b>105720</b>	-0.765625	0.062500	0.000000
<b>105730</b>	-0.046875	0.671875	-0.093750
<b>105740</b>	0.812500	1.078125	-0.203125
<b>105750</b>	1.015625	1.015625	0.015625
<b>105760</b>	0.515625	0.609375	0.218750

10577 rows × 3 columns

Figure 7.9: Accelerometer data of the bracelet on the drummer’s right wrist

### 7.2.4 Qualitative and Quantitative Data Analysis

In order to explore relationships between the participatory actions of the audience and the body motion of the performing drummer, we chose a twofold approach: first, a qualitative look at the video and audio recording; second, a numerical approach to the interaction data and its relation to the sensor data from the accelerometers.

#### Qualitative Analysis

Listening to the audio recording reveals that it takes about 20 to 30 seconds for the audience to get going with the interface, after which we hear changes in the instrument. Next comes a rather wild exploration of the parameters, which change almost randomly for 30 seconds before giving way to more informed choices. During this transition, the drummer not only continues his planned performance but also explores the changes himself, particularly on the snare drum, which he plays with his left hand. Looking back at the video, we can discern two strategies here. First, the drummer plays the snare drum less frequently, in case a lot of changes should

start to happen - when the audience turns up the decay of the snare, for example, it becomes harder for him to play more notes because they all blend into one another. Second, probably in the interests of exploring the sudden changes in sound, he uses both hands to quickly repeat drum hits at the same time, in order to better grasp the changes that are happening. Finally, during the last 30 seconds of the experiment, a general crescendo and overall increase in intensity of the music seems to affect both the interaction as well as the body motion of the drummer, and this continues until the breaking point after the section.

### Quantitative analysis

To simplify the process of data exploration, I will limit the analysis of the accelerometer data to the z-axis of the drummer's wrists. As the visual exploration of the video data showed, the movements of the left hand are more telling than those of the right hand because the right hand is used more continuously and the changes in hi-hat sound that it triggers are not as apparent as the changes in the snare drum sound played by the left hand. Figure 7.10 shows a cumulative plot of all interactions that were performed by the audience members over the course of the experiment period of about two minutes.

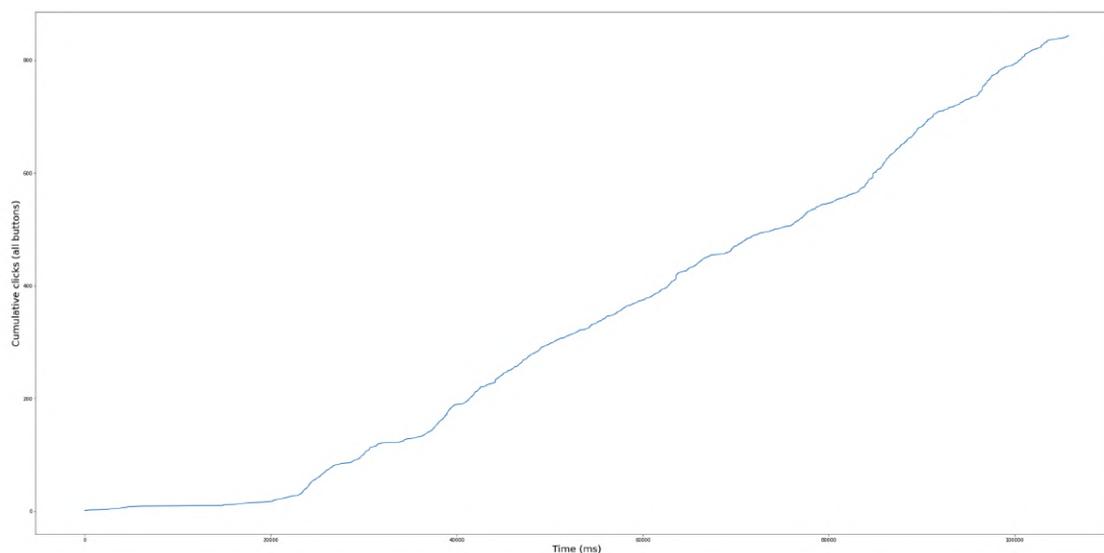


Figure 7.10: Cumulative plot of all interactions by audience members

**Plotting cumulative interaction data against acceleration of right and left wrists** Figure 7.11 shows the accelerometer data (z-axis only) taken from the right wrist, which was scaled to the cumulative interaction data for better comparison, whereas figure 7.12 shows a cumulative plot of both interaction and accelerometer data in order to explore the tendencies in the data.

Similar plots were then created for the left wrist used to play the snare drum, again only considering the z-axis and scaling it to the range of the interaction data

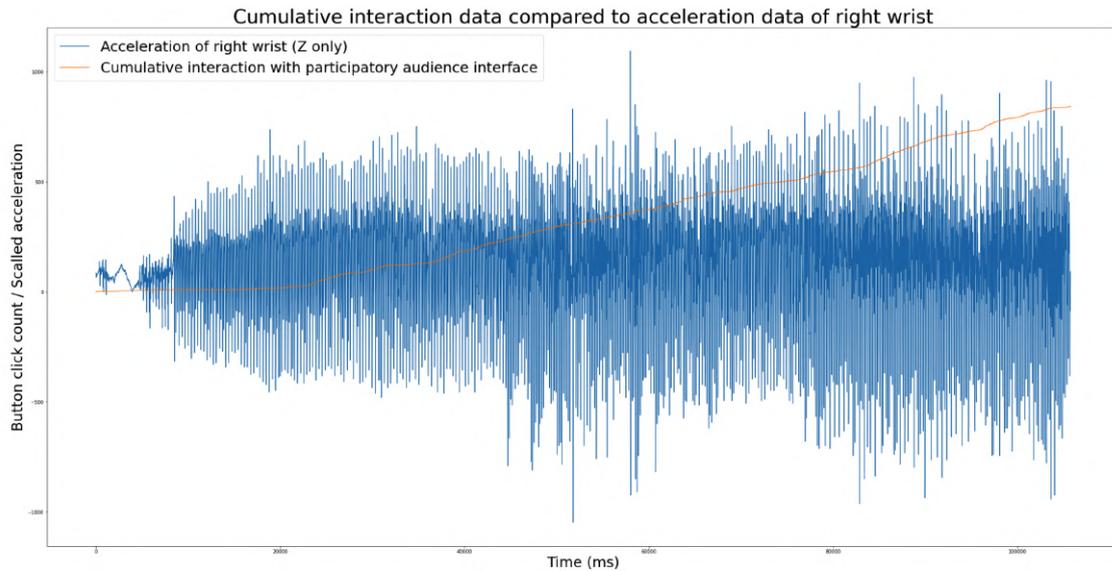


Figure 7.11: Cumulative interaction data compared to acceleration data of right wrist

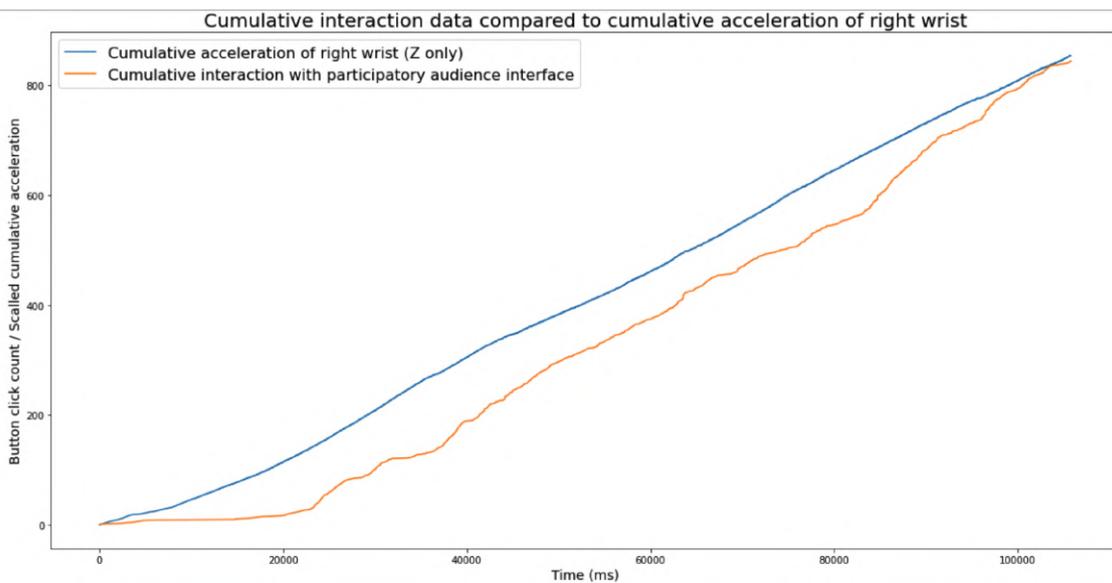


Figure 7.12: Cumulative interaction data compared to cumulative acceleration of right wrist

for better comparison, as shown in figures 7.13 and 7.14.

**A closer look at instrument-specific audience interaction data** For a better understanding of the effects that certain parameters have on the drummer’s motion, the interaction data is filtered for specific buttons that seem to be most impactful on the performance behavior. Figures 7.15 and 7.16 are graphs similar to the ones above but now showing the button affecting the decay value for the snare in both upward and downward directions only.

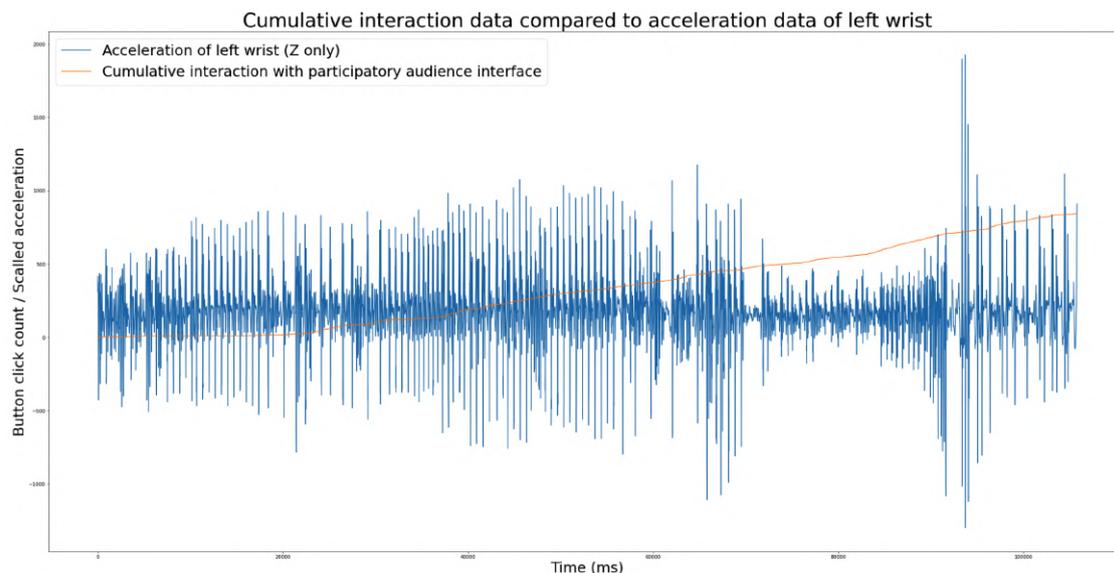


Figure 7.13: Cumulative interaction data compared to acceleration data of left wrist

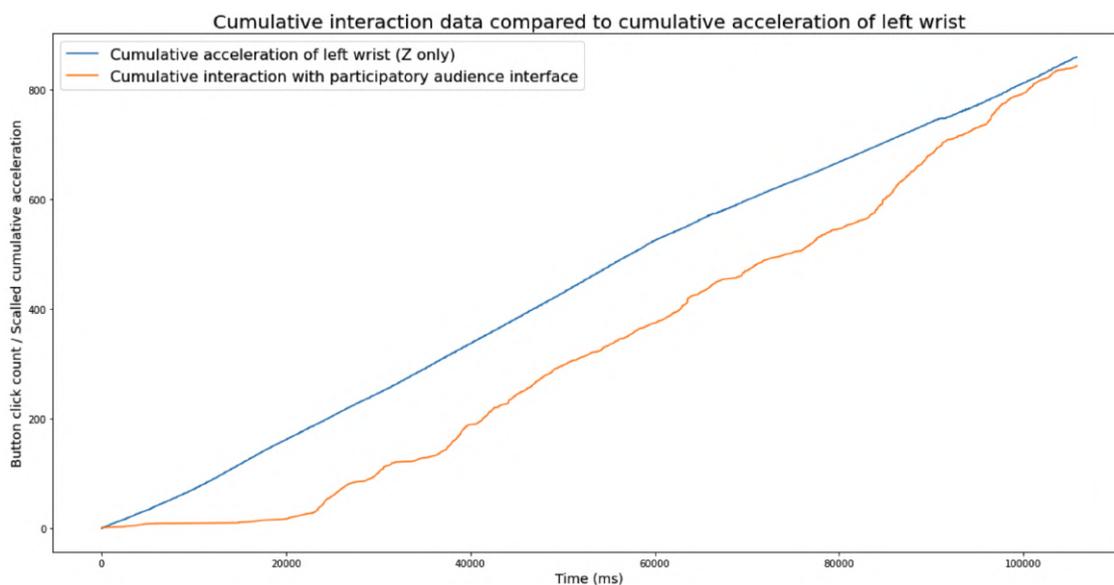


Figure 7.14: Cumulative interaction data compared to cumulative acceleration of left wrist

## 7.2.5 Discussion

As mentioned in the analysis section, the changes made to the drum sounds seem to affect the performer's body motion at certain points, based on video and accelerometer data analysis. During critical changes in particular, such as changes to the decay value of the snare drum, there are effects of more reticent play in the interests of exploring these changes during the performance. Another strategy of this exploration involves more repetition on the same instrument in order to better grasp how the parameters are changing. An evaluation of the personal interview with the performer is needed to confirm or discard these initial findings.

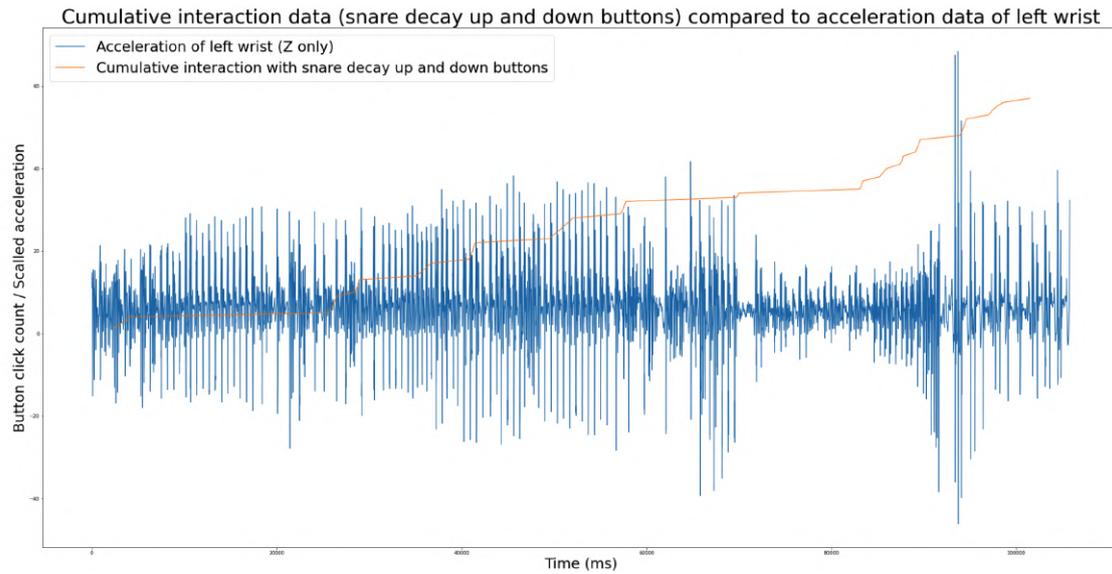


Figure 7.15: Cumulative interaction data (snare decay up and down buttons only) compared to acceleration of left wrist

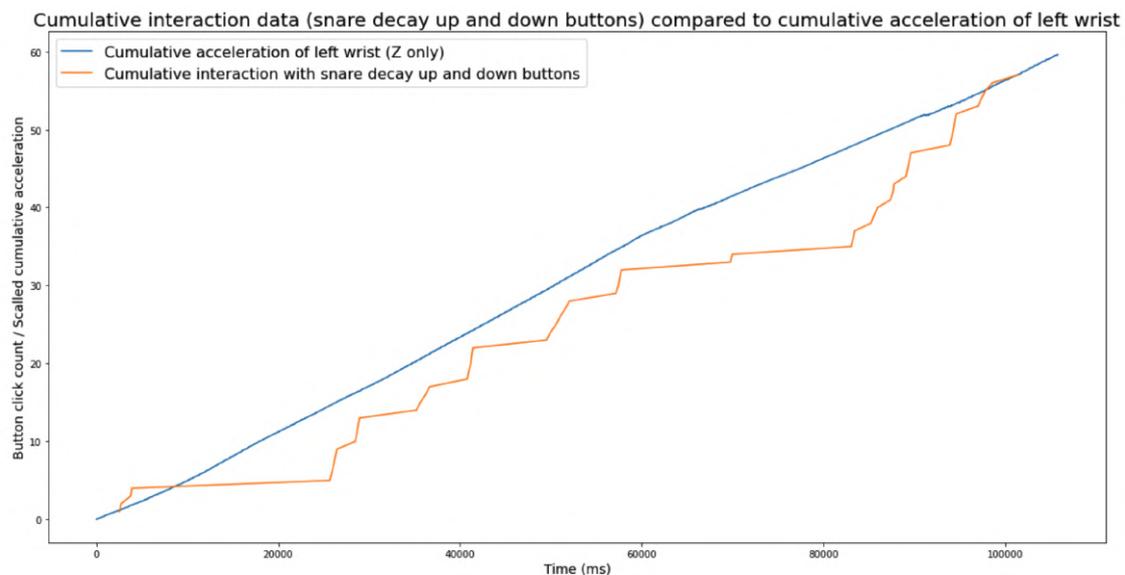


Figure 7.16: Cumulative interaction data (snare decay up and down buttons only) compared to cumulative acceleration of left wrist

The comparison of the interaction data with the sensor data of the accelerometers suggests some level of correlation between the total volume of participatory interactions and the drummer's body motion. However, a direct effect of the audience's participation upon the performer's motion cannot be derived from the analysed data. Confounding factors such as the musical structure of the piece, the interaction between the stage performers themselves, constraints in the interaction interface, and the listening focus of the drummer (including his monitoring system) could be at play here and demand more systematic analysis. The selection of a certain change parameter (snare decay buttons) may have a direct relation to the motion of the hand playing that instrument (the left hand), but it may also be that this relation

arose instead as an effect of the overall increase in musical intensity toward the end of the experimental section.

While collecting the interaction data, I saw that the audience members needed a certain amount of time to explore the interface, and that this impacted the ongoing performance. In order to minimize this distortion, an exploration phase should be implemented in future experiments. In addition, it is critical to have clear auditory feedback concerning the audience's individual actions or a better understanding of the effects these actions will have on the instrumental sounds. This would lead to a better understanding of both performer and audience as co-creating partners during performance.

Finally, the instrument parameters that were selected for the experiment should be reconsidered. Whereas decay values seem to relate to body motion to a certain degree, pitch values are not as critical to changing the percussive sounds. It might be better to replace these with other values of the ADSR curve, such as the attack values of the sounds. This might also affect the length of the chosen experiment - two minutes was very short as the section length for exploration. Similar setups with longer recording periods would generate more data that could also help to track some of the confounding factors and limit experiment noise, such as random actions among the audience.

## 7.2.6 Conclusions and Further Work

The question of how audience participation via personal smartphones affects performer's body motion was explored through both qualitative and quantitative data analysis. Quantitative analysis revealed a tendency of the performer to react both more sensitively and more experimentally during the performance. The numerical analysis showed some relation between specific buttons (namely, the snare decay buttons) and the performer's lefthand motion, and it is thought that the overall "musical intensity" of the performance might impact the extent to which audience members used the interface in order to take part in the performance.

However, more exploration of the data would be needed to show statistically significant direct relations between interaction data and performer motion. For this, one should calculate and filter the sensor data of all the accelerometers (including the ankle band on the performer's foot) and combine these with more elaborate video analysis tools in order to assess the motion quantity and qualities of the performing drummer more accurately. Additionally, we plan to integrate the MIDI data into the research at a later point, and we need to analyze the evaluative questionnaires of the audience members and the in-depth interviews with the performers.

### 7.3 Conclusion of the Preliminary Studies

Looking back at the two studies, it becomes clear that audience gestures represent a direct and immediate way for concert audiences to give feedback to performers through bodily gestures such as clapping or vocalizations. When technology is introduced to accommodate audience participation, this direct feedback becomes both more limited and more detailed at the same time. While audience members might still react via clapping or vocalizations, they also enjoy a new opportunity to engage with the performers in a more direct musical sense. The second study further showed that the audience's ability to improvise with performers reveals certain dialectic patterns that align in various ways. The study also showed that performers need to be open to adjusting their habits and even surrendering some of their performative agency to properly accommodate the audience's contributions.

# Chapter 8

## Article Summaries

This chapter summarizes and contextualizes the four articles presented in this article-based thesis. Figure 5.2 gave an overview of the studies and their connection to the theoretical framework presented in part 1 of the thesis. Studies 1-4 in the graph correspond to the published articles 1-4, respectively. While the first two articles are focused on the participatory design and present empirical studies, the third article instead reviews existing systems with the aim of building a theoretical model that accommodates speculation about future system designs. Article 4 then presents some alternative ideas for how to implement participation and is intended to inspire further work on interactive, autonomous music systems.

## 8.1 Article Overview

### Article 1

Jung, M. and Kummen, V. 2024 (published). "Hacking the Concert Experience: Exploring Co-Creative Audience Interaction at a Chiptune Live Performance" in G ullo, J.-O. and Hepworth-Sawyer, R. and Paterson, J. and Toulson, R.: *Innovation in Music: Technology and Creativity*, Routledge. <https://doi.org/10.4324/9781003118817-8>

### Article 2

Jung, M. and Clester, I. 2023 (published). "Hack the Show: Design and Analysis of Three Interaction Modes for Audience Participation" in Buffa, M. and Alexandraki, Ch. and Lebrun, J. and Roberts, Ch. (eds.): *Fostering Creativity through Web Audio*. *Journal of the Engineering Society (Special Issue)*. <https://doi.org/10.17743/jaes.2022.0111>

### Article 3

Jung, M. 2023 (published). "Intelligent Music Performance Systems: Towards a Design Framework" in *Studia Musicologica Norvegica*. Universitetsforlaget. Vol.49, Iss.1: 28-44. <https://doi.org/10.18261/smn.49.1.3>

### Article 4

Jung, M. 2023 (published). "Beyond Mutation: How Can We Acknowledge Symbiogenesis in Evolutionary Music Coding?" in *Proceedings of the International Conference on AI and Musical Creativity*. Brighton. Available at: <https://aimc2023.pubpub.org/pub/if7rl6o5>

## 8.2 Article 1

### Article 1 - Hacking the Concert Experience: Exploring Co-Creative Audience Interaction at a Chiptune Live Performance

Authors: Matthias Jung and Vegard Kummen

Published in Güllo, J.-O. and Hepworth-Sawyer, R. and Paterson, J. and Toulson, R.: *Innovation in Music: Technology and Creativity*, Routledge. Video documentation: <https://tinyurl.com/4xuphf7z>

### 8.2.1 Background, Research Aims and Questions

The first article ended up as a book chapter in a publication that emerged from the conference "Innovation in Music," which took place in June 2022 at the Royal College of Music (KMH) in Stockholm. The article's aim with this work was to explore audience participation practically after our review of many different concepts and approaches theoretically. Many of the participatory works in the field were too schematic and lacking in the profound musical integration of the audience into the compositional design. The project started with by designing a performance with an existing musical project (Vegard Kummen's band Kubbi) that integrated musical parts for the audience to provide at certain stages and in different performative roles. The research aim was to compare different paradigms for audience interaction both quantitatively and qualitatively to better understand the audience members' experiences and preferences. In addition, we wanted to get an idea of the performers' perspective on the interactive parts of the show in order to discuss the relationship between the experiences of both groups. Our main questions were thus directed toward the audience's and performers' experiences, the relations between them, and the consequences of this insight for the design of participatory systems for use in the given cultural context of chiptune live performance. The concrete questions suggested in the article are as follows:

How do audience members experience the performance when they are invited to participate in it musically?

How do participatory paradigms change the performance from the perspective of the performers on stage?

Which interaction paradigms are best suited to chiptune live music?

### 8.2.2 Methods and Research Design

The empirical study was designed as a research concert that lasted about one hour and was announced as a free event that people could attend<sup>1</sup>. We wanted an ecological study that would attract audience members just like any other band show,

---

<sup>1</sup>The research design, including the handling of personal data and the survey question texts, was confirmed with the Norsk senter for forskningsdata (NSD, now SIKT) in early 2022 before the concert study.

though attendees were admittedly informed about the study aspect of the concert in advance. The musical content of the performance was similar to the band's usual set, and the interactive parts were integrated into existing musical compositions. These compositions were designed in an iterative manner; we started with approaches we found in the literature, then implemented them in Max MSP, created the sound design, rehearsed them with the band, and finally adjusted the participatory design of the interface. The interface was realized as a browser-based smartphone application that logged the contributions quantitatively in order to determine how actively audience members were participating at any given time. After the concert, a study-specific self-report questionnaire was offered to the audience members. It was based on previous research work in the field, including aspects of the individual, technological, and social experience of the participatory concert. In addition, two semi-structured interviews were conducted after the concert, one with about one-third of the audience members and one with the performers.

### 8.2.3 Findings and Contributions

Data from the interaction log and survey of the first concert study can be found in this [open science repository](#). Video documentation of the study can be accessed [here](#).

Major findings of the participatory design process and interactive concert were that audience members were excited about the participation but did not necessarily feel that their musical experience was enhanced. We also found that a playful interactive paradigm to most enhance the musical experience. The data showed that performers preferred improvisational parts through which they felt inspired by the audience's agency. The self-reports suggested that audience members felt more connected to their peers and the performers during the interactive parts than during the non-interactive parts of the concert. Lastly, the experiment indicated that the benefit of participatory settings resided less in the musical outcome of the performance than in the refined listening experience reported by both audience members and performers.

The research contributes insights into the compositional design process involved in participatory work in an ecological concert setting and presents arguments for the decisions made regarding the specific aesthetic of chiptune music. It suggests and discusses a new measure of the activity of audience members contributing to musical aspects, named *Quantity of Participation (QoP)*. This measure, together with the questionnaire and interview results, offers insight into both the audience's and the performers' experiences of three different participatory settings that were explored during the standard musical set of the band Kubbi. Lastly, the work contributes by comparing those different paradigms of performer-audience interaction and discussing how to design participatory systems and how participatory settings affect the listening experience at live concerts.

## 8.3 Article 2

### [Article 2 - Hack the Show: Design and Analysis of Three Interaction Modes for Audience Participation](#)

Authors: Matthias Jung and Ian Clester

Published in Buffa, M. and Alexandraki, Ch. and Lebrun, J. and Roberts, Ch. (eds.): *Fostering Creativity through Web Audio*. Journal of the Engineering Society (Special Issue).

### 8.3.1 Background, Research Aims, and Questions

The second article was co-written with Ian Clester at Prof. Jason Freeman's lab at the Center for Music Technology at the Georgia Institute of Technology in Atlanta during my semester there as a visiting scholar. Together with Vegard Kummen, we planned a second concert with Kubbi, this time at Kristiansand's theatre as part of Norwegian Research Days 2022 with the full band including a keyboard player. With this research concert, we tried to incorporate the insights from the first concert and implement them in a new version of the interaction software while using the same self-report questionnaire after the concert. Our first experiment showed that the first interactive paradigm, wherein the audience acted in the musical background and took an environmental role, scored higher in the audience's ratings. To explore it further, we expanded the audience's interactive parts by adding light designs to the musical participation and hired a lighting designer for the project. We hoped to add an important component of the aesthetic experience of staged concerts and experimented with different interaction paradigms that involved color and directions for the 16 moving head lights that were used for the concert.

In addition, Ian helped us implement a more native version of the interaction software that replaced the Max patches that were used in the first concert, which, we suspected, caused delays when connected to the WebSocket design. This also allowed us to open-source the project so that other composers and performers could use the software, which is written in Javascript (clients) and Python (server). Moreover, the software now also allowed us to analyze individual, anonymized user behavior so we could acquire more information about how audience members made use of the interface.

Rather than focusing on the participants' experiences, the main research questions of this second article revolve around the behavior of audience members across the different interaction types. We wanted to use this behavior to inform our recommendations for the design of further compositions and interactive systems. The questions of the article are as follows:

How do different interaction modes affect collective audience behavior?

How can the audience members' contributions be structured on an individual level?

How can the findings be applied to the design of participatory systems?

### 8.3.2 Methods and Research Design

The design for approaching those questions this time tended toward the analysis of the richer log data, which is why only the questionnaire (not interviews) was used to capture the experiential dimension of the concert. The log data analysis again used the quantity measures for participation (QoP), but this time we could take it a step further and perform cluster analysis in order to find common user behaviors with the interactive software. In that sense, the software was used for both musical contribution and analysis at the same time. In addition, we used the answers in the self-report questionnaire to inform our discussion of the experiences of the audience members.

### 8.3.3 Findings and Contributions

Data from the interaction log and survey of the second concert study can be found in this [open science repository](#). Video documentation of the study can be accessed [here](#).

Results of the log data analysis show the highest engagement during the third interaction, in which audience members could control the directions of the moving head lights on stage. The next highest engagement occurred during the first interaction, in which participants controlled light colors and environmental sounds in the musical background. The lowest interaction numbers accompanied the second one, where audience members controlled parameters of the main synthesizer melody in the musical foreground in tandem with a rhythmically related light effect. The analysis of the individual contributions indicated five different clusters of user behaviors, and this finding was supported by a discussion based on the results of the post-show questionnaire. Most of the participants stayed moderately active throughout the interactive concert and explored the collaborative space without trying to control one of the buttons altogether. This indicates that the audience members were open to the concept, and only smaller groups of them tried to "hack" the buttons or became disengaged during the interactions.

Further work should include studies with a higher number of participants to validate the proposed user typology or develop a more refined understanding of interactive audience behavior with additional surveys and interviews. In addition, an analysis of the questionnaire data in relation to the individual, anonymized audience contributions themselves could begin to explain the connections between the audience members' backgrounds, musical training, and personality traits and their interactive behavior during the participatory concert. Finally, interviews with the performers would add the perspective from the stage and, in turn, round out the understanding of collaborative concerts.

## 8.4 Article 3

[Article 3 - Intelligent Music Performance Systems: Towards a Design Framework](#)

Author: Matthias Jung

Published in *Studia Musicologica Norvegica*. Universitetsforlaget. Vol.49, Iss.1: 28-44.

### 8.4.1 Background, Research Aims, and Questions

While the first two articles discuss experiments within an ecological concert setting, the third article looks at participatory systems that involve levels of autonomy within them. It acts as an extension of the response to question Q2.3 about the construction of participatory systems. The principal aim was to develop a better understanding of how autonomous music systems are integrated into participatory systems and what this integration does to our idea of technology-enhanced live performance scenarios. A secondary aim was to see whether the model, which was constructed with the help of musical proxemics and studies of musical AI, would be suitable for capturing some of the most prominent systems of recent decades and what tendencies those systems would show in relation to it. Lastly, those tendencies are made to inform speculation about trends in the design of Intelligent Music Performance Systems (IMPS). Though not specifically raised in the article, its main three questions were as follows:

What are the main aspects to be considered when designing IMPS?

Where can current models be placed in the model (case studies)?

How can we use the suggested model and conducted case study for speculating about future iterations of IMPS?

### 8.4.2 Methods and Research Design

To build its analytical model for IMPS, the article first reviews existing related frameworks from musicology, phenomenology, and design studies. In addition, word searches in prominent areas of publication such as the NIME (New Interfaces for Musical Expression) community are used to support the relevance of the topic. After integrating concepts of perceived distance (proxemics) into studies of autonomy, the article shifts to an exploration of cases following previous models. Those cases are presented together with their historical and cultural context, then placed into the analytical framework. Finally, the article uses methods from Speculative Design Research to better understand current tendencies and developments in the design of IMPS.

### 8.4.3 Findings and Contributions

The review of current models for understanding IMPS revealed a lack of theoretical concepts, especially in the domain of autonomous music systems. The combination

of Hall's theory of perceived distance together with Eigenfeldt et al.'s autonomy scale proved to be a useful approach to categorizing past and current cases of IMPS. Despite its small number of case studies, the analysis revealed a tendency toward one of the three axes of the model. Using Speculative Design Research, the article then extrapolated those tendencies to further cases of IMPS and considered preferable future systems.

The main contributions of this work reside in the creation of a new approach to conceiving IMPS, that might be helpful for other researchers in the space. The application of proxemics in musicology is not new, but its use in the context of performative technology represents a novel theoretical approach to the field. Moreover, the article contributes a new framing of IMPS in general and a theoretical discussion of the current tendencies and priorities of IMPS design processes.

## 8.5 Article 4

[Article 4 - Beyond Mutation: How Can We Acknowledge Symbiogenesis in Evolutionary Music Coding?](#)

Author: Matthias Jung

Published in Proceedings of the International Conference on AI and Musical Creativity. Brighton.

### 8.5.1 Background, Research Aims, and Questions

This article was inspired by the study and use of genetic algorithms in real-time generative music systems - for example, when navigating previously trained music systems with the input of a live human performer. A review of the biological metaphors around this algorithm class makes it clear that those principles have evolved since their introduction in the 1920s, which is why further investigation was warranted. The main aim of this investigation was to examine the development of those underlying, biologically inspired principles of IMPS and explore possible alternatives. It soon became clear that symbiosis and symbiogenesis merit further research in this regard. The article responded to the following research questions:

What are beneficial ways to simulate short and longer-term symbiotic relationships that would be helpful to musical applications?

How can we combine interdependent populations in a way that helps the creative outcome?

How can we usefully integrate environmentally inspired concepts such as the Gaia theory?

### 8.5.2 Methods and Research Design

Since this article answered a conference's call regarding alternative concepts, its method was to review the literature and then develop ideas regarding how an implementation might be approached in order to gauge its prospective value and utility. The article text was presented at the conference and discussed with participants, which led to more questions and ideas for further work.

### 8.5.3 Findings and Contributions

The main finding of this work to date is that it might be worthwhile to move ahead with an exploration of the biological principles of symbiosis and symbiogenesis in more depth. First of all, the different forms of symbiosis can be used as a metaphor for the different musical roles within a performance context. Secondly, on a more theoretical level, models from computational biology could be used to implement generative systems that are similar to neural networks or other biologically inspired interdisciplinary imports. The contributions of this work at this point are very limited but will be discussed a bit further in the conclusion to the thesis in chapter 9.



# Chapter 9

## Discussions and Conclusions

Following a discussion of the thesis findings in light of the suggested concepts, this chapter revisits the initial research aims and questions from the introductory chapter 1. It then summarizes the main empirical, methodological, and theoretical contributions of this work. Lastly, it reviews the limitations of the project and suggests further directions for investigation.

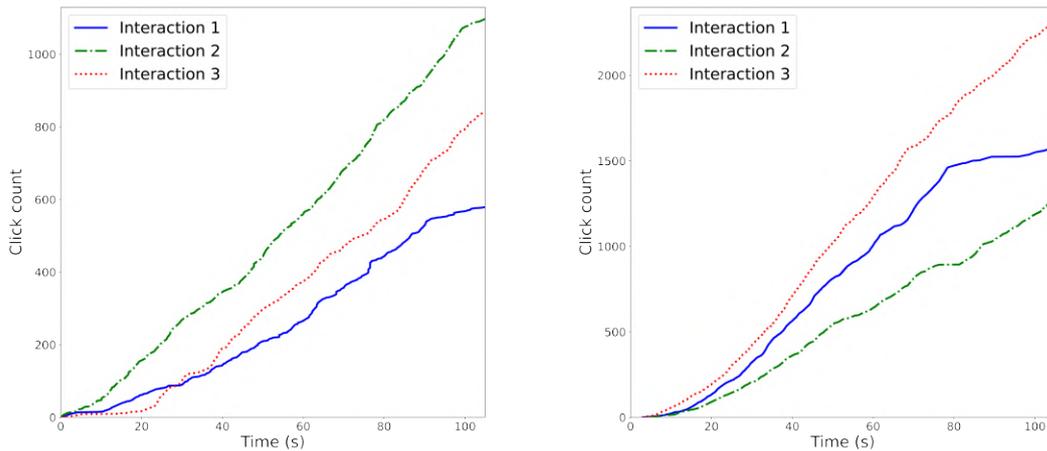
### 9.1 Final Discussions

The discussion below frames the findings of the thesis in relation to the suggested concepts summarized in chapter 5. It begins with audience gestures and audience instruments, revisiting the findings of studies 1 and 2, then considers IMPS and the alternative approaches to coding proposed in studies 3 and 4.

#### Audience Gestures and Audience Instruments

We will begin by reviewing the participatory studies and their insights with regard to audience engagement. Figure 9.1 shows cumulative plots for the clicks of all audience members during the first 100 seconds of the three interactive parts of the two participatory concerts. There are similarities among these plots despite the fact that any direct comparison of the concerts has limited validity due to the ecological differences of their settings (including different venues, numbers of participants and performers, and musical material). For example, during the first mode, audience members controlled the background of the performance; during the second mode, they controlled the melodic foreground; and during the third mode, a playful interactive paradigm was established for participation. In the first experiment, this paradigm involved control of the instruments; in the second concert, it involved musical effects and directional lighting. Looking at the extent of participation for interaction modes 1 and 2, we see that the interaction with the performance's background (mode 1) was less than the interaction with its foreground (mode 2). This relation was inverted for the second experiment. Note that the interface for the first concert experience had more buttons than the interface for the second. This might explain the different extents of participation but does not apply to the third

interaction mode, where the extent of participation for the second experiment was highest despite only four control buttons in the interface. This suggests that the number of buttons might not have had a simple or straightforward impact on the extent of participation of the audience members. Lastly, while we saw the highest level of participation during mode 3 of the second experiment, audience members participated at a middling level during mode 3 of the first experiment.



(a) Quantity of Participation for Concert Experiment 1

(b) Quantity of Participation for Concert Experiment 2

Figure 9.1: Comparison of Quantity of Participation

Table 9.1: Survey results for the three interaction modes for concert 1 and concert 2 (in brackets) compared.

	Mode 1	Mode 2	Mode 3	None
Which interaction mode did you enjoy the most?	87.5% (5.6%)	0% (11.1%)	12.5% (66.7%)	0% (11.1%)
Which interaction mode did make you feel most excited?	87.5% (27.8%)	0% (22.2%)	12.5% (33.3%)	0% (11.1%)
Which interaction mode enhanced your musical experience the most?	62.5% (44.4%)	0% (11.1%)	25% (22.2%)	12.5% (22.2%)
Which interaction mode encouraged you to participate the most?	75% (11.1%)	0% (22.2%)	25% (44.4%)	0% (16.7%)
During which interaction mode did you feel that the musicians reacted most to what the audience was doing?	37.5% (16.7%)	25% (27.8%)	0% (27.8%)	37.5% (27.8%)

When comparing the results of the self-report questionnaires for both experiments as depicted in table 9.1 we can see that Mode 1 was the clear favorite among

the three interaction modes in the first concert. At the same time, it displayed the lowest QoP of the three modes, indicating that active participation does not equate with greater enjoyment or excitement regarding the musical experience. On the contrary, it was Mode 3 which was most enjoyable and exciting for the audience members in the second concert experiment. In this case, high participation actually coincides with the highest QoP values, even though the interface featured only a few buttons on an x-y plane. In addition, audience members could not experience the stage performers' reactions to their contributions, suggesting that the interaction design should offer more opportunities for this engagement.

Figure 9.2 shows a graph of the user clusters as presented in article 2 (section 9.5). This data was only available for experiment 2 in the form of individual anonymized interactions. It shows that most audience members were quite open to participatory settings and stayed connected to the interface across all the interactive modes. The *reserved* and *exploratory* user groups displayed relatively low dwell times, while the *sporadic*, *engaged*, and *consistent* user groups all remained engaged throughout the concert. This is true for the first concert experiment as well, whose participants reported a general interest in taking part throughout the concert. However, only the expert users (musicians or individuals who scored high on openness in their personality) displayed high values of repeated button clicks—for example, the *consistent* and *engaged* user groups in experiment 2 and the expert users in the audience group interview after experiment 1. The biggest user group was called the *sporadic* group in experiment 2—its members stayed throughout the concert and participated from time to time without necessarily trying to gain control over any parameters. Similar user behavior was also reported during the audience interview for experiment 1, where most of the participants reported an interest but not overly active participation.

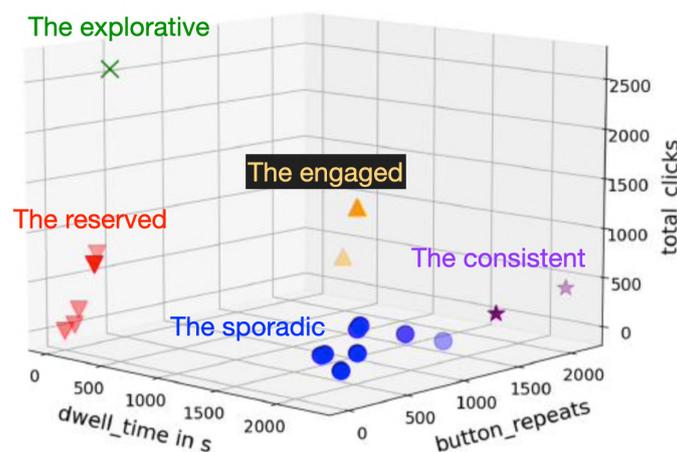


Figure 9.2: User clusters

Audience instruments are multi-user instruments with an asymmetrical structure between a smaller group of stage performers and a larger, much more diverse group of audience members. While this asymmetry creates certain challenges in terms

of replicating the typical back-and-forth interaction of musical ensembles, audience members prove to be generally open to contributing to performances even when most of them are not trained musicians. Audience instruments make it possible to turn audience gestures into musically meaningful contributions with which stage performers can interact because they are "translated" into their "musical" language in real time. They are thus epistemological tools (Magnusson, 2009) with multiple interfaces that tap into the current collective (emotional) state of the audience, unleashing it as tacit knowledge (Reuter, 2021) in both auditory and visual domains. In effect, a concert could itself be seen as an instrument that the stage performers, audiences members, and multitude of professionals curating, planning, and implementing the event "play" together. This recalls Simon Frith's understanding of performances as events that are initiated, designed, and supported by contributors far beyond the stage performers themselves (Frith, 1998). Audience instruments enable audience members to realize their gestures in real time and lean in to what they are listening to. They also adjust the expectations of both performers and audience members and nudge them toward a refined yet accessible co-creation process. Zooming out, audience instruments are one aspect of the larger concept of Intelligent Music Performance Systems, to which I will return in the following section.

## Intelligent Music Performance Systems

Designing Intelligent Music Performance Systems (IMPS) involves orchestrating a mix of agencies, including those of the designers and composers, the performers, the audience members, the intelligent objects, and so forth. These agencies are simultaneously embedded in culturally specific bodies and environments, as the model suggested in article 3 understood. The agency of autonomous, co-creating technologies can be seen as a representation of historical, enacted knowledge, and, as such, it affords new ways of interacting in performative contexts. AI in that regard is a representation of our past creative selves that enters the network—as a space of relationships—so that we can interact with it in a form of collective creation. These practices demand an updated understanding of group or distributed creativity.

The model also argues for the idea that the felt distance from other human and non-human agents is intertwined with their autonomy<sup>1</sup>. Because technology is now even entering our peripersonal space<sup>2</sup>, in some cases overlapping with or even penetrating our bodies, it constitutes an ontological change and disrupts the idea of performance at the same time. Fraunberger insists that people's "intimate entanglement with digital technologies is challenging the foundations of current HCI research and practice" (Fraunberger, 2019, 1). From the perspective of posthumanism and

---

<sup>1</sup>Note that the model suggested in article 3 only uses autonomy levels 0-5 as suggested by (Eigenfeldt et al., 2013). Levels 6 (Versatility) and 7 (Volition) are not especially evident in current systems because they tend to operate using training data that pre-defines their stylistic behavior (level 6), and they are not yet able to voluntarily decide where they will perform (level 7).

<sup>2</sup>Peripersonal space is the space within a person's physical reach. For an extensive discussion, see (Di Pellegrino and Làdavas, 2015)

ANT, he calls for design practices that go beyond user-centric designs because the human user is less central to the design when their agency is shared amongst other agents. He then points to Baumer and Brubaker's term *post-userism* (Baumer and Brubaker, 2017), which describes "instances where the classical conception of the user breaks down" (Frauenberger, 2019, 19).

This posthumanist view of IMPS design shifts our attention away from the human user and toward non-human agents and environmental aspects. Consequently, the dimensions *instruments*, *systems*, and *agents* suggested in article 3 all reveal aspects of this entanglement and address tendencies in system design within the proxemic space. As a consequence, the dimension of autonomy can also be approached from a non-user-centric perspective—that is, autonomous from whom/what? The perspective of the human user? The other agents? The environment? Beyond this, even, the idea of autonomy in relation to IMPS is not limited to the artificial agents; it concerns the other agents as well. Less frequently discussed here is the question of how dependent the system design might make us humans on AI technology, and to what extent it supports our creativity if we remain autonomous from it. That is, performers using AI as part of their instruments might come to depend more on it, whereas audience members who are not using it will experience things differently. Lastly, there are levels of autonomy between performers and audiences that do not involve AI at all, such as the audience's support of the artists, which often enables the performance in the first place, or the audience's dependency upon what the performers offer on stage. Consequently, the IMPS model must attempt to reconcile those autonomies with the perceived closeness of the agents, thereby centering the entanglement of these conditions among the many other aspects of IMPS.<sup>3</sup> The model argues for the virtues of systems to be found at the intersection of these design tendencies. Note that this is not meant to suggest a deterministic view or the ideas of conversion culture that have circulated in media studies (Jenkins, 2006). Instead, it points to an area of speculative design that might be fruitful for such explorations.

## Alternative Approaches to Coding

The final section of this discussion looks at ways to simulate short and longer-term symbiotic relationships to help musical applications; the inter-dependency of populations; and environmentally inspired concepts of computation. This work is just beginning but merits consideration here (and in the sections on "further work" in the area).

Figure 9.3 depicts the flow diagram of genetic algorithms used in music, together with the different forms of symbiosis found in nature. Note that symbiogenesis describes evolutionary mechanisms involving symbiotic forms specifically in terms of the transitions of those organisms from separate cultures into merged ones. In the

---

<sup>3</sup>Note that perceived closeness does not lead to either less or more autonomy as such. A well-designed IMPS gives users great creative autonomy despite the perceived closeness that is unavoidable in a performance space..

case of mitochondria, which were bacteria before they took on the power-generating function in our cells, an implementation of this principle would describe that evolutionary change.

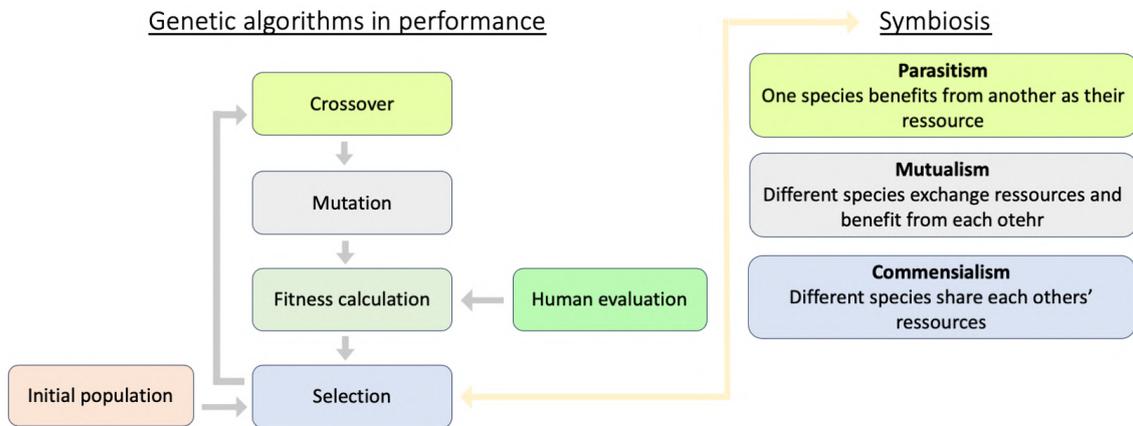


Figure 9.3: Schematic graph of genetic algorithms for music

Figure 9.4 presents some initial suggestions for implementing symbiotic concepts into participatory music contexts, as discussed in article 4<sup>4</sup>. While mitochondria themselves do not perform complex tasks, they still represent a productive metaphor for generating a creative outcome. Agre advocates for such a generative metaphor that constructs a "mapping between two semantic fields, and in doing so it inevitable draws certain elements of each field in the foreground and relegates others to the background" (Agre, 1997, 44).

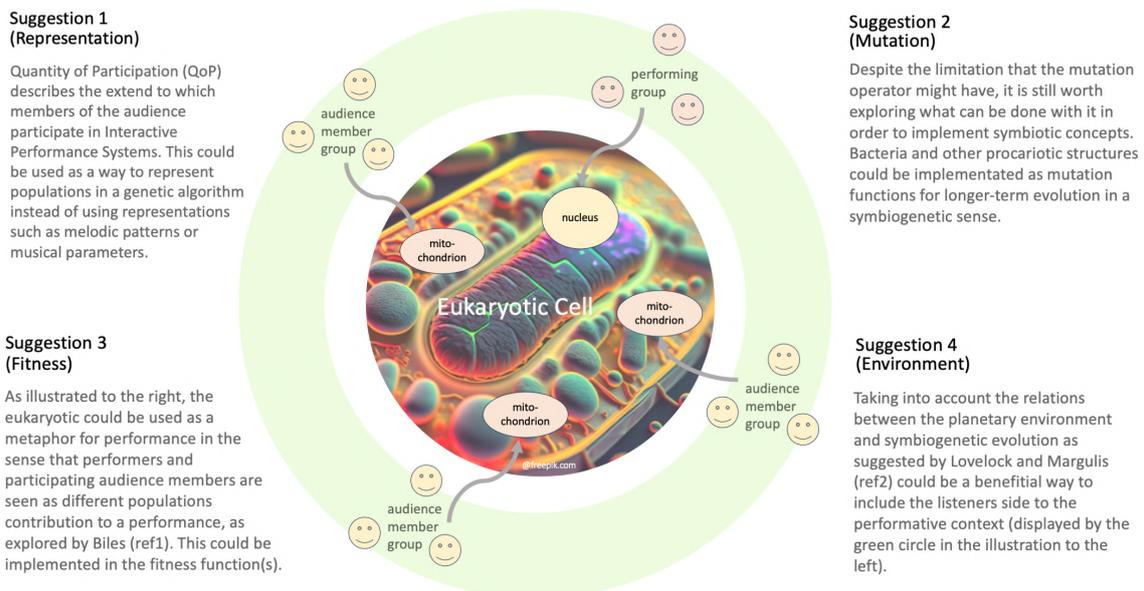


Figure 9.4: Suggestions for symbiotic concepts applied to the context of IMPS

Entering into a symbiotic relationship means surrendering some control to other

<sup>4</sup>The references in the graph are as follows: ref1 (Biles and Eign, 1995) and ref2 (Lovelock and Margulis, 1974), copyrighted picture used in line with license rules @freepik.com.

actors. In the case of participatory music systems, that is, composers and performers cede some agency to audience members. At the same time, they experience something new in what can be a fruitful trade-off between giving up control and gaining inspiration and collaboration. In that sense, Fraunberger's notion of the entanglement of human and non-human actors is symbiotic (Fraunberger, 2019).

Figure 9.5 attempts to capture this idea of entanglement by integrating the dependencies of the different actors to further develop the aforementioned concept of autonomy during co-creative processes in IMPS. Entanglement, unlike autonomy, is two-sided. On the other hand, symbioses can be more or less collaborative, which leads to the four quadrants of co-creative interaction. Returning to the idea of proxemics in co-creation, symbioses add certain qualities to those levels of perceived closeness that the perception itself does not encompass. How antagonistic or cooperative the perceived closeness turns out to be, however, depends on factors beyond simply the perception of closeness (including environmental ones, for example).

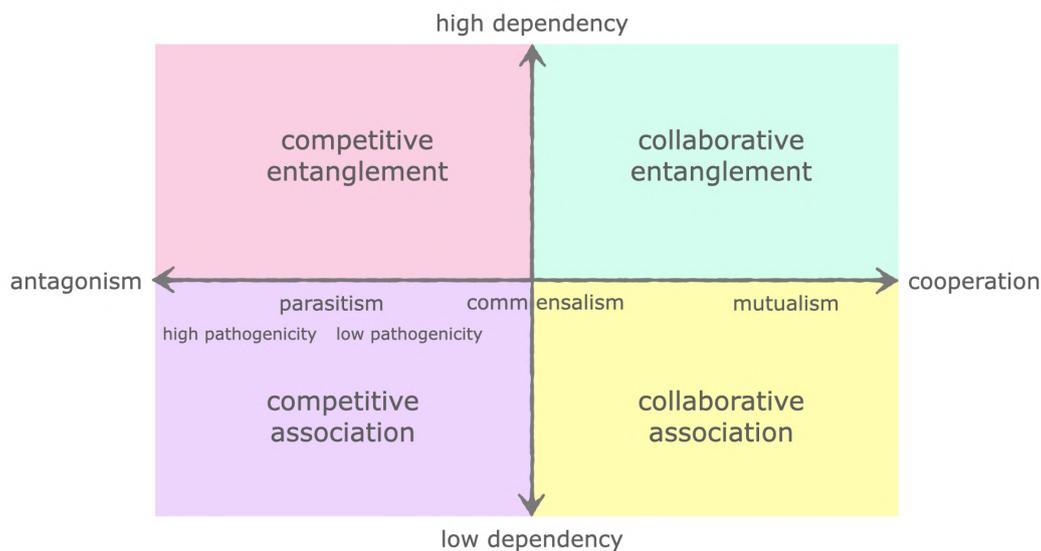


Figure 9.5: Integration of entanglement with symbiosis

In turn, artificial intelligence can be understood to emerge via the symbiotic collaboration of human and non-human agents, all with inherent forms of intelligence, as an extended or emergent condition. We might even speak of symbiotic intelligence—an intelligence that emerges as a consequence of the interactions among different agents and reflects the qualities of that collaborative process (that is, the dependencies and qualities of the perceived closeness). Ultimately, those collaborations can be tied to the understanding of affinity spaces, as suggested by Gee (Gee, 2005). Further exploration of symbiotic systems and their possible impacts upon the field of music technology goes beyond the scope of this thesis but will be addressed below under "further work."

## 9.2 Research Questions Revisited

This thesis's aim was to better understand the compositional affordances of participatory technology when used by performers, audiences, and artificial agents. It looked at how interactive performance technology in general, and participatory performative works in particular, impact the concert experience. The starting vantage point was Ihde's postphenomenological claim that technology that amplifies experience simultaneously takes other aspects of that experience away. By creating two participatory concert events, the research adopted a practical and exploratory approach that included analyzing the interactive behaviors and interviewing the stage performers and audience members. Following this work, the research focused on systemic aspects and the role of autonomous technologies and suggested a model for those systems that integrates notions of perceived distance and the autonomy of the different agents. Finally, the research revisited the existing concepts concerned with facilitating co-creative performance and proposed future avenues of biologically inspired approaches to co-creation with audience instruments.

### Research question 1

The first research question centered upon the experience of interactive live performance, and the relevant findings were summarized in section 8.2. Sawyer's reference to Brenneis's assertion that "the relatively passive role of European audiences may have caused Western researchers to neglect the pervasive importance of the audience in musical genres of other cultures" (Sawyer, 2003, 131) is an excellent prompt for this particular line of questioning<sup>5</sup>. The study's discussion encompassed the audience's and the performers' behavior and observations, as well as different participatory designs for audience instruments. It approached this area of focus with an exploratory design of three interaction modes to be used at a live show with a small audience at a live music venue. The participatory enhancements for the concert included both traditional elements of music composition, such as sliders for controlling musical parameters, as well as rather unconventional interfaces for changing the instrumental sounds used by the stage performers. Motivated by concepts of musical interaction and embodiment, the project recorded performer movement data that was later used in a preliminary study relating movement to participation.

Analyses of the interaction log for the audience instrument and the results of a post-concert survey and interviews with both audience members and performers indicated that both groups are generally open to participatory experiences. This was especially true if the participatory parts were clearly marked as such and the concert retained its familiar characteristics—that is, the artists remained the performative

---

<sup>5</sup>Brenneis, at the end of the cited article, asks a number of related questions about the relationship between performers and audiences: "To what extent are people involved not just as collaborators (...), but as active co-creators of what's going on! What expectations do performers have of the audience and audiences of the performers? What kinds of interactions are there between them? How does the audience contribute to the emergence of music, and on what terms does it define and shape meanings and values?" (Brenneis, 1990, 188).

center of the event and the participatory parts were merely enhancements, not replacements, regarding this material. Keeping this in mind, we sought to integrate the participatory parts as smoothly as possible, so we avoided any introductory review of the functionalities and musical controls of the audience instrument and instead offered an exploratory presentation. For example, we started the concert without the stage performers so the audience members could play with the sliders and sounds; when the stage performers did appear, the audience interaction was already well underway. This interactive mode worked well for the participants because it perpetuated their positioning as contextual rather than focal to the stage performance itself. Still, the audience instrument did afford their interaction with the central musical elements of the performance, including crucial melodic parameters such as volume, cutoff, and the decay of a sequencer. In this case, we found that the audience members did not fully exploit these possibilities during the participation, preferring to leave these choices in the hands of the performers. This hesitation, of course, might be why participants reported that their musical experience was not enhanced as significantly as it could have been. The third participatory paradigm might have been more successful because it did not take away from the traditional performative center of the stage musicians, but it did give the audience members access to the sounds they heard there. In addition, changes to the pitch and decay of the drum sounds were easier to perceive than other kinds of changes. In the end, we argued that this was an informal learning setting that corresponded to how musicians learn informally in the field—that is, by exploring instruments through trial and error rather than being lectured before starting to play. This third interaction also represented a novel interactive paradigm for both audience members and performers that both groups appreciated. Surprisingly, the drummer even suggested expanding the audience's control rather than dismissing or resenting it. This showed us that a playful interactive paradigm that expands on existing musical forms rather than simply "adding audience contributions" to the IMPS is a more meaningful form of musicking. It also confirmed that musical interaction should be understood as a two-way process. Focused, interactive musical play is key to this real-time interaction, and we concluded that it should be central to the gestural interfacing of audience instruments that work in conjunction with stage instruments as part of IMPS.

There was also data related to research question 1 that we did not pursue. In the questionnaire, we asked about the personal backgrounds of the audience members (personality traits, musical experience and preferences, and so on). After some attempts to relate this material to the experience data via correlation analyses, it became clear that this effort would lead the research in a different direction involving the individual and psychological level. While this could have been interesting, we remained more concerned with a behavioral analysis of the interaction data, since this—together with the questionnaire data—would tell us more about how the audience instrument was actually used during the interactions. In addition, we adjusted the interaction modes and included an interaction with the stage lighting because the first concert study indicated that audience members appreciated their environmental role during the participatory parts.

## Research question 2

The second research question expanded the investigation into the individual participatory level and the structure of the audience's behavior. Following the first experiment, as well, we adjusted and refined the interaction paradigms (or modes). Alongside a new implementation of the overall system, as described in section , we added a visual modality in the form of the stage lighting. The previous third mode was also replaced by a different paradigm wherein audience members could control colors and directions while also influencing musical parameters. We introduced a refined logging tool for the protocol that allowed us to observe individual contributions in addition to the summaries of overall participation. A third line of questioning related to the design of the systems, which eventually led to the case studies and IMPS model suggested as a result of research question 3.

The question of how different interaction modes impact both collective and individual audience behavior was approached through analysis of the interaction log data, which soon validated an observation we made during the first concert study: inviting audience members to control the performative center generated the lowest overall participation in the second concert study as well. Interestingly, it was the third mode that garnered the most audience contribution this time, again presumably related to the novelty of the approach that the participants explored. The analysis at the individual level of audience members, in turn, suggested five clusters spanning from relatively reserved to highly exploratory participants. Besides the level of engagement, other qualities surfaced in the behavioral cluster: while some audience members consistently used the audience instrument and made contributions most of the time, others only used it sporadically while otherwise focusing on the stage performance itself. This "sporadic" group was in fact the largest, suggesting that participation during concerts might work best if limited in time frame rather than available throughout. One performer comment from an interview after the first concert experiment turned out to capture a major insight from the experimental results of the whole study: performers appreciate audience interaction most when it is designed as a novel improvisational paradigm and limited to a short time frame. Like the performer-audience drum solo in interaction mode 3 during the first experiment, the sporadic nature of the contribution appealed to both audience members and performers during mode 3 of the second concert study as well. This could all be due to the general unfamiliarity of the dynamic in these concerts, but despite some hesitation, all participants seemed to find some pleasure and reward in the improvisatory meeting points.

These findings can inspire further designs of audience instruments that afford interactions between stage performers and audience members as improvisational partners. Such designs might encompass many individual contributions that move the system toward certain tendencies or control aspects outside the traditional roles of stage performers (such as the direction of the lights). Ideally, these contributions to the stage performance surprise and inspire the performers, who respond to them in turn and sustain a performance-audience loop that develops into a meaningful shared

understanding in the performance space. Outside the context of participatory systems, Mazzola has already noted that the "improvisationality of a performance genre is generally proportional to the degree of audience involvement" (Sawyer, 2003, 131). This is true for audience instrument design as well: the more open performers are to improvisation, the more they welcome participatory paradigms into their work. This line of inquiry also leads us back to audience reactions that do not directly impact the music in the performance, which we looked at in the first preliminary study 7.1. While this study was limited to audio feedback alone, the audience reactions clearly came across as gestures without participatory intention, and they too should inform the design of audience instruments as ecological phenomena. While some scholars have argued for "soft" audience participation (audience influencers, see (Freeman, 2010, 155)), we advocate for paradigms that create highly improvisational scenarios in a short amount of time (very much like a solo part) for direct musical interaction. This is not to dismiss background or atmospheric components but rather to best exploit what the audience instrument affords in a performance setting.

The perceived intensity of the participatory moment also drives participation itself: adjusting a volume slider for a background sound certainly feels different than adjusting the sound of a drummer who is soloing. Both concert studies demonstrated that it is not the number of options (such as interactive buttons) in the design of the audience instrument that determines the amount of participation. (The post-show questionnaire and log data also indicated that more participation did not lead to more enjoyment of the interactive paradigms.) Instead it appears to be the specific improvisational affordances of the instrument that determine the engagement, enjoyment, and felt social connections among the participants and their inclination to listen closely and contribute mindfully during the performance. Again, playful interactive paradigms resonate most with both groups and generate a feeling of connection, fun, and surprise. That said, traditional stage performer roles might also be expanded by technological systems that are able to translate them from one group to another, manage their respective contributions, and ultimately create other common ground for co-creative experiences in the performance space.

### Research question 3

Research question 3 addresses Intelligent Music Performance Systems, of which audience instruments are a part, and explores their past, current, and future developments. This area of focus includes the main design aspects and tendencies of IMPS as well as speculation about future iterations with the help of the model suggested in [article 3](#). As stated earlier, this line of questioning was motivated by our search for participatory paradigms for the implementation of audience instruments, and IMPS arose as an overarching system that could also encompass performer instruments and autonomous instruments. With this idea in mind, we approached the question via case studies of existing systems and models together with an analysis of their terminology, including a discussion of the concepts of artificial intelligence, artificial creativity, and related ideas.

The analysis showed that there are many aspects to the design of systems (see also Hoedl et al.'s 16 design dimensions (Hödl et al., 2017) or Mazzanti et al.'s six-dimensional evaluation model (Mazzanti et al., 2014), among others). These aspects became the basis for the investigation fields of the two participatory concert experiments and were integrated into five investigation categories: *Enjoyment, Arousal, HCI, Effectiveness and Motivation*, and *Surprise and Connectedness*. Further analyses of systems involving artificial agents revealed similar approaches to describing systems with a number of attributes, design dimensions, or evaluation categories, such as Tatar et Pasquier's taxonomy of musical metacreation, which describes musical agents using technical aspects such as their architecture, environment, in- and outputs, and musical tasks, among others (Tatar and Pasquier, 2019). Inspired by the theoretical perspectives of musical embodiment, proxemics, and system autonomy, the IMPS model attempted to reduce the many aspects of system design to three main interactive aspects. In particular, IMPS bring in participatory aspects and limit them to performance. The model is further restricted to interactive aspects, both human-to-human and HCI, and it is not exhaustive of the many aspects of the design of such systems mentioned above. Instead, it attempts to uncover the main relational tendencies by intertwining aspects of closeness (embodiment and proxemics) and autonomy. Rather than offering an exhaustive description or taxonomy of IMPS, the model focuses on co-creation among performers, audiences, and artificial agents, including their inter-dependencies and perceptive qualities. It acknowledges that the analyzed cases of IMPS reside in the corners of the model's three-dimensional space and therefore speculates about instances placed among those tendencies that can be related to the main categories of instruments, systems, and agents. Drawing upon Speculative Design Research, the study concludes that innovative approaches can be applied at the intersection of those categories by system designers taking into account those relational aspects. Lastly, the model does not see performers and audiences as separate categories but instead as interactional tendencies within a co-creative process. In doing so, it acknowledges the novel interactive paradigms that are surfacing alongside interactive AI systems—systems that allow participants to navigate databases with a large amount of accumulated cultural knowledge during the performance. All of this calls for novel ways of creating IMPS, which leads to the fourth and last area of investigation addressed below.

#### Research question 4

In the thesis's review of the underlying principles of intelligent systems, those dedicated to live performance were found to rely on evolutionary algorithms, and on genetic algorithms (or genetic coding) in particular. Since genetic algorithms are based on the selection principles described above, they were revisited with co-creative performance in mind. Many music systems applied evolutionary coding principles such as swarm algorithms or biological entities such as algae or slime mold to musical performances, while systems based on symbiotic principles—that is, those starting

with different species—were less common. This inspired a more general investigation of what symbio(gene)sis might bring to the design of IMPS.

An outcome of this particular line of questioning was that the engagement of audiences (measured as quantity of participation) could be used as a representational entity in system design. Further, it was found that the genetic algorithms themselves could be altered through the implementation of symbiotic operators together with existing mutation and fitness functions. Lastly, environmental theories were used as metaphors for implementing those relationships connected to the performance space, and symbiotic principles were integrated with entanglement, as suggested by (Frauenberger, 2019), producing the four areas of co-creative collaboration shown in figure 9.5. This gave rise to the notion of symbiotic intelligence as a preliminary finding that merits implementation in future design work.

### 9.3 Overarching Contributions of the Thesis

This thesis project has investigated how interactive performance technology both amplifies and reduces—or, in effect, alters—the concert experience. The main motivation for the project was to better understand audience behavior at live performances in tandem with the instrumental realms that intelligent music technology offers to the design of performance systems. This section summarizes the contributions of the work accomplished throughout the research period and its implications for the field. It has three sections: contributions from the empirical studies, methodological contributions, and theoretical contributions.

#### Empirical Contributions for the Design and Use of Audience Instruments

Following upon the creation of two audience instruments with three participatory modes each (one of them openly accessible), two concert experiments specifically responded to the thesis’s research questions. Through them, we found that, in the context of electronic music performances, audience members preferred interactive paradigms that either expanded their traditional role in the atmospheric background of the concert space or afford playful and novel interactive paradigms linking the auditorium floor and the stage. On the other hand, paradigms that invited audience members to collectively control central performative elements, such as main melodies, were less attractive, given the bandwidth bottleneck of a single audio channel as well as the general cultural preference to leave such things in the hands of the stage performers. Those performers, in turn, reported that the elements added and adjusted by the audience were at once irritating and inspiring at first. Following a period of adjustment, however, the performers were able to interact with those contributions. It soon became clear that we needed to restrict the interactive parts to limited periods of time that were announced in advance. Given those premises, both audience members and performers opted for short, improvisational paradigms

along the lines of a traditional instrumental solo. Useful to this dynamic was the performers' direct reaction to the audience's contribution on their instruments.

In addition, the concert studies contributed insight into listening behaviors at participatory concerts. While the audience gestures on their instrument needed to be transparent, they were not always audible, yet all participants paid a lot of attention to the collective performance. Framing the concert as co-creative, that is, led to a refined form of listening during the performance and shifted participants away from the notion of a pre-prepared show and toward the notion of a collective exploration using afforded instruments in an improvisational space that conflates the roles of performers and listeners. These insights have great potential for the creation of participatory works and the design of interactive performance systems.

## Methodological Contributions for the Analysis of Audience Gestures and Instruments

The thesis contributed novel analytical methods for participatory scenarios. Results of cluster analyses of audience-reaction audio data gave rise to the possibility of framing those collective responses as *audience gestures*—a concept developed further with the help of 4E cognition. Based on theoretical inputs such as emergence and attuning, audience gestures as communicative forms of expression informed the nature of the interaction among the stage performers, other audience members, and non-human agents via audience instruments as a part of IMPS. Effects of emergence arose in the performer-audience interaction as a result of a comparative analysis of performer motion data and log data from the audience interface. On the one hand, we saw that participatory paradigms led to refined listening, as suggested above; on the other hand, we saw that emergence effects led to greater participation, which influenced the performers in turn via an interactive loop that can accelerate toward emergent and "more intense" moments in the performance. Further analysis of the log data sets at the level of individual audience members uncovered the five participatory types in a normal distributed "curve of engagement." Other qualities came up in the post-concert interviews as well—for example, audience members preferred to use the interface sporadically for shorter periods of time rather than consistently throughout the performance. This, again, calls for "pointed" interaction modes that offer improvisational opportunities within shorter time frames to audience members who are motivated to exploit them. The methodological contributions here included the establishment of the Quantity of Participation (QoP), as well as the clustering of the user groups based on the interaction data.

## Theoretical Contributions for Performative Co-Creation

As discussed above, the concepts of audience gesture and audience instruments are positioned within the more inclusive model of Intelligent Music Performance Systems (IMPS). As a thesis contribution, this model enables fresh approaches to existing performative systems and theoretical models through the lenses of embodiment,

participation, and autonomy. It relates Hall's four proxemic zones to the concepts of embodiment and participation and integrates them with notions of autonomous technologies. This offers future IMPS researchers and designers a new perspective upon performance technologies and performance system design.

In addition, the model enabled the thesis's critique of the algorithm class of genetic algorithms as a coding convention and suggested alternatives including symbio(gene)sis, a biologically inspired coding metaphor to be integrated into existing genetic algorithms or harnessed as an alternative approach for future implementations. The thesis then revisited the concept of autonomous agents via entanglement theories and the post-phenomenological repositioning of the human performer away from the center of things and toward the co-creative web of participation that IMPS enable. Lastly, the concept of entanglement was used to replace the autonomy levels and integrated with symbiotic principles to produce four quadrants to help describe and diagram the relations among those three musical agents within the performance space.

## 9.4 Limitations

Before suggesting further work for future studies, I will review the limitations of the empirical and theoretical work accomplished in the thesis project.

### Limitations of the Empirical Work

The participatory concert studies were limited in a number of ways. First, both concerts had relatively low numbers of participants. Furthermore, the studies are limited in their ecological validity because both concerts were announced as research concerts, and the participants were aware that their contributions were part of an experiment to which they consented. These conditions might have hindered natural behavior, and the participatory setting itself was new to most of the participants. The results of the studies are also limited by certain technical issues that arose during the concerts, as described in the articles. Lastly, the comparison between the experiments is limited because they involved different groups of audience members, different venues, and different compositions (though the participatory aspects were similar). In addition, the audience's interaction with the stage lighting during the second concert also complicates a direct comparison between the experiments. On a more general level, the studies are limited in the sense that they were both conducted in a specific musical genre with a specific cultural context (that is, chiptune music).

### Limitations of the Theoretical Models

The model for IMPS suggested in [article 3](#) is also limited across multiple dimensions. First, the case studies to test and validate the model are relatively few and do not

represent or reflect the many systems that exist. These cases instead derive from the author's sense of historically and culturally diverse systems that are apparent in the scientific literature and therefore well documented. A broader and more inclusive review of existing work would involve field studies and the integration of projects that are not as thoroughly connected to the academic world. In addition, the labeling of the systems in relation to the numeric system created for the model is the product of the author's perspective alone. This reading is therefore a strictly hermeneutic analysis, and the labeling would have to be done by more than one person to achieve a better evaluation of the model and its tendencies. Finally, the suggestions in [article 4](#) are merely first steps or initial ideas about approaching genetic coding, and their actual implementation is not considered yet.

## 9.5 Further Work

Since this PhD project was limited by a timeframe of three years, certain aspects remain to be investigated, and additional studies and work are required to validate, develop, and extend the findings to date. The empirical work was also constrained by the timeframe, and future projects could address the interfaces and technology in general. There are also further questions that arose, as we will see below.

### Audience participation

**Audience interfaces** Following the first experiment, a number of suggestions for further work arose in the first article. Some were addressed in the second experiment, including audience interactions with the visual domain of performance in the form of the stage lighting. Others, such as a study with bigger audiences and movement analysis, are left to future experiments. Movement data from the audience or other non-conscious inputs would improve the ecological validity of future studies in that the participatory composition would not directly interfere with the audience's behavior so they could participate without having to focus on the participation interface. Accelerators in their phones or infrared camera systems would both gather movement data in this way, as would bracelets or haptic interfaces. Some explorations have been started using finger rings that include temperature display, for example. Though they take attention away from the performance, dedicated interaction interfaces may be better than personal phones because the latter are often used for other purposes, such as communicating or sharing content via social media. Small, tactile, more ergonomic devices might encourage audience members to interact in a more playful and haptic way and merit further exploration.

**Emotional response as input data** One outcome of the second concert experiment was the conclusion that interaction with musical parameters could limit the extent of the audience's participation. In such scenarios, musically trained individuals have an advantage over amateurs, and the participatory performance might not be as accessible to everyone as intended. To address this, it could be interesting to

use audience feedback that is metalevel in character, such as emotional responses to which every audience member could contribute. Figure 9.6 imagines an interface for further study at the top left of the graph depicting nine sliders for emotional states that audience members could control. It is based on the GEMS model for musical emotions, as suggested by (Zentner et al., 2008).

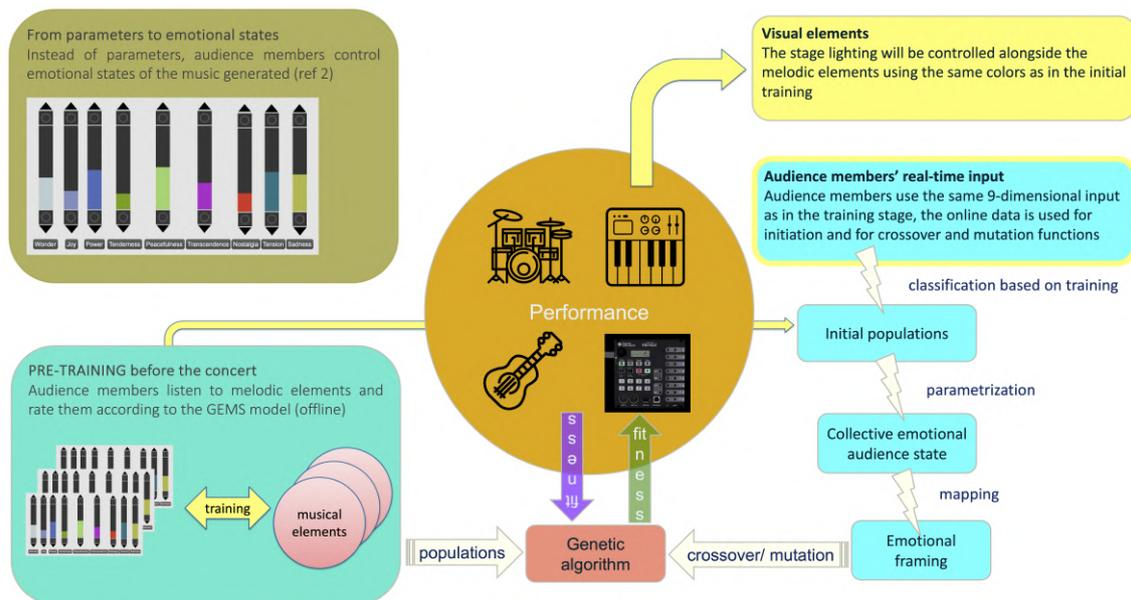


Figure 9.6: Concept for further study with emotion-based audience interface

**Non-real-time approaches** Since real-time approaches can be challenging, especially when it comes to implementing neural networks, further work could include the training of models before the concert with the help of audience members according to their specific feedback. Figure 9.6, bottom left, sketches such a scenario, where audience members train a model on their emotional responses that is then applied in combination with a genetic algorithm to the real-time context in tandem with the audience interface described in the last paragraph. In this way, the audience's agency could be "time-stretched" over a longer period. This might also open up possibilities for composers and performers to interact with this data before the actual show.

**Displaying audience contributions** Displaying, and assuring transparency of, the audience's contributions can be a challenge in participatory works. The concert experiments not only showed that interacting with the musical background is a natural way to display audience contributions in the audio domain but also that interacting with stage lighting is a natural way to extend the display of those contributions. Further work in this area might involve the exploration of 3D audio systems or concert stage designs beyond lighting in terms of engaging audiences further.

**Individual versus collective participation modes** Returning to the issue of transparency in distributed music systems, further work should also involve experimentation with the use of different individual and collective modes of inputting

content into the performance (this also involves the display and sharing of those creative contributions). For instance, in an individual mode of some sort, users might be able to explore and rehearse their contributions or store them for a later point in time. See figure 9.6 for one proposal for limiting individual interactions to the period before the event and then generating creative consensus on site; other forms of shifting between individual and collective contribution modes would be worth exploring as well.

**Audience simulation** Finally, further studies should explore the use of the anonymized data generated by audience members at a concert to train models of audience reaction that might be useful to composers and performers after the event. This might not only help them better understand the reactions of their live audience but also inspire them during the composition process or in future concerts. It is important, of course, to clear up any ethical concerns regarding the use of audience members' data long after the concert in question. After all, the live interaction at a music concert is not easy to simulate and is considered to be highly complex and non-deterministic. Still, future work should address questions like the following: "Does it help to model audience gestures (and the inherent contradictions among them) in the interests of creative participation?"

## Data Collection and Analysis

While the studies' respective audience interfaces were used for music creation, they were also analytical instruments with which to study the inclusion of audiences in staged performances. Thus, further work also remains with respect to analyzing this data in various ways and expanding on the methods used throughout this thesis.

**Correlations in user experience and behavior data** Further investigation should involve the anonymized information about the backgrounds of the audience members and any potential correlation with the experiences reported in the post-concert self-report. An initial cross-correlation matrix showed certain significant valid correlations, but more work involving causality and correlation analysis would be useful. A similar analysis of those reported backgrounds and the corresponding user clusters might even indicate what demographic, musical, and psychological backgrounds correlate with what behavioral clusters.

**Continuous measures** Data collection for further experiments should also include measuring audience feedback that is recorded in parallel to the quantitative data in order to get a better picture of what is actually happening at any given time. As the questionnaire is presently structured, it only allows audience members to report on all interaction modes at the same time, and this should be changed in future experiments. It would also be possible to allow audience members to report during breaks in the musical performance so that their experiences are as fresh as possible. Lastly, one could construct more controlled experiments with A and B

groups with and without interfaces or use the audience interface for the reporting of reflections about music alone (thus using the participatory interface for recording audience feedback alone).

**Cross-analysis with musical content** Finally, it would be interesting to look into musical and aesthetic dimensions of audience participation and use musical descriptors in relation to the participatory quantitative and qualitative results from the perspectives of both performers and audience members. This could be in the form of recorded audio and/or video from the concert, stems from the pre-produced musical elements from the performer's instruments, or MIDI data from the performer and audience interfaces. Further, the changes the audience made to the musical parameters could also provide further insight into what the participants changed creatively in a given scenario. Together with a qualitative investigation of the audience's musical intentions, fresh insight could be gained from these experiments.

## Symbio(gene)sis

**Integrated and circular systems for audiences** The outcome of the model presented in article 3 suggests that the role of instruments is moving toward systems that are created to be played more collaboratively. Such holistically designed systems would be albums that audience members can listen to and even train to reflect their interests. While this might sound idealistic, given the constraints the music industry imposes upon creative fields, the connections between album listening and concert attendance in light of more participatory systems raise interesting questions: How can Intelligent Music Performance Systems connect to the everyday listening experience of music fans and audiences? What are the novel forms of music dissemination possible via systems that integrate personal and social music listening, and how could those systems be designed? What does this mean for the design and creation of those systems themselves? How does this shape the balance between surprise and active participation by audiences?

**Implement a system that is inspired by symbio(gene)sis** Further work should engage with the implementation of a system rooted in a genetic algorithm while integrating ideas from biological theory involving symbiotic life. In addition, algae could be well suited to the integration of symbiogenesis into musical processes, since these entities propagate within just a few hours and the development of the species could be quite carefully observed. This project would entail a much deeper investigation of theoretical biology and should start with the study of symbiotic systems that could be further developed to become part of music.

**Aesthetics of IMPS** Finally, future research should explore the aesthetic artifacts of musical co-creation that involve symbiotic intelligence (or artificial intelligence broadly speaking). What are the auditory characteristics of IMPS and is there a way to identify and classify artifacts of AI with certain techniques that were used

while implementing those systems? Is sound even the domain those artifacts would show up in? If not, what other forms of expression can different ways of implementation take? As Edgar Varese famously noted, new music requires new instruments. What does this mean for a collaborative form of music? If we as humans are not the center of intelligence but instead simply part of a distributed, symbiotic intelligence involving our environment and other agents around us, how might we aesthetically implement those metaphors? In what way are symbiotic algorithms creatively specific as a subclass of genetic algorithms?

Audience participation can make creative events more accessible by allowing audience members to contribute personally. Along the way, participatory compositions can heighten the listening experience because both performers and audiences must focus harder on their contributions. Audience participation is also an idealistic narrative for co-creation - one that enables amateurs to better understand performed musical works, actively learn about their composition and structure, and contribute via their data to the training of intelligent systems. In short, this transforms the interaction into a manifestation of symbiotic intelligence distributed across performing, contributing, and artificial agents. Future work should look into how this can be done, what systems are best suited for achieving it, and how we can use participatory systems to not only enhance access to systems themselves but also visualize the hidden knowledge of the audience during live events.

# Bibliography

- Achternbusch, H. (1978). *Die Atlantikschwimmer (final movie scene)*. Herbert Achternbusch Filmproduktion (Buchendorf).
- Agre, P. (1997). *Computation and Human Experience*. Cambridge University Press.
- Albert, J. V. (2013). *Interactive Musical Partner: A System for Human/Computer Duo Improvisations*. PhD thesis, Louisiana State University and Agricultural & Mechanical College.
- Amabile, T. M. (1983). The Social Psychology of Creativity: A Componential Conceptualization. *Journal of Personality and Social Psychology*, 45(2):357–376.
- Ariza, C. (2009). The Interrogator as Critic: The Turing Test and the Evaluation of Generative Music Systems. *Computer Music Journal*, 33(2):48–70.
- Attali, J. (1985). *Noise: The Political Economy of Music*. Manchester University Press.
- Auslander, P. (2021). *In Concert: Performing Musical Persona*. University of Michigan Press.
- Barad, K. (2003). Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter. *Signs: Journal of Women in Culture and Society*, 28(3):801–831.
- Baumer, E. P. and Brubaker, J. R. (2017). Post-userism. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 6291–6303.
- Beer, R. D. (1995). A Dynamical Systems Perspective on Agent-Environment Interaction. *Artificial Intelligence*, 72(1-2):173–215.
- Bertelsen, O. W. and Bødker, S. (2003). Activity Theory. In JM, C., editor, *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science*, pages 291–324. Elsevier.
- Bickmore, T. W. and Picard, R. W. (2005). Establishing and Maintaining Long-Term Human-Computer Relationships. *ACM Transactions on Computer-Human Interaction*, 12(2):293–327.

- Bijker, W. E., Hughes, T. P., and Pinch, T. (1987). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. MIT Press.
- Biles, J. A. (2007). Improvising with Genetic Algorithms: GenJam. In Miranda, E. R. and Biles, J. A., editors, *Evolutionary Computer Music*, pages 137–169. Springer.
- Biles, J. A. and Eign, W. G. (1995). GenJam Populi: Training an IGA via Audience-Mediated Performance. In *Proceedings of the 1995 International Computer Music Conference*. ICMA.
- Birnbaum, D., Fiebrink, R., Malloch, J., and Wanderley, M. M. (2005). Towards a Dimension Space for Musical Devices. *Proceedings of the International Conference on New Interfaces for Musical Expression*, 5:192–194.
- Blackwell, A. F., Cocker, E., Cox, G., McLean, A., and Magnusson, T. (2022). *Live Coding: A User’s Manual*. MIT Press.
- Blackwell, T. and Bentley, P. (2002). Improvised Music with Swarms. *Proceedings of the 2002 Congress on Evolutionary Computation. CEC’02 (Cat. No.02TH8600)*, 2:1462–1467.
- Blackwell, T., Bown, O., and Young, M. (2012). Live Algorithms: Towards Autonomous Computer Improvisers. In *Computers and Creativity*, pages 147–174. Springer.
- Boden, M. A. (2004). *The Creative Mind: Myths and Mechanisms*. Psychology Press.
- Boden, M. A. (2010). The Turing Test and Artistic Creativity. *Kybernetes*, 39(3):409–413.
- Bødker, S. (2006). When Second-Wave HCI Meets Third-Wave Challenges. *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles*, pages 1–8.
- Bødker, S. (2015). Third-Wave HCI, 10 Years Later: Participation and Sharing. *Interactions*, 22(5):24–31.
- Bongers, B. (2007). Electronic Musical Instruments: Experiences of a New Luthier. *Leonardo Music Journal*, 17:9–16.
- Bourriaud, N., Pleasance, S., Woods, F., and Copeland, M. (2002). *Relational Aesthetics*. Les Presses du Réel Dijon.
- Bown, O. (2015). Attributing Creative Agency: Are We Doing It Right? *Proceedings of the International Conference on Computational Creativity*, pages 17–22.

- Bown, O. (2021). *Beyond the Creative Species: Making Machines That Make Art and Music*. MIT Press.
- Bown, O. and Brown, A. R. (2018). Interaction Design for Metacreative Systems. In Filimowicz, M. and Tzankova, V., editors, *New Directions in Third Wave Human-Computer Interaction: Volume 1—Technologies*, pages 67–87. Springer.
- Bown, O., Eldridge, A., and McCormack, J. (2009). Understanding Interaction in Contemporary Digital Music: From Instruments to Behavioural Objects. *Organised Sound*, 14(2):188.
- Brenneis, D. (1990). Musical Imaginations: Comparative Perspectives on Musical Creativity. *Theories of Creativity*, pages 170–189.
- Brown, A. R. and Sorensen, A. (2009). Interacting with Generative Music through Live Coding. *Contemporary Music Review*, 28(1):17–29.
- Brown, R. T. (1989). Creativity: What are we to measure? In Glover, J. A., Ronning, R. R., and Reynolds, C. R., editors, *Handbook of Creativity*, chapter 1, pages 3–32. Springer Science & Business Media.
- Bulley, J. (2018). *Sounding Materiality: Explorations in Resonant Practice*. PhD thesis, Goldsmiths, University of London.
- Bulley, J. and Jones, D. (2014). Living Symphonies Installation. Goldsmiths, University of London.
- Burnett, M. and Gallagher, S. (2020). 4E Cognition and the Spectrum of Aesthetic Experience. *The Journal for the Philosophy of Language, Mind and the Arts*, (2).
- Burrows, D. (1987). Instrumentalities. *Journal of Musicology*, 5(1):117–125.
- Caillon, A. and Esling, P. (2021). RAVE: A Variational Autoencoder for Fast and High-quality Neural Audio Synthesis. *arXiv: 2111.05011*.
- Campbell, M., Greated, C. A., and Myers, A. (2004). *Musical Instruments: History, Technology, and Performance of Instruments of Western Music*. Oxford University Press.
- Carr, C. (2023). *Dadabots*. Keynote Talk at the AI and Musical Creativity Conference at the University of Sussex, Brighton.
- Castells, M. (2010). *The Rise of the Network Society*. Wiley-Blackwell.
- Cerf, W. (1961). *An Introduction to Metaphysics*. JSTOR.
- Clark, A. (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford University Press.

- Clark, A. and Chalmers, D. (1998). The Extended Mind. *Analysis*, 58(1):7–19.
- Clarke, E. F. (2005). *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. Oxford University Press.
- Clayton, M. (2013). Entrainment, Ethnography and Musical Interaction. In Clayton, M., Dueck, B., and Leante, L., editors, *Experience and Meaning in Music Performance*, pages 17–39. Oxford University Press.
- Clément, A., Moreira, L., Rosa, M., and Bernardes, G. (2021). Musical Control Gestures in Mobile Handheld Devices: Design Guidelines Informed by Daily User Experience. *Multimodal Technologies and Interaction*, 5(7).
- Cole, R. (2020). The Problem with AI Music: Song and Cyborg Creativity in the Digital Age. *Popular Music*, 39(2):332–338.
- Collins, N. (2011). Live Coding of Consequence. *Leonardo*, 44(3):207–211.
- Collins, N. and McLean, A. (2014). Algorave: Live Performance of Algorithmic Electronic Dance Music. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 355–358.
- Collins, N., McLean, A., Rohrhuber, J., and Ward, A. (2003). Live Coding in Laptop Performance. *Organised Sound*, 8(3):321–330.
- Cook, J. (1994). Agent Reflection in an Intelligent Learning Environment Architecture for Composition. In M. Smith, A. S. and Wiggins, G. A., editors, *Music Education: An Artificial Intelligence Approach*, pages 3–23. London: Springer Verlag.
- Cook, P. (2001). Principles for Designing Computer Music Controllers. *Proceedings of CHI 2001, Extended Abstracts*, pages 491–492.
- Cope, D. (1991). *Computers and Musical Style*, volume 6. Citeseer.
- Cox, A. (2011). Embodying Music: Principles of the Mimetic Hypothesis. *Music Theory Online*, 17(2).
- Csikszentmihalyi, M. (1999). 16 Implications of a Systems Perspective for the Study of Creativity. *Handbook of Creativity*, 313.
- Csikszentmihalyi, M. (2014). *The Systems Model of Creativity: The Collected Works of Mihaly Csikszentmihalyi*. Springer.
- Culkin, J. M. (1967). A Schoolman’s Guide to Marshall McLuhan. *The Saturday Review*, pages 53, 70.
- Dahlstedt, P. (2021). Musicking with Algorithms: Thoughts on Artificial Intelligence, Creativity, and Agency. In *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*, pages 873–914. Springer.

- Dahlstedt, P. and Nordahl, M. G. (2001). Living Melodies: Coevolution of Sonic Communication. *Leonardo*, 34(3):243–248.
- Davis, T. (2019). Instrumental Intentionality: An Exploration of Mediated Intentionality in Musical Improvisation. *International Journal of Performance Arts and Digital Media*, 15(1):70–83.
- De Preester, H. (2011). Technology and the Body: The (Im) Possibilities of Re-embodiment. *Foundations of Science*, 16:119–137.
- Deleuze, G. and Guattari, F. (1988). *A Thousand Plateaus: Capitalism and Schizophrenia*. Bloomsbury Publishing.
- Di Pellegrino, G. and Làdavas, E. (2015). Peripersonal space in the brain. *Neuropsychologia*, 66:126–133.
- Doğantan-Dack, M. (2022). *Rethinking the Musical Instrument*. Cambridge Scholars Publishing.
- Dunne, A. and Raby, F. (2013). *Speculative Everything: Design, Fiction, and Social Dreaming*. MIT Press.
- Eerola, T. and Vuoskoski, J. K. (2012). A Review of Music and Emotion Studies: Approaches, Emotion Models, and Stimuli. *Music Perception: An Interdisciplinary Journal*, 30(3):307–340.
- Eglash, R. (2006). Technology as Material Culture. In Tilley, C., Keane, W., Küchler, S., Rowlands, M., and Spyer, P., editors, *Handbook of Material Culture*, pages 329–340. Sage.
- Eigenfeldt, A., Bown, O., Pasquier, P., and Martin, A. (2013). Towards a taxonomy of musical metacreation: Reflections on the first musical metacreation weekend. In *Proceedings of the Ninth Conference on Artificial Intelligence and Interactive Digital Entertainment*.
- Elsden, C., Chatting, D., Durrant, A. C., Garbett, A., Nissen, B., Vines, J., and Kirk, D. S. (2017). On Speculative Enactments. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 5386–5399.
- Euler, L. (1741). Solutio Problematis ad Geometriam Situs Pertinentis. *Commentarii Academiae Scientiarum Petropolitanae*, pages 128–140.
- Fiebrink, R. (2017). Machine Learning as Meta-instrument: Human-Machine Partnerships Shaping Expressive Instrumental Creation. In Bovermann, T., de Campo, A., Egermann, H., Hardjowirogo, S.-I., and Weinzierl, S., editors, *Musical Instruments in the 21st Century*, pages 137–151. Springer.
- Fiebrink, R. and Caramiaux, B. (2018). The Machine Learning Algorithm as Creative Musical Tool. In Dean, R. T. and McLean, A., editors, *Oxford Handbook on Algorithmic Music*, pages 181–208. Oxford University Press.

- Fiebrink, R. and Cook, P. R. (2010). The Wekinator: A System for Real-Time, Interactive Machine Learning in Music. *Proceedings of the Eleventh International Society for Music Information Retrieval Conference*.
- Fiebrink, R., Trueman, D., and Cook, P. R. (2009). A Meta-Instrument for Interactive, On-the-fly Machine Learning. *Proceedings of the Conference of New Interfaces for Music Expression*.
- Franinović, K. and Serafin, S. (2013). *Sonic Interaction Design*. MIT Press.
- Frauenberger, C. (2019). Entanglement HCI the Next Wave? *ACM Transactions on Computer-Human Interaction*, 27(1):1–27.
- Freeman, J. (2010). Web-Based Collaboration, Live Musical Performance and Open-Form Scores. *International Journal of Performance Arts and Digital Media*, 6(2):149–170.
- Frith, S. (1998). *Performing Rites: On the Value of Popular Music*. Harvard University Press.
- Gee, J. P. (2005). Semiotic Social Spaces and Affinity Spaces: From the Age of Mythology to Today’s Schools. In *Beyond Communities of Practice: Language, Power and Social Context*, pages 214–232. Cambridge University Press.
- Gemmeke, J. F., Ellis, D. P., Freedman, D., Jansen, A., Lawrence, W., Moore, R. C., Plakal, M., and Ritter, M. (2017). Audio Set: An Ontology and Human-Labeled Dataset for Audio Events. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 776–780. IEEE.
- Getzels, J. W. (1987). Creativity, Intelligence, and Problem Finding: Retrospect and Prospect. *Frontiers of Creativity Research*, pages 88–102.
- Gibbs, A. (2010). Sympathy, Synchrony, and Mimetic Communication. In Gregg, M. and Seigworth, G. J., editors, *The Affect Theory Reader*, pages 186–205. Duke University Press.
- Gibson, J. J. (1977). The Theory of Affordances. In Shaw, R. and Bransford, J., editors, *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, pages 67–82. Routledge.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates, Inc.
- Gifford, T., Knotts, S., McCormack, J., Kalonaris, S., Yee-King, M., and d’Inverno, M. (2018). Computational Systems for Music Improvisation. *Digital Creativity*, 29(1):19–36.
- Gillick, J., Deng, W., Ryokai, K., and Bamman, D. (2021). Robust Laughter Detection in Noisy Environments. *Interspeech Proceedings*, pages 2481–2485.

- Glăveanu, V. P. (2013). Rewriting the Language of Creativity: The Five A's Framework. *Review of General Psychology*, 17(1):69–81.
- Godøy, R. I. and Leman, M. (2010). Why Study Musical Gestures? In Godøy, R. I. and Leman, M., editors, *Musical Gestures: Sound, Movement, and Meaning*, pages 3–11. Routledge.
- Godøy, R. I., Song, M., Nymoën, K., Haugen, M. R., and Jensenius, A. R. (2016). Exploring Sound-Motion Similarity in Musical Experience. *Journal of New Music Research*, 45(3):210–222.
- Goffman, E. F. (1974). *Frame Analysis: An Essay on the Organization of Experience*. Harvard University Press.
- Grunwald, A. (2015). Die hermeneutische Erweiterung der Technikfolgenabschätzung. *Technikfolgenabschätzung, Theorie und Praxis*, 24(2):65–69.
- Guilford, J. P. (1950). Creativity. *American Psychologist*, 5:444–454.
- Guilford, J. P. (1967). *The Nature of Human Intelligence*. McGraw-Hill Series in Psychology. McGraw-Hill, New York.
- Guzik, E. E., Byrge, C., and Gilde, C. (2023). The Originality of Machines: AI Takes the Torrance Test. *Journal of Creativity*, 33(3).
- Hall, E. T. (1966). *The Hidden Dimension*. Doubleday, New York.
- Haraway, D. (2006). A Cyborg Manifesto: Science, Technology, and Socialist-feminism in the Late 20th Century. In Stryker, S. and Whittle, S., editors, *The International Handbook of Virtual Learning Environments*, pages 117–158. Springer.
- Hauser, S., Oogjes, D., Wakkary, R., and Verbeek, P.-P. (2018). An Annotated Portfolio on Doing Postphenomenology through Research Products. *Proceedings of the Designing Interactive Systems Conference*, pages 459–471.
- Hawkins, J. (2021). *A Thousand Brains: A New Theory of Intelligence*. Hachette UK.
- Heidegger, M. et al. (1977). *The Question Concerning Technology*. Harper & Row.
- Hershey, S., Ellis, D. P., Fonseca, E., Jansen, A., Liu, C., Moore, R. C., and Plakal, M. (2021). The Benefit of Temporally-Strong Labels in Audio Event Classification. In *International Conference on Acoustics, Speech and Signal Processing*, pages 366–370. IEEE.
- Hickinbotham, S. and Stepney, S. (2016). Augmenting Live Coding with Evolved Patterns. In *Evolutionary and Biologically Inspired Music, Sound, Art and Design: 5th International Conference, EvoMUSART 2016, Porto, Portugal, March 30–April 1, 2016, Proceedings 5*, pages 31–46. Springer.

- Hödl, O., Bartmann, C., Kayali, F., Löw, C., and Purgathofer, P. (2020). Large-Scale Audience Participation in Live Music Using Smartphones. *Journal of New Music Research*, 49(2):192–207.
- Hödl, O., Fitzpatrick, G., Kayali, F., and Holland, S. (2017). Design Implications for Technology-Mediated Audience Participation in Live Music. *Proceedings of the 14th Sound and Music Computing Conference*, pages 28–34.
- Holland, S. (2013). Artificial Intelligence in Music Education: A Critical Review. In Miranda, E. R., editor, *Readings in Music and Artificial Intelligence*, pages 239–274. Routledge.
- Holland, S., Wilkie, K., Mulholland, P., and Seago, A. (2013). Music Interaction: Understanding Music and Human-Computer Interaction. In Holland, S., Wilkie, K., Mulholland, P., and Seago, A., editors, *Music and Human-Computer Interaction*, Springer Series on Cultural Computing, pages 1–28. Springer.
- Hunter, P. G., Schellenberg, E. G., and Schimmack, U. (2008). Mixed Affective Responses to Music with Conflicting Cues. *Cognition & Emotion*, 22(2):327–352.
- Iazzetta, F. (1996). Formalization of Computer Music Interaction through a Aemi-otic Approach. *Journal of New Music Research*, 25(3):212–230.
- Ihde, D. (1979). *Technics and Praxis: : A Philosophy of Technology*. Springer.
- Ihde, D. (2009). *Postphenomenology and Technoscience: The Peking University Lectures*. SUNY Press.
- Jarina, R. and Olajec, J. (2007). Discriminative Feature Selection for Applause Sounds Detection. *Eighth International Workshop on Image Analysis for Multimedia Interactive Services*, pages 13–16.
- Jenkins, H. (2006). *Convergence Culture*. New York University Press.
- Jensenius, A. R. (2022). *SOUND ACTIONS: Conceptualizing Musical Instruments*. MIT Press.
- Johnson, D. G. and Verdicchio, M. (2019). AI, Agency and Responsibility: The VW Fraud Case and Beyond. *AI & Society*, 34:639–647.
- Jordà, S. (2005). Multi-User Instruments: Models, Examples and Promises. *Proceedings of the Conference on New Interfaces for Musical Expression*, pages 23–26.
- Jordanous, A. (2016). Four PPPPerspectives on Computational Creativity in Theory and in Practice. *Connection Science*, 28(2):194–216.
- Juslin, P. N. (2013). From Everyday Emotions to Aesthetic Emotions: Towards a Unified Theory of Musical Emotions. *Physics of Life Reviews*, 10(3):235–266.

- Juslin, P. N. (2019). *Musical Emotions Explained: Unlocking the Secrets of Musical Affect*. Oxford University Press.
- Kantosalo, A., Ravikumar, P. T., Grace, K., and Takala, T. (2020). Modalities, Styles and Strategies: An Interaction Framework for Human-Computer Co-Creativity. In *Proceedings of the 11th International Conference on Computational Creativity*, pages 57–64. Association for Computational Creativity.
- Kantosalo, A. and Takala, T. (2020). Five C’s for Human-Computer Co-Creativity: An Update on Classical Creativity Perspectives. In *Proceedings of the 11th International Conference on Computational Creativity*, pages 17–24. Association for Computational Creativity.
- King, E. and Gritten, A. (2017). Dialogue and Beyond. In Rink, J., Gaunt, H., and Williamon, A., editors, *Musicians in the Making: Pathways to Creative Performance*, volume 1, pages 306–321. Oxford University Press.
- Kirke, A. and Miranda, E. R. (2021). Performance Creativity in Computer Systems for Expressive Performance of Music. In Miranda, E. R., editor, *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*, pages 521–584. Springer.
- Kirshenblatt-Gimblett, B. (2004). Performance Studies. In Bial, H., editor, *The Performance Studies Reader*, pages 43–55. Routledge.
- Kvifte, T. (1989). *Instruments and the Electronic Age: Towards a Terminology for a Unified Description of a Playing Technique*. Solum Forlag Oslo.
- Kvifte, T. (2008). What Is a Musical Instrument? *Svensk Tidskrift för Musikforskning*, 90(1):45–56.
- Latour, B. (2005). *Reassembling the Social : An Introduction to Actor-Network Theory*. Clarendon Lectures in Management Studies. Oxford University Press.
- Leavy, P. (2017). *Handbook of Arts-Based Research*. Guilford Publications.
- Lee, S. W., de Carvalho Jr, A. D., and Essl, G. (2016). Crowd in C [loud]: Audience Participation Music with Online Dating Metaphor Using Cloud Service. *Proceedings of the 2nd Web Audio Conference*. Accessed March 15, 2023.
- Lee, S. W. and Freeman, J. (2013). Echobo: A Mobile Music Instrument Designed for Audience to Play. *Proceedings of the 13th International Conference on New Interfaces for Musical Expression*, pages 450–455.
- Leman, M. (2008). *Embodied Music Cognition and Mediation Technology*. MIT Press.
- Lemonnier, P. (2013). *Technological Choices: Transformation in Material Cultures since the Neolithic*. Routledge.

- Levitin, D. J. (2008). *The World in Six Songs: How the Musical Brain Created Human Nature*. Penguin.
- Lopes, P. (2017). Interacting with Wearable Computers by Means of Functional Electrical Muscle Stimulation. *Proceedings of the First Biannual Neuroadaptive Technology Conference*, pages 118–122.
- Lopes, P., Ion, A., Mueller, W., Hoffmann, D., Jonell, P., and Baudisch, P. (2015). Proprioceptive Interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 939–948. ACM.
- Loughran, R. and O’Neill, M. (2020). Evolutionary Music: Applying Evolutionary Computation to the Art of Creating Music. *Genetic Programming and Evolvable Machines*, 21:55–85.
- Lovelock, J. E. and Margulis, L. (1974). Atmospheric homeostasis by and for the biosphere: The gaia hypothesis. *Tellus*, 26(1-2):2–10.
- Lubart, T. (2005). How Can Computers Be Partners in the Creative Process? Classification and Commentary on the Special Issue. *International Journal of Human-Computer Studies*, 63(4-5):365–369.
- MacKenzie, D. and Wajcman, J. (1985). *The Social Shaping of Technology: How the Refrigerator Got Its Hum*. Open University Press.
- Magnusson, T. (2009). *Epistemic Tools: The Phenomenology of Digital Musical Instruments*. PhD thesis, University of Sussex.
- Maibom, H. L. (2020). *Empathy*. Routledge.
- Malafouris, L. (2008). At the Potter’s Wheel: An Argument for Material Agency. In Knappett, C. and Malafouris, L., editors, *Material Agency: Towards a Non-Anthropocentric Approach*, pages 19–36. Springer.
- Manoj, C., Magesh, S., Sankaran, A. S., and Manikandan, M. S. (2011). Novel Approach for Detecting Applause in Continuous Meeting Speech. In *Proceedings of the 3rd International Conference on Electronics Computer Technology*, volume 3, pages 182–186. IEEE.
- Matuszewski, B., Schnell, N., and Bevilacqua, F. (2019). Interaction Topologies in Mobile-Based Situated Networked Music Systems. *Wireless Communications and Mobile Computing*.
- Mazzanti, D., Zappi, V., Caldwell, D. G., and Brogni, A. (2014). Augmented Stage for Participatory Performances. *Proceedings of the Conference of New Interfaces for Music Expression*, pages 29–34.
- Mazzola, G. and Cherlin, P. B. (2008). *Flow, Gesture, and Spaces in Free Jazz: Towards a Theory of Collaboration*. Computational Music Science. Springer Science & Business Media.

- Mazzola, G., Park, J., and Thalmann, F. (2011). *Musical Creativity: Strategies and Tools in Composition and Improvisation*. Computational Music Science. Springer, Berlin Heidelberg, 1. ed. edition. Literaturangaben.
- McCormack, J. (2021). Achieving Greatness: The Making of Eden (Video). <https://youtu.be/hGKEMEO5V3g>, Timecode 37:22, Accessed September, 10, 2023.
- McCormack, J., Hutchings, P., Gifford, T., Yee-King, M., Llano, M. T., and d’Inverno, M. (2020). Design Considerations for Real-Time Collaboration with Creative Artificial Intelligence. *Organised Sound*, 25(1):41–52.
- McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. MIT Press.
- Mejias, U. A. (2007). *Networked Proximity: ICTs and the Mediation of Nearness*. Teachers College, Columbia University.
- Menardi, G. (2011). Density-Based Silhouette Diagnostics for Clustering Methods. *Statistics and Computing*, 21(3):295–308.
- Minsky, M. (1961). Steps toward Artificial Intelligence. *Proceedings of the IRE*, 49(1):8–30.
- Minsky, M. (1988). *Society of Mind*. Simon and Schuster.
- Miranda, E. R. (2021). *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*. Springer Nature.
- Miranda, E. R. and Al Biles, J. (2007). *Evolutionary Computer Music*. Springer.
- Miranda, E. R. and Wanderley, M. M. (2006). *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*, volume 21 of *The Computer Music and Digital Audio Series*. AR Editions, Inc.
- Mueller, F. F., Lopes, P., Strohmeier, P., Ju, W., Seim, C., Weigel, M., Nanayakkara, S., Obrist, M., Li, Z., Delfa, J., et al. (2020). Next Steps for Human-Computer Integration. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–15.
- Mulligan, M. (2023). The Music Industry’s Tipping Point Is Right Here, Right Now. <https://rainnews.com/mark-mulligan-07152023>. Accessed December 17, 2023.
- Newen, A., De Bruin, L., and Gallagher, S. (2018). *The Oxford Handbook of 4E Cognition*. Oxford University Press.
- Nika, J. and Chemillier, M. (2012). Improtek: Integrating Harmonic Controls into Improvisation in the Filiation of OMax. *Proceedings of the International Computer Music Conference*, pages 180–187.

- Nika, J., Déguernel, K., Chemla, A., Vincent, E., and Assayag, G. (2017). DYCI2 Agents: Merging the "Free", "Reactive", and "Scenario-Based" Music Generation Paradigms. *Proceedings of the International Computer Music Conference*.
- Noë, A. (2004). *Action in Perception*. MIT Press.
- Norman, D. A. (1988). *The Psychology of Everyday Things*. Basic Books.
- Norman, D. A. (1999). Affordance, Conventions, and Design. *Interactions*, 6(3):38–43.
- Nymoen, K. (2013). *Methods and Technologies for Analysing Links between Musical Sound and Body Motion*. PhD thesis, University of Oslo.
- Oliveros, P. (2010). *Sounding the Margins: Collected Writings 1992-2009*. Deep Listening Publications.
- O'Neill, S. and Sloboda, J. (2017). Responding to Performers. In Rink, J., Gaunt, H., and Williamon, A., editors, *Musicians in the Making: Pathways to Creative Performance*, volume 1, pages 322–340. Oxford University Press.
- Oord, A. v. d., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., and Kavukcuoglu, K. (2016). Wavenet: A Generative Model for Raw Audio. *arXiv Preprint arXiv:1609.03499*.
- Orning, T. (2017). Music as Performance – Gestures, Sound and Energy: A Discussion of the Pluralism of Research Methods in Performance Studies. *Journal for Research in Arts and Sports Education*, 1(5).
- Pena, W. (1987). *Problem Seeking: An Architectural Programming Primer*. AIA Press, 3rd edition.
- Ravikumar, P. T. and Wyse, L. (2019). MASSE: A System for Music Action Selection Through State Evaluation. *Proceedings of the International Conference on Computational Creativity*, pages 132–139.
- Rebelo, P. (2009). Dramaturgy in the Network. *Contemporary Music Review*, 28(4-5):387–393.
- Renwick, R. C. (2017). *Topologies for Network Music*. PhD thesis, Queen's University Belfast.
- Reuter, A. (2021). Pop as Process: The Digitalization of Groove, Form and Time. *Dancecult: Journal of Electronic Dance Music Culture*, 13(1):2–21.
- Rhodes, M. (1961). An Analysis of Creativity. *The Phi Delta Kappan*, 42(7):305–310.
- Rowe, R. (1993). *Interactive Music Systems: Machine Listening and Composing*. MIT Press.

- Russel, S. and Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Pearson Education, 2nd edition.
- Russell, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178.
- Sacharin, V., Schlegel, K., and Scherer, K. R. (2012). *Geneva Emotion Wheel Rating Study (Report)*. University of Geneva, Swiss Center for Affective Sciences.
- Saltz, D. Z. (2003). The Art of Interaction: Interactivity, Performativity, and Computers. In Auslander, P., editor, *Performance: Critical Concepts in Literary and Cultural Studies*, pages 332–350. Routledge.
- Saunders, R. and Bown, O. (2015). Computational Social Creativity. *Artificial Life*, 21(3):366–378.
- Saunders, R. and Gero, J. S. (2002). How to Study Artificial Creativity. *Proceedings of the 4th Conference on Creativity & Cognition*, pages 80–87.
- Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., and Fitch, W. T. (2021). Music as a Coevolved System for Social Bonding. *Behavioral and Brain Sciences*, 44:e59.
- Savery, R. and Weinberg, G. (2022). Robotics. In Clancy, M., editor, *Artificial Intelligence and Music Ecosystem*, pages 52–67. CRC Press.
- Sawyer, R. K. (2003). *Group Creativity: Music, Theater, Collaboration*. Lawrence Erlbaum.
- Sawyer, R. K. (2006). Group Creativity: Musical Performance and Collaboration. *Psychology of Music*, 34(2):148–165.
- Sawyer, R. K. (2011). *Explaining Creativity: The Science of Human Innovation*. Oxford University Press.
- Sawyer, R. K. (2019). Individual and Group Creativity. In Kaufman, J. C. and Sternberg, R. J., editors, *The Cambridge Handbook of Creativity*, Cambridge Handbooks in Psychology, pages 567–586. Cambridge University Press.
- Sawyer, R. K. and DeZutter, S. (2009). Distributed Creativity: How Collective Creations Emerge from Collaboration. *Psychology of Aesthetics, Creativity, and the Arts*, 3(2):81.
- Schechner, R. (2004). Performance Studies: The Broad Spectrum Approach. In Bial, H., editor, *The Performance Studies Reader*, pages 7–9. Routledge.
- Schuller, B. W. (2013). *Intelligent Audio Analysis*. Springer.
- Seif, G. (2018). The 5 Clustering Algorithms Data Scientists Need to Know. <https://www.kdnuggets.com/2018/06/5-clustering-algorithms-data-scientists-need-know.html>. Accessed January 8, 2024.

- Sismondo, S. (2010). *An Introduction to Science and Technology Studies*, volume 1. Wiley-Blackwell.
- Sloboda, J. A. and Juslin, P. N. (2001). Psychological Perspectives on Music and Emotion. In Sloboda, J. A. and Juslin, P. N., editors, *Music and Emotion: Theory and Research*, pages 71–104. Oxford University Press.
- Sloboda, J. A. and Lehmann, A. C. (2001). Tracking Performance Correlates of Changes in Perceived Intensity of Emotion during Different Interpretations of a Chopin Piano Prelude. *Music Perception*, 19(1):87–120.
- Small, C. (1998). *Musicking: The Meanings of Performing and Listening*. Wesleyan University Press.
- Sternberg, R. J. and O’Hara, L. A. (1998). Creativity and Intelligence. In J, S. R., editor, *Handbook of Creativity*, pages 251–272. Cambridge University Press.
- Tanimoto, S. L. (2013). A Perspective on the Evolution of Live Programming. In *1st International Workshop on Live Programming*, pages 31–34. IEEE.
- Tatar, K. and Pasquier, P. (2019). Musical Agents: A Typology and State of the Art Towards Musical Metacreation. *Journal of New Music Research*, 48(1):56–105.
- Taylor, P. A. (2013). Participation and the Technological Imaginary: Interactivity or Interpassivity? In Delwiche, A. and Henderson, J. J., editors, *The Participatory Cultures Handbook*, pages 247–256. Routledge.
- Thom, B. (2000). BoB: An Interactive Improvisational Music Companion. *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 309–316.
- Thom, B. (2003). Interactive Improvisational Music Companionship: A User-Modeling Approach. *User Modeling and User-Adapted Interaction*, 13(1-2):133–177.
- Toksöz Fairbairn, K. (2022). *Dis/cord: Thinking Sound Through Agential Realism*. Punctum Books.
- Tomás, E., Gorbach, T., Tellioglu, H., and Kaltenbrunner, M. (2022). *Embodied gestures*. TU Wien Academic Press.
- Torrance, E. P. (1966). Torrance Tests of Creative Thinking. *Educational and Psychological Measurement*.
- Turchet, L. and Barthelet, M. (2017). Envisioning Smart Musical Haptic Wearables to Enhance Performers’ Creative Communication. *Proceedings of International Symposium on Computer Music Multidisciplinary Research*, pages 538–549.

- Turchet, L., West, T., and Wanderley, M. M. (2020). Touching the Audience: Musical Haptic Wearables for Augmented and Participatory Live Music Performances. *Personal and Ubiquitous Computing*, 25:49–769.
- Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236):433–460.
- Uhle, C. (2011). Applause Sound Detection. *Journal of the Audio Engineering Society*, 59(4):213–224.
- Varela, F. J., Rosch, E., and Thompson, E. (1992). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.
- Vear, C., Benford, S., Avila, J. M., and Moroz, S. (2023). *Human-AI Musicking: A Framework for Designing AI for Music Co-Creativity*. Talk at the AI and Musical Creativity Conference at the University of Sussex, Brighton.
- Verbeek, P.-P. (2005). *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. Penn State Press.
- Verbeek, P.-P. (2015). Introduction to Mediation Theory. *Interactions*, 22(3):26–31.
- Voigts, E., Auer, R. M., Elflein, D., Kunas, S., Röhnert, J., and Seelinger, C. e. (2024). *Artificial Intelligence – Intelligent Art? Human-Machine Interaction and Creative Practice*. transcript Verlag.
- Von Hornbostel, E. M. and Sachs, C. (1914). Systematik der Musikinstrumente: Ein Versuch. *Zeitschrift für Ethnologie*, 46(H. 4/5):553–590.
- Weeks, P. (1996). Synchrony lost, synchrony regained: The achievement of musical co-ordination. *Human Studies*, 19(2):199–228.
- Weinberg, G. (2005). Interconnected Musical Networks: Toward a Theoretical Framework. *Computer Music Journal*, 29(2):23–39.
- Wessel, D. (2006). An Enactive Approach to Computer Music Performance. *Le Feedback dans la Creation Musical*, pages 93–98.
- Wiggins, G. A. (2006). Searching for Computational Creativity. *New Generation Computing*, 24:209–222.
- Wikström, P. (2009). *The Music Industry: Music in the Cloud*. Polity Press.
- Yang, N., Sha, R., Sankaranarayanan, R., Sun, Q., and Weinberg, G. (2021). Drumming Arm: An Upper-Limb Prosthetic System to Restore Grip Control for a Transradial Amputee Drummer. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 10317–10323. IEEE.
- Yee-King, M. J. (2011). An Autonomous Timbre Matching Improviser. *Proceedings of the International Computer Music Conference*.

- Zentner, M., Grandjean, D., and Scherer, K. R. (2008). Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement. *Emotion*, 8(4):494–508.

# Appended Articles

## Article 1 - Hacking the Concert Experience: Exploring Co-Creative Audience Interaction at a Chiptune Live Performance

### Abstract

*With this pilot study we explore three paradigms of interaction between audience members and musical performers at a chiptune live concert. The central question is how both audience members and performers experience the concert during active audience participation via personal smartphones. The concert experiment was conducted during a live performance of the chiptune group "Kubbi". The three different interaction paradigms are investigated based on a mixed methods approach that uses smartphone interaction logs, audience survey data from a self-report questionnaire, as well as interview data from two interviews that we conducted after the concert with a group of audience members and the performing musicians separately. The work contributes with insights into the musical integration of participatory paradigms, a better understanding of the experience of the interactive performance, and reflections on the general benefits of participatory approaches in live electronic music concerts and specifically chiptune music.*

Keywords: Audience Participation, Interactive Performance, Chiptune Music

Version: accepted manuscript

## Article 2 - Hack the Show: Design and Analysis of Three Interaction Modes for Audience Participation

### Abstract

*Artists and composers have explored the potential of mobile phones to enable audience interaction in live music concerts for the last two decades. However, studies of the compositional process and the systems' use by audience members remain scarce. This paper presents the design of three interactive modes for audience participation which invite audience members to influence both musical parameters and stage lighting of an electronic live music performance. Analysis of log data from the concert reveals users' interaction preferences and suggests five different types of participatory user behaviors, which are supported by results of a post-concert questionnaire. Finally, the paper considers design aspects for audience participation and proposes ideas for future implementations of distributed music systems.*

Keywords: Audience Participation, Distributed Music Systems, Interactive Performance

Version: accepted manuscript

# Hack the Show: Design and Analysis of Three Interaction Modes for Audience Participation\*

MATTHIAS JUNG<sup>1</sup> AND IAN CLESTER<sup>2</sup>  
(matthias.jung@uia.no) (ijc@gatech.edu)

<sup>1</sup> Department of Popular Music, Faculty of Fine Arts, University of Adger, Kristiansand, Norway

<sup>2</sup> Center for Music Technology, Georgia Institute of Technology, Atlanta, GA, USA

Artists and composers have explored the potential of mobile phones to enable audience interaction in live music concerts for the last two decades. However, studies of the compositional process and the systems' use by audience members remain scarce. This paper presents the design of three interactive modes for audience participation which invite audience members to influence both musical parameters and stage lighting of an electronic live music performance. Analysis of log data from the concert reveals users' interaction preferences and suggests five different types of participatory user behaviors, which are supported by results of a post-concert questionnaire. Finally, the paper considers design aspects for audience participation and proposes ideas for future implementations of distributed music systems.

## 0 INTRODUCTION

The motivation for this study stems from interest in involving audience members as active participants in musical live performances. Interaction technologies for participatory concerts in real-time have been getting more accessible in recent years and they are expected to be a promising way to meet the interest of concert audiences in getting involved with performing artists. The main intention of this research is to determine how the design of audience interaction technology affects the participatory behavior of audience members at live music concerts. The study aims to explore the design and use of three different interaction modes, which we hope will inform the creation of future iterations of participatory systems. Specifically, we propose to address the following research questions:

- How do different interaction modes affect collective audience behavior?
- How can the audience members' contributions be structured on an individual level?
- How can the findings be applied to the design of participatory systems?

After a brief review of related work, we explain the experimental design and procedure undertaken in the study, including the musical and technical implementation and

data collection methods at the live event. Then, we present the results of the interaction log analysis and the self-report questionnaire. Finally, we discuss our results and their limitations with a focus on design considerations, before closing with suggestions for future work and concluding remarks.

## 1 RELATED WORK

### 1.1 Distributed music

The concept of distributed creativity goes far beyond musical live events and has been explored from many different perspectives. Bourdieu theorizes the field of cultural production taking into account a complex network of different art forms and audiences by embedding the artistic processes into a framework of creative agents [1, p. 49]. Glavanu echoes the idea of creative aspects in the receptive act and describes distributed creativity as "a conception that extends previous accounts of cognition and makes them engage with the sociality, materiality and temporality of the world"[2, p. 28]. For the field of music, distributed creativity has been investigated by e.g. Clarke [3] and Hagberg [4]. Taylor suggests distributed music to be a musical genre and historically reviews participatory compositions starting in the 1970s with works such as *Cassette 100* (1971) and *Ugnayan* (1974) by Filipino professor Jose Maceda (1917-2004), which involve tape players and radio receivers as a speaker array held by audience members [5].

---

\*Corresponding author email address: matthias.jung@uia.no.  
Last updated: June 28, 2023

## 1.2 Personal phones for audience participation

An early example for a participatory composition that involves audience members' phones is *Dialtones* (2001) by Levin, which made use of mobile phones to play back pre-composed audio on the phones' speakers but did not include active involvement of participants [6]. Early smartphone-centered compositions including audience participation in the early 2010s, such as *Orkestra* (2010) by Nicholas Bryan (described in [7]) or *echobo* (2012) by Lee [8], often used native mobile apps for audience participation. Since the availability of the Web Audio API<sup>1</sup> and web application frameworks such as Node.js<sup>2</sup>, many participatory works in the later 2010s use browser-based solutions for the audience to participate with their personal smartphones. A characteristic example can be found in *soundworks*, developed by Matuszewski et al. since 2015 as a framework for the creation of distributed multimedia applications [9].

## 1.3 Interactive approaches

Artists have explored many different ways that audience members can be involved in staged performances. *Reflecting Paddles* (1991) by Loren and Rachel Carpenter used a camera-system to assess colors of flippable green-and-red-paddles that are held by audience members for controlling a staged visual [10]. Dahl et. al used real-time Twitter data for sonification and visualization in their composition *TweetDreams* (2011) [11]. *Crowd in C* (2016) by Lee et al. encouraged interaction among audience members with a dating metaphor that includes short melodies associated with user profiles [12]. *Open Symphony* (2017) by Wu et al. let audience members vote among five different playing styles that are suggested to specific performers of acoustic instruments [13].

*Open Band* (2017) used a multi-user web chat to produce sounds directly from individual contributions [14], which was the basis for *Playsound* (2019), where sonification of the language prompts was executed based on the semantic features, accessing an online sample data base [15]. Hoedl et al. used ultrasonic sound for audience participation in a compositional process involving audience members [16]. Turchet explored haptic wearable technology to display music-induced tactile signals to audience members during both a DJ performance and a mandoline performance [17]. With their composition *Autopia* (2021), Lorway et al. explored aspects of Computational Creativity with an AI bot in a live-coding performance, that uses audience members' evaluations of melodies to train the fitness function of a genetic algorithm [18].

## 1.4 Typologies for distributed music works

Multiple attempts can be found to typologize distributed music systems. For more general typologies of interactive music systems the reader is referred to Weinberg [19] and

<sup>1</sup><https://github.com/WebAudio/web-audio-api>

<sup>2</sup><https://github.com/nodejs/node>

Renwick [20]. For more specific frameworks for audience participation we recommend reviewing Freeman [21], Allison [22], York [23], and Matuszewski [24]. For a better understanding of the perspective of the audience members, Hoedl et al. presented a study among musicians and potential audience members resulting in 16 key issues for the design of technology-mediated audience participation [25]. Xambó and Roma suggest a set of 13 composition dimensions for analyzing networked music works that include "the role of the performer, the role of the audience, the location of sound and the type of feedback" [26, p. 55].

## 2 EXPERIMENTAL DESIGN AND PROCEDURE

The following sections will first present the compositional process and interaction design for the participatory concert, specify the three different interactive modes, and finally describe the concert setting for the experiment.

### 2.1 Compositional process and interaction design

The interactive concert was designed in collaboration with the electronic music trio *Kubbi* and lighting technician Mike Holm. It builds on a previous study with two musicians of the trio in April 2022 and incorporates feedback from the audience as described in [27]. For the audience interface we chose a simple, slider-based tool that integrated well with the performers' and lighting technician's interfaces, but we replaced the parameter names with symbols since names proved to be a point of confusion in earlier experiments.

The concert features three interactive parts that were integrated into an existing *Kubbi* show. These parts will be referred to as Mode 1, 2, and 3, and described in detail below. In the audience interface they appear as Level 1, 2, and 3. Mode 1 lasted 334 seconds at the beginning of the concert, Mode 2 lasted 212 seconds in the middle of the concert, and Mode 3 lasted 318 seconds towards the end of the concert. The different lengths of the interactive parts are a result of the musical implementation of the parts, and as such were not set to be equally long.

The performance was rehearsed in *Kubbi's* rehearsal space together with Mike Holm, who used a screen for visualization of the stage lighting. The interactive modes were stress-tested with a tool for simulating 100 audience members sending random contributions to the server. The following section describes the outcome of this process and explains the three modes that were used for the participatory parts of the concert.

### 2.2 Three modes for audience interaction

At the beginning of the concert, audience members were presented with a mixer-like interface including eight sliders as shown in Fig. 1. The symbols in front of the sliders correspond to the specific soundscapes, and the colors correspond to colors of the stage lighting. During Mode 1, audience members adjust the volume levels of these soundscapes, and at the same time control the intensity of

the moving head light which is hard-wired to each of the soundscapes.

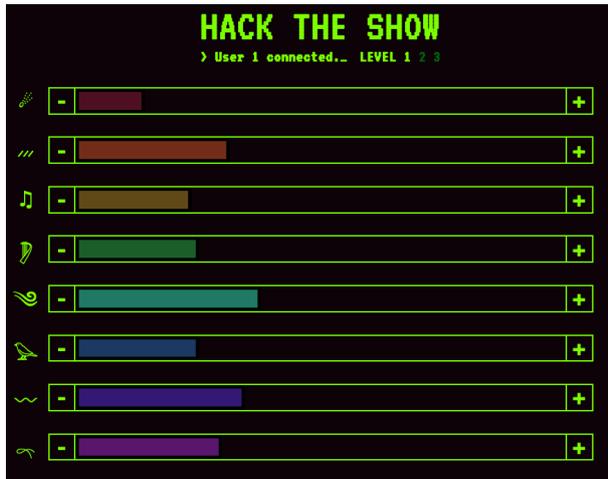


Fig. 1. Audience interface for Mode 1.

The interface for Mode 2 is shown in Fig. 2 and opens a filter module of a four-step sequencer. The filtering brings out the harmonics of the sequenced tones, so that it also affects the perceived pitch of the tones. This is combined with a cutoff filter to bring out the main melody within the electronic music performance. The four steps of the sequencer similarly affect the color tone of four moving lights distributed from left to right across the stage and pointed downwards from above the stage, so that the melody line moves from stage left to stage right in time and space simultaneously. The higher the slider value, the more red tone is mixed into the white tone of the lights for each of the slider elements of the audience interface.

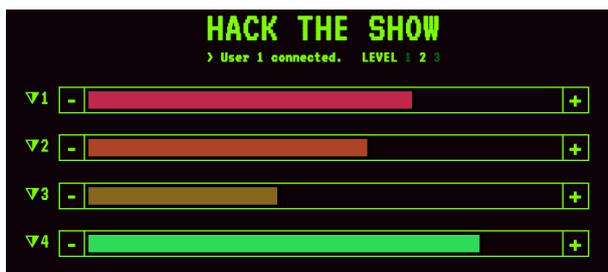


Fig. 2. Audience interface for Mode 2.

The interface for Mode 3 features an x-y plane that allows audience members to move a point across two dimensions via the counter objects on the four sides of the plane as shown in Fig. 3. The y-values determine the balance between the low-pitch and high-pitch parts of a chord-riff generated by the keyboard player. Lower values towards the bottom end of the plane feature the low-pitch elements, and higher values along the y-axis towards the top of the plane feature the higher pitched parts of the melody. Values along the x-axis determine the parts of the signal of the chord-riff was being sent to a reverb effect. Higher values along the x-axis lead to sending bigger parts of the signal

to the reverb, ending up in a wetter signal towards the right end of the plane. At the same time, the current position in the x-y plane determined the direction of the stage lights, which were constantly moving between left/right (x-axis) and bottom/top (y-axis) as an effect of the audience members' input.

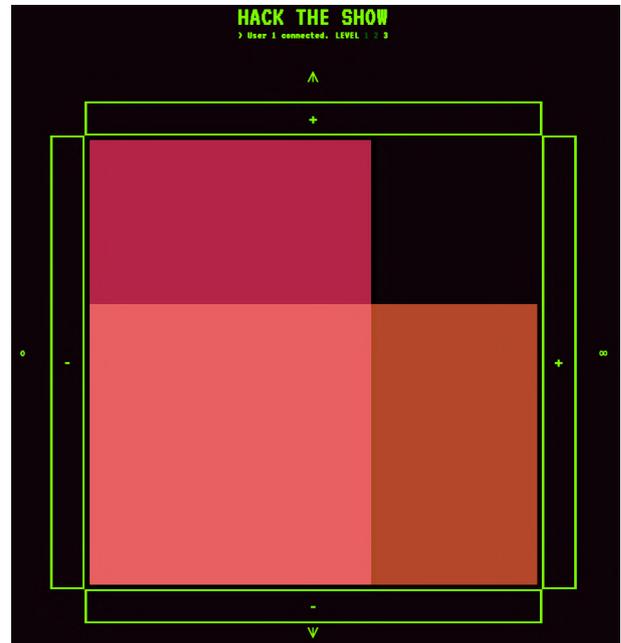


Fig. 3. Audience interface for Mode 3.

In all three modes, audience members control the sliders via the + and - buttons, which send commands to the server (described in 2.3) to increment or decrement the corresponding parameter. This design enables everyone to have some control over parameter values without allowing any individual to unilaterally set a parameter, with the result that the sliders—and their corresponding musical parameters—tend to move smoothly, reflecting the net input across all audience members, rather than rapidly jumping around between absolute values.

Apart from the slider elements, + and - buttons, and symbols, the audience interface also indicated the current interaction (Level 1, 2 or 3), the user's connection status (Connected or Connecting...), and the user number, which was taken from the IP address from the individual audience member. This user number was then later used in the post-concert survey to connect the individual interactions to the survey results in an anonymous way. Whenever no interaction was active, the interface showed the message "Nothing to see here".

### 2.3 Implementation

After running into performance limitations with our first approach,<sup>3</sup> we decided to instead build a custom web ap-

<sup>3</sup>The approach made use of Max/MSP's Miraweb package <https://github.com/Cycling74/miraweb> as described in [27].

plication to enable audience interaction.<sup>4</sup> This application took the form of a web client written in vanilla JavaScript communicating with a server written in Python (using Sanic<sup>5</sup>) over WebSockets. The client was responsible for sending user actions to the server and displaying the current state (with a UI that varied depending on the current interaction mode). The server received actions from each client, logged them for later analysis (see Sec. 2.5), updated the global state (consisting of the current parameter values), and finally sent updates to clients about state changes (computing differences between the global state and each client's known local state and only sending the delta). Note that user actions represented incremental changes (e.g. 'decrement slider 2') rather than absolute changes ('set slider 1 to 30'), corresponding to the + and - button interactions described in 2.2. This design reduces the risk of race conditions, such as one user's action overwriting another's because it was processed second.

In addition to serving clients, our server also created virtual MIDI input/output devices (using Mido<sup>6</sup> with python-rtmidi<sup>7</sup>) in order to communicate with controllers, instruments, and stage lighting via MIDI CC events. A controller connected to the server allowed the performer to switch the current interaction mode from on-stage. And the server controlled various parameters (mixer levels, filter cutoffs, and EQ/reverb routing, as described in Sec. 2.2) via MIDI CC output.

Ultimately, using standard web technologies (HTML, CSS, JS, WebSockets) enabled us to build an application more tailored to the concert and attain greater confidence about its reliability. In addition to simplifying styling and animation, we were able to construct custom widgets such as the x-y plane (essentially overlapping, translucent sliders) for Mode 3. We also built a small, browser-based stress-testing tool that allowed us to easily spin up hundreds of concurrent connections and gain confidence about the system being performance-ready.



Fig. 4. System architecture.

## 2.4 Concert setting

The live concert was part of the Norwegian Research Days and was staged in *Teateret* in Kristiansand (Norway) on September 27, 2022. The performance was part of a program of several events and followed an interview of

<sup>4</sup>Our application's source code is available at <https://github.com/ijc8/hack-the-show> as free software.

<sup>5</sup><https://sanic.dev/>

<sup>6</sup><https://mido.readthedocs.io/>

<sup>7</sup><https://spotlightkid.github.io/python-rtmidi/>

Vegard Kummen and Matthias Jung by Prof. Rolf Kristensen. There were approximately 30 audience members attending the show, which lasted about 45 minutes and was free of charge. Fig. 5 shows a picture of the stage taken from the back of the auditorium. It captures the 16 moving heads that were used as the interactive lighting display for the performance. Audiovisual documentation of the three modes can be found at <https://tinyurl.com/hacktheshow>.



Fig. 5. Frontal view of the stage.

## 2.5 Data collection

**Interaction log** All audience contributions were stored in a JSON file, including a timestamp derived from the server's clock upon receiving a command (when did an interaction happen?), user IP (who caused the interaction?), parameter number (what parameter did they change?), and value change (how much did they change it?). Parameters refer to the sliders for Modes 1 and 2, and for the x and y axes for Mode 3. Parameter changes are adapted to the MIDI-standard (0-127), and were set to + and -2 for Mode 1 and 2, and to + and -1 for Mode 3. These step sizes proved to be best suited for the size of the audience as a result of our previous studies, and also seemed a reasonable choice after testing them with the audience simulation tool for up to 100 virtual audience members during rehearsals.

**Post-concert questionnaire** Right after the concert, audience members were asked to fill out a self-report questionnaire including 56 questions about their personal backgrounds, listening habits and general concert behavior, as well as their experience of the participatory concert. This paper makes only brief use of this background information to give an idea about the demographic and then focuses mainly on the reported experience. Most of the answers were collected on 7-point Likert-type items, and as ordinal data for the respondents' preferences of the three interactive modes tested (Mode 1, 2, 3, or None of them). 7-point Likert-scale was used following [17], whose method was found to be well suited for our survey. Participants accessed the questionnaire after the concert via a QR-code on their personal phones. All audience members were informed about the study and the anonymized use of the log-file data before the concert. Participants of the survey were informed about data use at the beginning of the questionnaire. All data was collected anonymously and did not in-

clude any any information that can be traced back to individuals.

### 3 RESULTS

The results of the study will be presented in two parts. First, an analysis of all contributions, as recorded in the log data, will be presented. Methodological approaches for analyzing log files can be found e.g. at [28]. For an approach that combines log data with survey data, we refer to [29]. Then, the same log data will be analysed with respect to individual contributions. In the second part of this section, the results of the self-report questionnaire, which was answered by some of the audience members, will be presented.

#### 3.1 Interaction log

**Quantities of participation** The log file shows 18 different IP addresses, which translates to roughly two thirds of the audience members joining the interactions<sup>8</sup>. Looking into the three interaction modes, 15 audience members participated in Mode 1, eight audience members in Mode 2, and 15 audience members in Mode 3. The total number of records in the log file, which translates to the total number of all single clicks recorded is 9,791<sup>9</sup>. These can be divided into 2,944 for Mode 1, 1,940 for Mode 2, and 4,907 for Mode 3. The lowest contribution of single IP addresses over all three interaction modes is one click, the highest one 2,633. The overall time the log recorded is 2,268 seconds (37 minutes and 48 seconds), whereas the individual interaction modes lasted for 334, 212, and 318 seconds respectively. During Mode 2 no events were logged between 127 and 163 seconds into the mode due to technical difficulties on the server side. The total click numbers for Mode 2 would therefore most certainly have been higher than reported.

**Click rates** Given the total click numbers together with the lengths of the three interaction modes, the average number of clicks per second can be calculated for each mode. These are 8.82 clicks per second for Mode 1, 11.02 clicks per second for the Mode 2, and 15.44 clicks per second for Mode 3. For Mode 2, the 36 seconds when there was no logging of events were subtracted from the total time before calculating the click rate for better comparability. Taking into account the different numbers of participants, we can also consider the click rate on an individual level, which translates to an average individual click rate of 0.59 clicks per second per participant for Mode 1 (SD 0.46), 1.38 clicks per second per participant for Mode 2 (SD 1.45), and 1.03 clicks per second per participant for Mode

3 (SD 1.21). On average, individual participants clicked 1.00 times per second (SD 1.04), when interaction was possible.

**Participation during the three modes compared** Fig. 6 shows a cumulative plot for the first 200 seconds of the three different interaction modes in comparison. The number of audience members participating varied between the modes, but the plot gives an impression of the audience's total activity over time and shows the progression of their contributions. The next section considers these contributions on an individual level and makes an attempt to cluster those based on user behavior. We cannot exclude the fact that the order of the modes influenced the participatory behavior of the audience members. Moreover, the curve of Mode 2 shows a plateau between 127 and 163 seconds which is caused by the before mentioned technical issues with logging events and it would most certainly have continued and surpassed the curve of Mode 1 after 200 seconds.

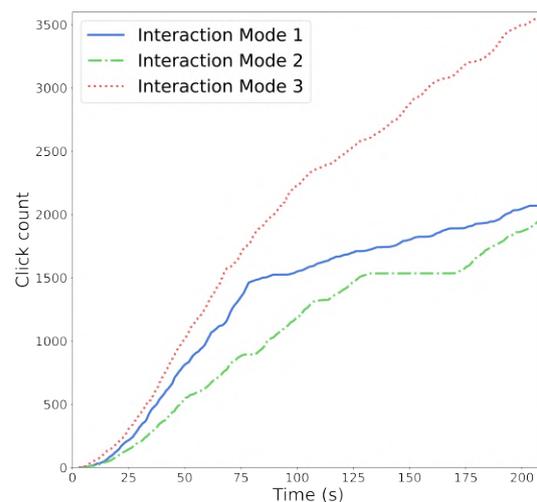


Fig. 6. Cumulative clicks for the first 200 seconds of the three interaction modes compared.

**Individual participation analysis** For analyzing the individual audience members' contributions, we employ K-means clustering [30]. This clustering method will suggest the best number of user clusters, which then will be analyzed in more detail. As feature dimensions for analysis, the log parameters *timestamp*, *parameter number*, and *value change* of the JSON log file are taken into account for each of the IP addresses. However, it is considered more beneficial to process these features in the sense that the timestamps are transferred into a total dwell time for each participant, which is common practice in user analysis [31]. We combine the record of parameter numbers and value changes into a second feature dimension called *button repeats*, which reflects the number an individual par-

<sup>8</sup>For simplicity and to avoid more invasive identification methods, we assume that audience members stayed connected to the venue's Wi-Fi and thus retained the same IP for the duration of the concert.

<sup>9</sup>The interaction events, executed as button presses on the participants' smartphones, are referred to as *clicks* throughout the text.

ticipant clicks the same button repeatedly, i.e. the + or - of the same slider. As a third feature dimension, the total number of clicks that a participant contributes is considered an important parameter for participation in the concert. The three-dimensional feature array is then scaled, before K-means clustering is performed. Fig. 7 shows plots of the summed squared errors and a silhouette score for the cluster numbers. The summed squared error indicates the distance to the found cluster centers, whereas the silhouette score specifies how well the individual contributions are matched to the cluster they belong to. Higher numbers within the range from -1 to 1 indicate a better fit of individuals to their cluster. The maximum silhouette score is 0.56 found for an ideal number of five clusters. Fig. 8 shows these clusters plotted along the original (unscaled) feature dimensions. The different clusters will be visually inspected and discussed in Sec. 4.

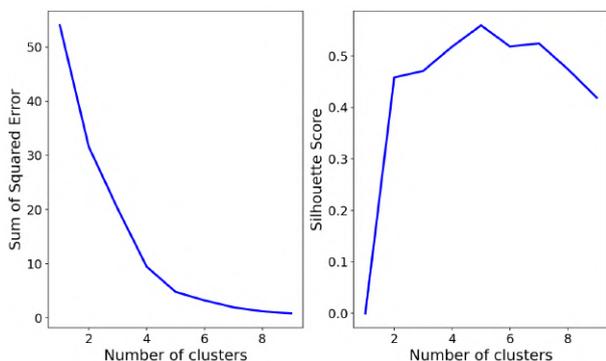


Fig. 7. Error and silhouette scores for K-means clustering.

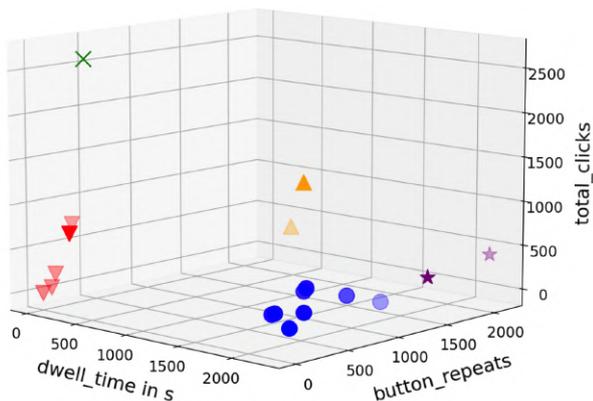


Fig. 8. K-means clusters found for individual participants.

### 3.2 Questionnaire

The completed and valid questionnaires (which included an active IP from the experiment) included eight respondents, five male and three female. This low number of respondents is unfortunate and has to be explained by the nature of the concert study as one of several performances of an open concert event. Some audience members might have left the performance early, some of the respondents were not active participants throughout the experiment, and

some of those who were active did not fill out the questionnaire fully. The ages of the respondents ranged from 21 to 30 years (mean 25.88, SD 4.39). 50% of the respondents underwent five years of musical training or more, 37.5% considered themselves as professional musicians, and 50% as semiprofessional musicians. All respondents participated in some form of musical collective such as a band or choir. Many of them are familiar with the music played at the concert and enjoy listening to that musical style.

Table 1 shows an overview of the preferred interaction modes. Mode 1 was reported to be most enjoyable, exciting, musically enhancing, and encouraging to participate. Reaction from the performers was also reported to be highest for Mode 1, however only with 37.5%. In the additional comments section of the questionnaire and during informal talks with audience members after the concert, the preference towards Mode 1 was also confirmed with the argument that the audience members' contributions were most clearly mapped to changes of soundscapes and light intensities. Table 2 shows the results of the questions referring to the audience members' experience of the participatory concert. Per Boone [32], we treat the Likert-type items as continuous data to get a better impression of the central tendencies compared to medians and variability. In addition, we calculate standard deviation instead of standard error, as suggested by Turchet et al. [17]), for reporting the spread of the items. Means ranged from 3.25 for the recognition as a collaborative act to 6.38 for the enjoyment of the event. The lowest standard deviation was 0.71 for the enjoyment of the general style of music, the musicians' performance, and the interactive parts. The highest standard deviation was 1.70 for the effect that the interaction had on the performance.

Table 1. Survey results for the three interaction modes compared.

	Mode 1	Mode 2	Mode 3	None
Which interaction mode did you enjoy the most?	<b>87.5%</b>	0%	12.5%	0%
Which interaction mode did make you feel most excited?	<b>87.5%</b>	0%	12.5%	0%
Which interaction mode enhanced your musical experience the most?	<b>62.5%</b>	0%	25%	12.5%
Which interaction mode encouraged you to participate the most?	<b>75%</b>	0%	25%	0%
During which interaction mode did you feel that the musicians reacted most to what the audience was doing?	<b>37.5%</b>	25%	0%	<b>37.5%</b>

Table 2. Questions and results (Mean  $\pm$  Standard Deviation) of the post-concert questionnaire (evaluated on 7-point Likert-type items: 1 - strongly disagree, 2 - disagree, 3 - slightly disagree, 4 - neutral, 5 - slightly agree, 6 - agree, 7 - strongly agree)

Likert-type question	Mean $\pm$ SD
I generally enjoyed the event.	6.38 $\pm$ 1.06
I am familiar with music similar to the one played at the concert.	5.63 $\pm$ 1.30
I generally like the style of music played at the concert.	6.75 $\pm$ 0.71
I enjoyed the musicians' performance overall.	6.75 $\pm$ 0.71
I enjoyed the audience interaction parts.	5.25 $\pm$ 0.71
I felt excited during the audience interaction parts.	4.88 $\pm$ 0.85
The interaction software worked well for me.	4.50 $\pm$ 1.07
The interaction software was easy to use.	5.88 $\pm$ 1.25
I enjoyed using the interaction software.	5.38 $\pm$ 0.74
The interaction parts enhanced my musical experience.	4.00 $\pm$ 1.20
I felt that my interactions with the software had an effect on the performance.	3.50 $\pm$ 1.70
I understood what part of the music was created by the audience.	3.75 $\pm$ 1.58
I felt engaged to participate when it was possible.	4.38 $\pm$ 1.19
I feel the audience has some level of ownership of the creative work.	3.50 $\pm$ 1.20
I felt that the performance was surprising.	4.50 $\pm$ 1.51
I enjoyed the surprising elements of the performance.	6.13 $\pm$ 1.13
I felt connected to the musicians.	5.00 $\pm$ 1.20
I felt the performers reacted to what we did as an audience.	4.25 $\pm$ 1.28
I felt connected to other audience members.	3.38 $\pm$ 1.41
The performance felt like a collaborative act of all people present.	3.25 $\pm$ 1.17

## 4 DISCUSSION

### 4.1 Quantities of participation and interaction preferences

With about two thirds of the audience members participating in the interactive parts with approximately 1.00 clicks per second per participant, the basis for analysis is considered sufficient. The remaining audience members might have not stayed throughout the concert, because the performance wasn't the only event that was attended by visitors of the Research Day event. The lower click rate of Mode 1 (0.59 clicks per second) can be related to a decline of clicks at about 75 seconds into the interaction as Fig. 6 shows. This sudden flattening of the interaction curve is thought to be connected to the musicians starting to play their instruments and the focus of the audience shifting towards the stage activity, and as a consequence away from their smartphones.

Mode 2 displays a more steady course of interaction, but the total number of participants went down to only eight audience members for this part. This dip is also apparent

in the survey results, suggesting that the audience experienced Mode 1 as more enjoyable, exciting, and encouraging (see Table 1). The preference for Mode 1 could also be interpreted as an appreciation of a clear and transparent interaction mode, where sound and volume were most obviously connected to the intensity of the corresponding colored stage light. The preference for this mode might also be explained by the fact that audience members value a more conventional audience role in the background or environment of the concert, compared to Mode 2, during which participants were invited to interact with more central parameters of the melody. The comparably low number of eight participants during this interaction might finally be explained by the fact that audiences in general are least engaged in the middle of concerts and their engagement is higher especially at the beginning and towards the end. The lower number of four sliders used for the melody parameters during Mode 2 could also have reduced engagement. Finally, the technical problems on the server side might have left some participants unsatisfied, because during this period their interactions didn't affect the performance. This might also have had an effect on the engagement during the remaining time of Mode 2 when interaction was possible again.

Mode 3, in contrast, had the highest total activity, which might be due to its more unconventional approach, i.e. an interaction with directional lighting in combination with pitch and reverberation effects. It also might partly explain the high number of 15 participants exploring this interaction mode, and hence the highest total click activity out of the three modes. It seemed to keep the participants' attention throughout the interaction, which suggests that this mode was the most sustainably engaging one for audience members. However, these numbers are in contradiction with the questionnaire results, where only 25% reported Mode 3 to be the most engaging one. In that regard, it has to be questioned that higher click numbers directly translate to the experience of more engagement during the interactive parts.

Interestingly, survey respondents rated Mode 1 to be the most interactive, in the sense that musicians reacted to the audience's contributions (see last question of Table 1), and none of the respondents had the impression that this was the case for Mode 3, which could be regarded as something rather unrelated to the musical composition and hence detached from the performers' role in the concert. Lastly, the fact that 37.5% of the survey respondents reported that they did not have a sense of performer response during any of the interaction modes, points to the fact that there is room for modes that facilitate more direct performer-audience-interaction in the compositional design. These might be tools that change parameters that are bound to individual performers, or those that let performers respond to audience contributions more recognizably.

### 4.2 Interaction types

As the cluster analysis (see Sec. 3.1) shows, audience members used the interface very differently. The five clus-

ters, that were inferred by K-means clustering, suggest five different user interaction types. Referring to Fig. 8, the biggest number of users falls into a cluster that is characterized by a high dwell time and low total click and button repeat numbers (eight circles). This seems to be an interested, but rather back-holding group of participants, whom we suggest to refer to as *the sporadic*. The second cluster can be found scoring low in all three dimensions (five downward triangles). This group seems to only participate for a short period before closing the interface again, and are therefore referred to as *the reserved*. Another cluster with two participants can be found with an intermediate dwell time, and intermediate click and repeat numbers (two upward triangles). Since they generally display a greater engagement with those click numbers, we refer to them as *the engaged*. Another small cluster of two participants appear at relatively low total click numbers with a high button repeat rate and high dwell times (two stars). These users seems to be persistent with their contributions, in order to have a maximum effect of specific button controls, and are therefore called *the consistent*. Lastly, we find a single user with a low dwell time, extremely high click number, but low button repeat rate. This is unconventional, since such a high click number would suggest a high chance of clicking similar buttons repeatedly. However, since this user is using many buttons in a non-repeating way, we suggest the term *the explorative*. On the other hand, perhaps this single user should be considered an outlier (and the same could be said of the smaller groups of only two users) which could be grouped into one of the bigger clusters. In any case, these findings would need to be reproduced to gain greater validity and should only be regarded as a first attempt of structuring participatory behavior.

### 4.3 User experience of the interactive concert

As the questionnaire results given in Table 2 show, audience members enjoyed the event and musicians' performance more than they enjoyed the interactive parts, while many were familiar and liked the style of music performed at the concert. However, audience members felt only slightly excited during the interactive parts. Moreover, the lower numbers in functionality of the interface might be related to networking issues reported by some participants. While respondents reported to enjoy the use of the interface, they reported the experienced musical enhancements as neutral. Notably, the effect of their contributions, the transparency of the interaction and the level of ownership were reported with negative tendencies. These issues seem to be most prominent during Mode 2 (see Table 1). Engagement to participate was reported to be slightly positive. Participants felt that the performance wasn't very surprising, but reported that they enjoyed the surprising elements of the concert. This is an interesting result, since it points to the fact that the interaction modes could be designed to offer more surprising elements in future iterations. Lastly, audience members felt slightly connected to the performing musicians, however the performers' reactions to the audience's participation was rated rather neutral. The connec-

tion felt among audience members and the feeling of co-creation during the concert were rated as slightly negative.

### 4.4 Limitations of the study

The results of this study are limited in a number of ways. First, the interactive parts of the concert are very specific in their musical implementation and for that reason are hard to generalize. To produce comparable results for other musical genres, the interaction modes would have to be well adapted. Moreover, the number of participants of the concert was very small, and more studies with bigger audiences would have to be conducted to validate the results of the cluster analysis and the self-report questionnaire. This is especially true given the demographics of the sample, which in this case consisted of individuals in their 20s with strong backgrounds as musicians. In addition, the results of the interaction log of Mode 2 is limited, since we experienced some technical issues with the logging for a part of this mode as described in Sec. 3.1. Finally, the order of the interactive modes limits the comparability of the study and we cannot exclude that a different order would have led to similar results.

### 4.5 Design considerations

Concerning the compositional design of the interaction modes, conventional sliders for accessing musical parameters as implemented in Mode 1 and 2 proved to be intuitive for audience members given the familiarity of such controls in common media tools. Removing the textual description from the sliders as used in previous experiments avoided confusion and the use of symbols rather than text in the interface turned out to be a better choice. However, the number of sliders offered in the interface should be carefully considered and adjusted to the size of the audience. In our study, offering eight sliders for 30 audience members (who were not all constantly active) in Mode 1 worked well. Reducing the number to only four sliders for Mode 2 might have been a reason for lower participation and should be reconsidered.

Another possible reason that Mode 1 was more popular than Mode 2 is the fact that parameter changes affected musical background elements rather than more central parts in the musical foreground. We found that collective control over musical parameters can be problematic and suggest that interaction designs should avoid scenarios that create this type of bottleneck. The idea to collaborate musically as part of an audience with a few musical performers on a stage on a main melody also resulted in less transparency and ultimately lowered the audience members' motivation to participate. This is also supported by the survey results, suggesting that these forms of audience involvement do not necessarily contribute to a feeling of musical co-creation and collective agency. Potential solutions could be the use of multiple melodic elements to be accessed by different groups of audience members or a more controlled distribution of the access to certain parameters at certain points in time.

The design of an x-y plane as implemented for Mode 3 demonstrated the highest number of total clicks and as such suggest high audience interest in interaction modes that go beyond the traditional slider-design. This type of design could be further developed with options for directly interacting with the plane rather than using increment and decrement buttons to control the collective state. In addition, those contributions could use more refined methods for aggregating the individual levels in future design solutions, which will become a necessity for performances with bigger audiences.

Concerning the mapping of the interface controls to sound and light parameters, we found that a simple and clear design as implemented in Mode 1, where soundscapes in the musical background were clearly mapped to certain soundscapes and light colors, seem to better fit the audience's traditional role as being part of the stage surrounding and create rather atmospheric or environmental contributions than influencing central musical elements. Nevertheless, participants valued Mode 3 in terms of a more unconventional form of involvement that is only partly influencing the musical content (in this case only one synthesizer's pitch and reverb parameters) but happens on a somewhat different level (in this case the directional lighting). We see potential in designs that further explore the space of controlled directional and colored stage lighting. Future designs could enable a more individualized access to specific parts of the lighting or let groups of audience members control those together. Ultimately, it may be worth exploring more experimental and playful modes of interaction that offer the audience interactions with the stage design rather than the control of main musical parameters.

#### 4.6 Future work

With validation of the results of this study in mind, future work should include concert studies with bigger audiences. This would entail the design of more refined interactive modes that replace the discrete changes by aggregated ones in a refined way. Solutions might also be found with asynchronous modes, that let participants create musical elements before the actual concert, or a more individualized design that displays interactions first for the client before sharing them with the entity of the audience. This also concerns the design process of the interaction modes, that can include audience members in a more profound manner. In addition, the analysis of recorded media data such as audio, video, and MIDI could help better understand the nature of the audience members' participatory behavior, and in conjunction with log data and self-reports give a better qualitative understanding of audience participation. Finally, the implementation of a greater level of interactivity, that takes into account reactions of the performing musicians, should be explored in more detail with future studies.

## 5 CONCLUSION

For this study on audience participation, the authors designed three interaction modes in collaboration with an electronic live music trio and a lighting designer. Audience members could control soundscapes and colored lighting (Mode 1), parameters of the central melody together with intensities of rhythmically flashing lights (Mode 2), and pitch and reverb of a synthesizer in conjunction with light directions on an interface involving an x-y plane (Mode 3). The following three paragraphs relate to the three research questions as suggested in the introduction section respectively.

Cumulative plots of the interaction logs and analysis of the self-report questionnaire showed a preference of the audience towards a simply structured, environmental participation (Mode 1) over the control of rhythmically bound, central musical parts of the performance (Mode 2). Most active participation was however observed during Mode 3, which went beyond conventional slider controls in offering a directional lighting interface.

Cluster analysis using K-means clustering suggested five different user types, that were found along the three feature dimensions *total click number*,  *dwell time*, and *button repeats*. A first attempt of a user typology for the interactive concert was given, which will need validation by future studies with larger sample sizes.

Future work on interaction design should include studies with a higher number of participants and, as a consequence, more work on refined, possibly asymmetric interaction modes. This especially concerns the temporal domain, where the involvement of parallel user actions on a collective interface seems limited beyond a higher number of participants. In addition, other modes for interaction beyond the discrete, slider-based control should be investigated. Future designs could also include gamified elements, multi-modal feedback, and support social interaction among audience members. Finally, machine-learning-inspired modes for participation based on previous training by audience members before the concert seems a valid approach. Ultimately, we see greater potential for audience participation in the control of parts of the performative staging, such as lighting, stage design, or musical background elements, than in the direct control of main musical parameters.

## 6 ACKNOWLEDGMENT

The authors are grateful to all individuals involved in the concert event including the musicians of the group *Kubbi*: Vegard Kummen, Tobias Øymo Solbakk, Kristian Olav Vågen, lighting designer Mike Holm, and the organizer and host of the Research Day event Prof. Rolf Kristensen, as well as Prof. Jason Freeman and his lab for advising on the implementation and analysis connected to this study.

## 7 REFERENCES

- [1] P. Bourdieu, *The field of cultural production: Essays on art and literature* (Columbia University Press, 1993).
- [2] V. P. Glaveanu, *Distributed Creativity: Thinking Outside the Box of the Creative Individual*, 1st ed., SpringerBriefs in Psychology (Springer International Publishing, Cham, 2014), URL <https://doi.org/10.1007/978-3-319-05434-6>.
- [3] E. F. Clarke and M. Doffman, *Distributed creativity: Collaboration and improvisation in contemporary music* (Oxford University Press, 2017 December), URL <https://doi.org/10.1093/oso/9780199355914.001.0001>.
- [4] G. L. Hagberg, “The ensemble as plural subject,” in E. F. Clarke and M. Doffman (Eds.), *Distributed Creativity: Collaboration and Improvisation in Contemporary Music*, vol. 2, pp. 300–313 (Oxford University Press, 2017), URL <https://doi.org/10.1093/oso/9780199355914.003.0025>.
- [5] B. Taylor, “A History of the Audience as a Speaker Array,” *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 481–486 (2017 May), URL <http://doi.org/10.5281/zenodo.1176324>.
- [6] G. Levin, S. Gibbons, G. Shakar, Y. Sohrawardy, J. Gruber, J. Lehner, *et al.*, “Dialtones (A Telesymphony): Final Report,” (2001 September), URL [http://www.flong.com/storage/pdf/reports/dialtones\\_report.pdf](http://www.flong.com/storage/pdf/reports/dialtones_report.pdf).
- [7] J. Oh and G. Wang, “Audience-participation techniques based on social mobile computing,” *Proceedings of the International Computer Music Conference (ICMC)* (2011 August), URL <https://www.gewang.com/publish/files/2011-icmc-audience.pdf>.
- [8] S. W. Lee and J. Freeman, “echobo: A mobile music instrument designed for audience to play,” *Proceedings of the 13th International Conference on New Interfaces for Musical Expression (NIME)*, pp. 450–455 (2013 May), URL <http://doi.org/10.5281/zenodo.1178594>.
- [9] B. Matuszewski, “A Web-based Framework for Distributed Music System Research and Creation,” *Journal of the Audio Engineering Society*, vol. 68, no. 10, pp. 717–726 (2020 October), URL <https://doi.org/10.17743/jaes.2020.0015>.
- [10] L. C. Carpenter, “Video imaging method and apparatus for audience participation,” US Patent 5,365,266 (1994), URL [https://patents.google.com/patent/WO1993012614A1/lista\\_geral\\_produtos\\_quimicos.pdf](https://patents.google.com/patent/WO1993012614A1/lista_geral_produtos_quimicos.pdf).
- [11] L. Dahl, J. Herrera, and C. Wilkerson, “Tweet-Dreams: Making Music with the Audience and the World using Real-time Twitter Data,” *Proceedings of the 11th International Conference for New Interfaces for Musical Expression (NIME)*, pp. 272–275 (2011 June), URL [https://www.nime.org/proceedings/2011/nime2011\\_272.pdf](https://www.nime.org/proceedings/2011/nime2011_272.pdf).
- [12] S. W. Lee, A. D. de Carvalho Jr, and G. Essl, “Crowd in C [loud]: Audience participation music with online dating metaphor using cloud service,” *Proceedings of the 2nd Web Audio Conference* (2016 April), URL [https://webaudioconf.com/\\_data/papers/pdf/2016/2016\\_65.pdf](https://webaudioconf.com/_data/papers/pdf/2016/2016_65.pdf).
- [13] Y. Wu, L. Zhang, N. Bryan-Kinns, and M. Barthes, “Open symphony: Creative participation for audiences of live music performances,” *IEEE MultiMedia*, vol. 24, no. 1, pp. 48–62 (2017 May), URL <https://doi.org/10.1109/MMUL.2017.19>.
- [14] A. Stolfi, M. Barthes, F. Goródsy, and A. D. de Carvalho Junior, “Open Band: A Platform for Collective Sound Dialogues,” *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences* (2017 August), URL <https://doi.org/10.1145/3123514.3123526>.
- [15] A. S. Stolfi, A. Milo, and M. Barthes, “Playsound. space: Improvising in the browser with semantic sound objects,” *Journal of New Music Research*, vol. 48, no. 4, pp. 366–384 (2019 September), URL <https://doi.org/10.1080/09298215.2019.1649433>.
- [16] O. Hödl, C. Bartmann, F. Kayali, C. Löw, and P. Purgathofer, “Large-scale audience participation in live music using smartphones,” *Journal of New Music Research*, vol. 49, no. 2, pp. 192–207 (2020 February), URL <https://doi.org/10.1080/09298215.2020.1722181>.
- [17] L. Turchet, T. West, and M. M. Wanderley, “Touching the audience: musical haptic wearables for augmented and participatory live music performances,” *Personal and Ubiquitous Computing*, pp. 1–21 (2020 March), URL <https://doi.org/10.1007/s00779-020-01395-2>.
- [18] N. Lorway, E. Powley, and A. Wilson, “Autopia: An AI collaborator for live networked computer music performance,” *Proceedings of the 2nd Conference on AI Music Creativity (MuMe + CSMC)* (2021 July), URL <http://repository.falmouth.ac.uk/4307/>.
- [19] G. Weinberg, “Interconnected musical networks: Toward a theoretical framework,” *Computer Music Journal*, vol. 29, no. 2, pp. 23–39 (2005 Summer), URL <https://doi.org/10.1162/0148926054094350>.
- [20] R. C. Renwick, *Topologies for Network Music*, Ph.D. thesis, Queen’s University Belfast (2017 February), URL <https://core.ac.uk/download/pdf/96665533.pdf>.
- [21] J. Freeman, “Web-based collaboration, live musical performance and open-form scores,” *International Journal of Performance Arts and Digital Media*, vol. 6, no. 2, pp. 149–170 (2010), URL [https://doi.org/10.1386/padm.6.2.149\\_1](https://doi.org/10.1386/padm.6.2.149_1).
- [22] J. T. Allison, Y. Oh, and B. Taylor, “NEXUS: Collaborative Performance for the Masses, Handling Instrument Interface Distribution through the Web,” *Proceedings of the 13th International Conference on New Interfaces for Musical Expression (NIME)*, pp. 1–6 (2013 May), URL

[https://www.nime.org/proceedings/2013/nime2013\\_287.pdf](https://www.nime.org/proceedings/2013/nime2013_287.pdf).

[23] A. York, “Transforming Musical Performance: The Audience as Performer,” *Innovation in Music: Performance, Production, Technology, and Business*, pp. 306–330 (2019 June), URL [https://doi.org/10.26530/OAPEN\\_1005168](https://doi.org/10.26530/OAPEN_1005168).

[24] B. Matuszewski, N. Schnell, and F. Bevilacqua, “Interaction Topologies in Mobile-Based Situated Networked Music Systems,” *Wireless Communications and Mobile Computing* (2019 March), URL <https://doi.org/10.1155/2019/9142490>.

[25] O. Hödl, G. Fitzpatrick, F. Kayali, and S. Holland, “Design Implications for Technology-Mediated Audience Participation in Live Music,” *Proceedings of the 14th Sound and Music Computing Conference 2017*, pp. 28–34 (2017 May), URL <http://eprints.cs.univie.ac.at/5451/>.

[26] A. Xambó and G. Roma, “Performing Audiences: Composition Strategies for Network Music using Mobile Phones,” *Proceedings of the 20th International Conference on New Interfaces for Musical Expression (NIME)*, pp. 55–60 (2020), URL <https://doi.org/10.5281/zenodo.4813192>.

[27] M. Jung and V. Kummen (in press), “Hacking the concert experience - exploring co-creative audience interaction at a chiptune live performance,” *Innovation in Mu-*

*sic: Performance, Production, Technology, and Business* (2023).

[28] R. M. Alguliyev, F. J. Abdullayeva, and S. S. Ojagverdiyeva, “Log-File Analysis to Identify Internet-addiction in Children,” *International Journal of Modern Education & Computer Science*, vol. 13, no. 5 (2021 October), URL <https://doi.org/10.5815/ijmecs.2021.05.03>.

[29] I. Arroyo and B. P. Woolf, “Inferring learning and attitudes from a Bayesian Network of log file data,” in C.-K. L. et al. (Ed.), *Artificial Intelligence in Education*, pp. 33–40 (IOE Press, 2005 July), URL <https://dl.acm.org/doi/abs/10.5555/1562524.1562535>.

[30] O. Kramer, *Machine Learning for Evolution Strategies* (Springer, 2016), URL [https://doi.org/10.1007/978-3-319-33383-0\\_5](https://doi.org/10.1007/978-3-319-33383-0_5).

[31] X. Yi, L. Hong, E. Zhong, N. N. Liu, and S. Rajan, “Beyond Clicks: Dwell Time for Personalization,” *Proceedings of the 8th ACM Conference on Recommender Systems*, pp. 113–120 (2014 October), URL <https://doi.org/10.1145/2645710.2645724>.

[32] H. N. Boone, D. A. Boone, et al., “Analyzing Likert Data,” *Journal of Extension*, vol. 50, no. 2, pp. 1–5 (2012 April), URL [https://archives.joe.org/joe/2012april/pdf/JOE\\_v50\\_2tt2.pdf](https://archives.joe.org/joe/2012april/pdf/JOE_v50_2tt2.pdf).

## THE AUTHORS



Matthias Jung

Matthias holds a Bachelor from Paderborn University in Popular Music and Media and a Master in Electronic Media from Stuttgart Media University. For his Master thesis he joined IOSONO, where he was engaged with auralization models for wavefield synthesis. Before starting his Ph.D. in Norway, Matthias worked at Berlin’s state academy as audio production specialist and headed the program team of Most Wanted: Music conference and Hybrid Music Lab - both centered around artistic and technological aspects of musical creativity. He recently spend a semester at Georgia Tech as a visiting scholar.



Ian Clester

Ian Clester composes music and programs. His research interests include the design of languages, environments, and tools for composing and performing music with computational or generative aspects. His performance practice includes electric guitar and live coding, and he is a founding member of the MIT Laptop Ensemble. His work has been recognized with best paper (Web Audio Conference 2022) and best poster (Audio Mostly 2021) awards. He received his B.Sc. in EECS and music and his M.Eng. in EECS from MIT, and he is currently pursuing a Ph.D. in music technology at Georgia Tech.

## Article 3 - Intelligent Music Performance Systems: Towards a Design Framework

### Abstract

*This article attempts to define and typologise the main principles for the design of Intelligent Music Performance Systems (IMPS). It presents a three-dimensional framework based on studies of proxemics combined with findings from AI design research. Each of these dimensions – embodiment, participation and autonomy – is presented together with existing taxonomies, then integrated into an analytical framework. This framework informs the discussion of nine historical cases of IMPS from the 1950s to the present to gain a refined understanding of their interactive design. The discussion leads back to three main tendencies in IMPS design – instruments, systems and agents – and the article concludes by combining the proposed framework with ideas from Speculative Design Research.*

Keywords: Intelligent Music Performance Systems, Musical Proxemics, Interactive Composition, Speculative Design Research

Version: published article

# Intelligent Music Performance Systems: Towards a Design Framework

Matthias Jung

*Department of Popular Music, Faculty of Fine Arts, University of Agder, Kristiansand (Norway)*

Matthias holds a Bachelor from Paderborn University in Popular Music and Media and a Master in Electronic Media from Stuttgart Media University. Before joining the Ph.D. program *Popular Music Performance* at UiA in Norway, he worked at Berlin's state academy as music production specialist and headed the program team of *Most Wanted: Music* conference. His recent research focuses on bio-inspired coding and creative symbioses within music performance systems, some of which he carried out lately at Georgia Tech University in Atlanta as a visiting scholar.

[matthias.jung@uia.no](mailto:matthias.jung@uia.no)

## Abstract

This article attempts to define and typologize the main principles for the design of Intelligent Music Performance Systems (IMPS). It presents a three-dimensional framework based on studies of proxemics combined with findings from AI design research. Each of these dimensions – *embodiment*, *participation*, and *autonomy* – is presented together with existing taxonomies, then integrated into an analytical framework. This framework informs the discussion of nine historical cases of IMPS from the 1950s to the present, to gain a refined understanding of their interactive design. The discussion leads back to three main tendencies in IMPS design – *instruments*, *systems*, and *agents* – and the article concludes by combining the proposed framework with ideas from Speculative Design Research.

## Keywords

Intelligent Music Performance Systems, Musical Proxemics, Interactive Composition, Speculative Design Research

## 1 Introduction

Intelligent Music Performance Systems (IMPS) are performance systems which involve artificially intelligent components in the generation of their creative result. These systems make use of computers or other creative technology to create musical performances and represent a subclass of Interactive Digital Musical Instruments or Interactive Music Systems (to emphasize their systemic approach; Rowe 1992, 43). Those latter classes, in turn, belong to the family of Digital Musical Instruments (DMI) and, more generally, to New Interfaces for Musical Expression (NIME) – terms coined in the last 30 years or so (see Figure 1). In the 1990s and 2000s, the term Interactive Performance System also appears in the literature (Pressing 1990, 12) to point to an early form of IMPS. Note that not all IMPS necessarily make use of digital technology; their intelligence can also stem from non-digital concepts, as is indicated in Figure 1. Most of today's IMPS, however, are in fact digital systems.

IMPS typically work with a combination of sensors and digital media to enable real-time interactions among performers, instruments, and environments. Interaction models include a human input layer and an output layer generated, in most cases, by a computer (Winkler 1998 as cited in Drummond 2009, 126). The generated output can be fed back into the system and similarly serves as an input for the human performer, resulting in a circular process also referred to as a human–machine interaction loop in the larger field

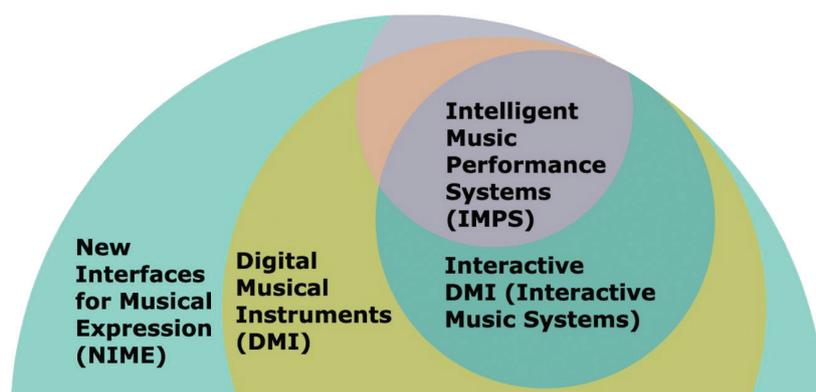


Figure 1. Delimitation of IMPS.

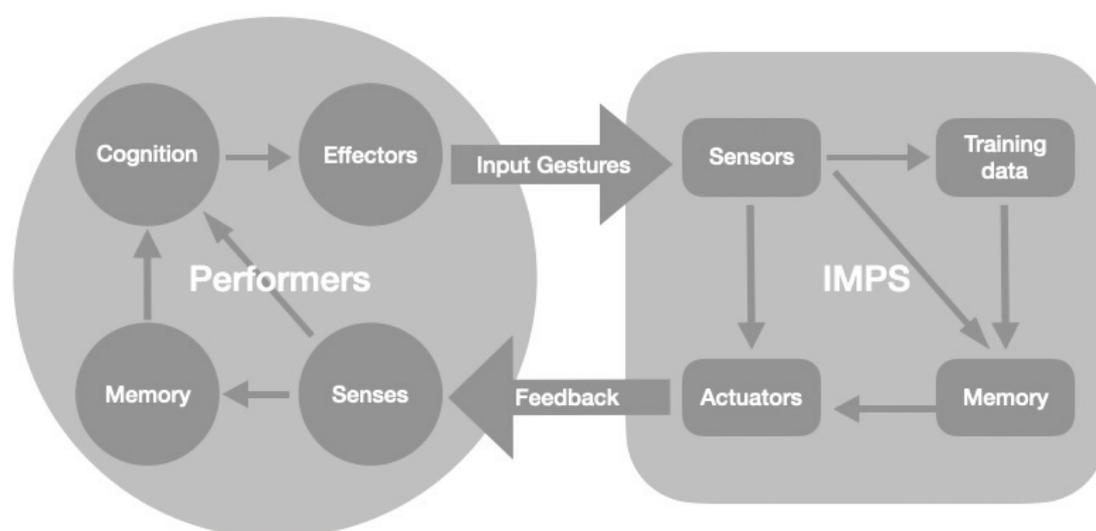


Figure 2. Performer–system interaction loop.

of sonic interaction design (Visell 2013). Collins labels those sounds which can be adapted by players of computer games as *interactive diegetic sounds* and those interactive processes which implement a form of bodily involvement as *kinetic gestural interaction* (Collins 2008, 125–127). Likewise, DMI which incorporate haptic elements can be understood to provide a twofold feedback (auditory and haptic), as proposed by Miranda and Wanderley (2006, 71–83). Researchers have assigned DMI, including AI components, a computational memory and cognition layer as well (Bongers 2007). Figure 2 is adapted from Drummond (2009, 131) and attempts to integrate the aforementioned models of human–computer interaction into the interaction between human performers and IMPS. The literature typically refers to the artificially intelligent components of IMPS as (autonomous) musical agents.<sup>1</sup>

IMPS have been developed for a multitude of musical genres, from classical music (for example, *Open Symphony* [Wu et al. 2017]) to experimental electronic music (for example, *Autopia* [Lowray et al. 2021]). Many systems may be used independent of their native musical culture, though the development of IMPS is always subject to cultural context and should be recognized as such in the analytical process. Much analytical work has been targeted towards considerations of system design, and many articles conclude with design recommendations of value to the community of creators of those systems. A word search for *design goals*, *design considerations*, *design principles* and *design guidelines* in two main

1 For a general introduction to the concept of agent systems, see Woolridge (2009).

publication venues for IMPS, the NIME proceedings and the ICCC proceedings,<sup>2</sup> produced 50 results for articles with these words in their titles since 2010. The results were based on 1,330 papers (NIME) and 400 papers (ICCC) within the timeframe 2010–2022, and demonstrate a growing need for orientation in a field which is becoming much more complex. Regarding the current state of research, Tatar and Pasquier (2019) analyzed 78 musical agents using 13 design criteria, but their work was directed towards autonomous agents and not limited to real-time contexts. Gifford et al. (2018) analyzed 23 real-time systems and suggested a taxonomy focused on computational co-improvisation with human performers (Gifford et al. 2018, 33). This recent work is recommended for readers who are not familiar with the general field of IMPS.

The present article reviews the existing research through the lens of performance systems by applying a proxemic perspective as proposed by Hall (1966) which is connected to models for autonomous musical agents. This approach suggests three dimensions for IMPS design (section 2). A derived framework (section 3) is then applied to a small number of historical cases from the 1950s to the present time (section 4). Section 5 then introduces the concept of Speculative Design Research and applies it to the proposed IMPS design model. Section 6 summarizes the article's contributions.

## 2 Three design dimensions for IMPS

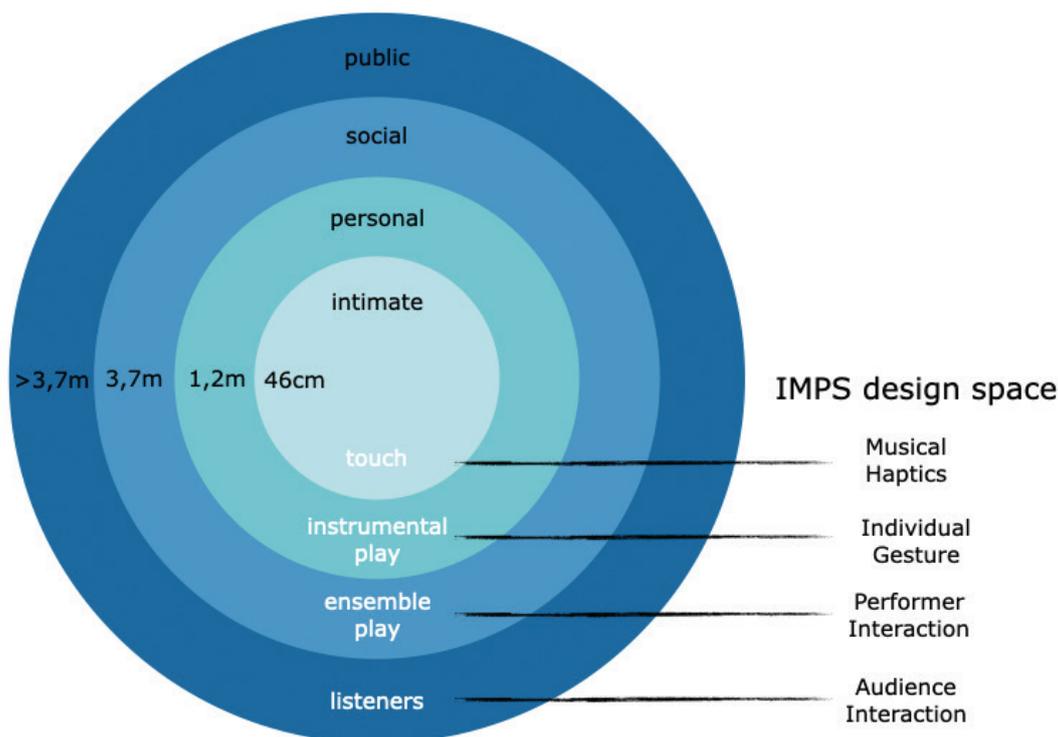
In *The Hidden Dimension*, Hall proposes a perception-based model for interpersonal distance including four proxemic zones: intimate, personal, social, and public (Hall 1966, 121–129). Since the 1960s, this model has offered a relevant perspective on how physical distance affects the behavior of animals and humans. Though it was not initially intended for the analysis of creative practice, it has also been applied to music listening by Collins and Dockwray (2015) and Dockwray (2016). In this section, it will be used to structure the physical distance between human and non-human agents in the interests of adapting it to the design of IMPS, as illustrated in Figure 3.

### 2.1 Embodied instrument design

The first design dimension relates to both the intimate and the personal zones of Hall's model (the two inner zones in the graph). They are the proxemic zones within which a performer would typically interact with an instrument. Tactile interaction is part of the intimate zone, given the immediate contact between instruments, controllers or sensors and the human body. The space within which an individual performer acts is the personal zone. Technological trends have also introduced the potential for merged devices and bodies, including physically implanted instruments. Regarding this realm of (musical) wearables and implants, Mueller et al. (2020) conclude that the most imminent application might involve what they call 'integrated' devices, which are neither separate from nor implanted into the body, but instead overlap with it and its senses. Ultimately, this initial IMPS design dimension evokes Leman's (2008) understanding of embodied instruments and digital technology as central design components for distributed systems.

Regarding the design for these two proxemic zones, Cook proposed 22 principles at the first NIME conference (Cook 2001), encompassing aspects such as respect for the bandwidth of the performer; an operative guideline that the instrument should generate music

2 NIME proceedings stands for the *Proceedings of the International Conference on New Interfaces for Musical Expression* (accessible online via <https://www.nime.org>), and ICCC proceedings stands for the *Proceedings of the International Conference of Computational Creativity* (accessible online via <https://computationalcreativity.net>).



**Figure 3.** Hall's proxemic model applied to the IMPS design space.

right away (and offer subtlety for interaction later); and the insight to think in terms of pieces rather than controllers when designing NIME (Cook 2001). When designing DMI, in turn, it is critical to consider the mapping of the embodied gestures and the output of the system. West et al. (2021) recently conducted a study that explored design criteria according to the themes of *control*, *legibility*, and *sound* among DMI users. They position such mapping at the heart of most forms of electronic artistic practice and argue that elaborate mappings must inform fully integrated musical instruments.

Regarding the analysis of NIME, Birnbaum et al. (2005, 194) suggest a seven-dimensional model accounting for what expertise is required to play the instrument, how the instrument is controlled, how many degrees of freedom it offers, what feedback modalities it uses for its output, how many players can interact with it, how it is distributed in space, and what role sound plays.

## 2.2 Designing for social interactivity

The following section engages with the two outer proxemic zones of Hall's model, the social (at a distance of 1.2 to 3.7 meters) and the public (any distance beyond 3.7 meters). These zones typically engage the affordances of a performance system for interpersonal communication and collective creation. For systems which assume a performative center (stage), the social zone might involve the interaction among the stage performers or among the audience members, given that members of these groups are usually positioned near one another. The public zone, on the other hand, might involve the interaction *between* these groups.

Distributed creativity is the subject of many studies of IMPS and approached from various perspectives. Glăveanu (2014) proposes a separation between the creative material and the social dimension of distributed systems, one which was picked up by Kantosalo and Takala (2020) while theorizing the creative behavior of a collective. Placing the creative contribution at the center of their model, they specified 'five Cs' in play during a collective

creative act – *context, collective, collaboration, community, and contributions* – all of which are connected to each other as well. Regarding the aesthetic dimensions of relational art, one should begin with Bourriaud (2002).

On a more concrete design level, Weinberg suggests a theoretical framework consisting of ‘Interconnected Musical Networks’ for the field of music composition and performance (Weinberg 2005, 31–37). He recognizes a series of typologies and architectures which he separates into systems for smaller groups of up to ten participants and those for groups above ten participants. Similar typologies appear in the work of Renwick (2017, 117–138) and Matuszewski (2019, 4), the latter of whom developed six different graphs for interactive music performance systems: *disconnected, unidirectional, bi-directional, centrifugal star, centripetal star* and *forest graph*.

Systems intended to accommodate larger groups of audience members are typically referred to as participatory compositions or systems for audience participation. Earlier examples of real-time systems in the pre-smartphone era include *Reflective Paddles* (Carpenter 1991), *Dialtones* (Levin 2001) and *Glimmer* (Freeman, 2005), whereas smartphone-centered systems include *Echobo* (Lee and Freeman 2013), *Open Symphony* (Wu et al. 2017) and *The Singularity* (York 2019). Regarding aspects of design for participatory systems, Lee and Freeman suggest five criteria: *accessibility, musical safety, initiation, attractiveness* and *transparency* (Lee & Freeman 2013, 451–452). Hoedl proposes 16 key design issues based on an audience survey (Hoedl 2017, 31–33),<sup>3</sup> and Xambó and Roma detail 13 compositional dimensions which account for *performer and audience roles, location, and feedback types*, among other things (Xambó & Roma 2020, 56–57). Mazzanti et al. present a six-dimensional model for the categorization of compositional designs which includes audience interaction within the following categories: *control design freedom, system versatility, audience interaction transparency, audience interaction, focus* and *active/passive audience affinity* (Mazzanti et al. 2014, 30).

### 2.3 Designing for autonomy

Certain aspects of artificially intelligent implementations, sometimes also referred to as artificially or computationally creative systems, have gained momentum in recent decades. Advances in computational power have been especially impactful on the field of real-time performance systems, given that AI processes typically involve heavy computation loads. Regarding aspects of design, McCormack et al. propose the so-called AI design circle (McCormack et al. 2020, 3) to argue that creative practice and AI system design work best when they mutually inform each other during both coding and performance.

Designing AI systems for musical creation involves different levels of non-human autonomy. Eigenfeldt et al. (2013, section ‘A Taxonomy of Musical Metacreation’) specify eight levels of metacreative paradigm for creative AI systems, where 0 involves no metacreativity and 7 describes the design of a musical agent which demonstrates volition (Table 1).<sup>4</sup> Note, however, that this definition does not accommodate interactivity as such, and feedback between human and machine agents is not included in the theoretical model.

3 These 16 design issues fall within five categories in Hoedl (2017, 33): *music-related information* (skilfulness, masterfulness), *motivation* (distinctiveness, obtrusiveness, expressiveness, sociability, exposure), *behaviour* (mood, diversity, objects, communication), *mobile technologies* (readiness, openness), and *opinion* (appropriateness, contradiction, creativity).

4 The term metacreativity is used for systems which generate creative artefacts without human contribution. A full explanation can be found Pasquier et al. (2016).

**Table 1.** Levels of autonomy in metacreative musical systems, as proposed by Eigenfeldt et al. (2013)

Level	Autonomy paradigm	Definition
0	No metacreativity involved	
1	Independence	The use of any process on a musical gesture that is beyond the control of the composer.
2	Compositionality	The use of any process to determine the relationships between pre-defined musical gestures.
3	Generativity	The generation of musical gestures.
4	Proactivity	Systems/agents that are able to initiate their own musical gestures.
5	Adaptability	a) Agents behave in different ways over time due to their own internal evolution; b) agents interact and influence one another.
6	Versatility	Agents determine their own content without predefined stylistic limits.
7	Volition	Agents exhibit volition, deciding when, what and how to compose/perform.

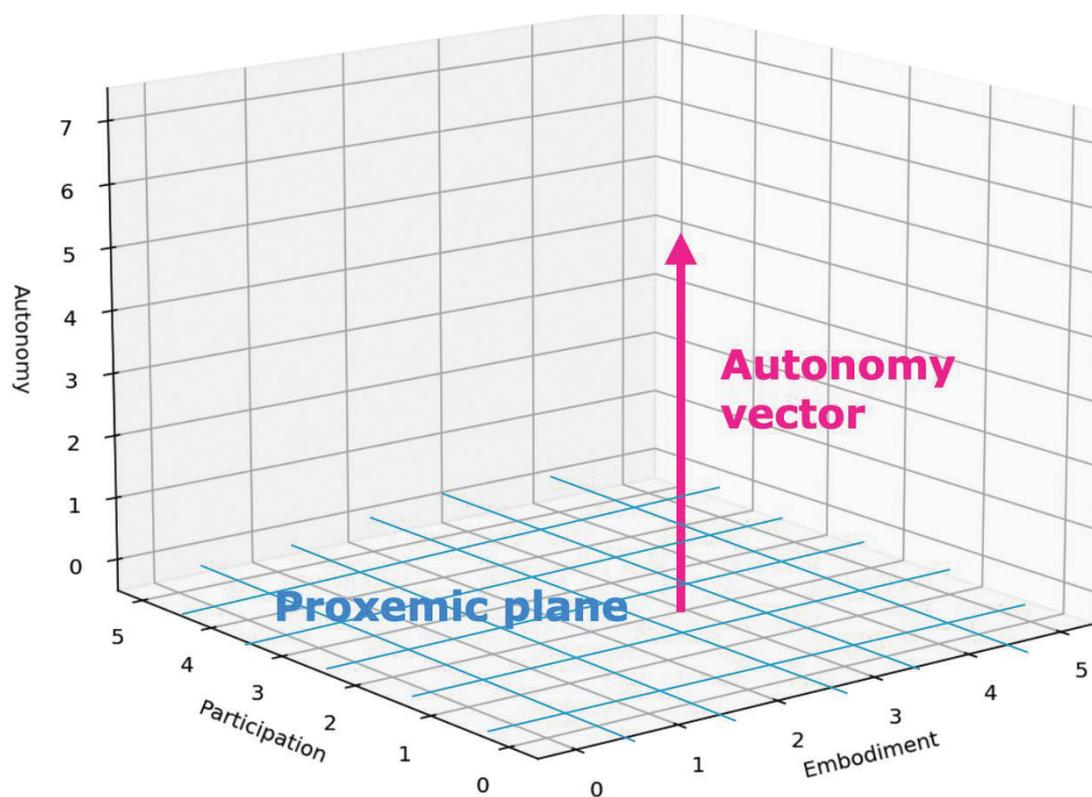
Designing for autonomy can affect all proxemic zones. On an intimate and personal level, for example, autonomy might involve a gesture recognition system trained by user data or algorithms which draw upon some form of external knowledge database for interactive purposes. At a social or public level, autonomy might involve humanoids or other interactive social bots with which performers and audiences interact. (For an extensive discussion of musical dialogue, see King and Gritten [2017]; for a focus on musical interaction with AI-based systems, see Llano et al. [2020]. For a discussion of non-human agency with a focus on human aspects, see Dahlstedt [2021, 882–887], whose findings are based on Latour’s more general theory of non-human agents [Latour 2005, 63–86].)

### 3 Towards a model of IMPS design

Figure 4 is a visual representation of the proposed model for IMPS design. The x-axis refers to the first two proxemic zones, that is, the degree to which a system implements an embodied approach. Traditionally, this represents the instrument design level, including the ways in which a performer might interact with the system on a haptic and individual level, including such gestures as pressing keys, plucking strings, waving the arms, touching screens, or conveying body gestures to a camera system. All of these gestures fall within the first two proxemic zones and relate to the individual body of a system user, which is why this is referred to as the dimension of *embodiment* design. The axis uses a five-step quantization for categorizing IMPS, where 0 indicates no embodied design, and 5 indicates a high degree of embodiment in the system.

The second design dimension, *participation* design, is visualized along the y-axis and relates to the two outer proxemic zones of Hall’s model, that is, the social and public zones which concern the interactions between performers in the system. This could be, for example, the affordances of a system concerning each performer’s sonic contributions or the integration of audience contributions within the public zone. Like the *embodiment* axis, the *participation* axis makes use of a 5-point scale with a 0 point. Note that the placement on the x and y-axes does not always translate to a literal distance from a system’s physical interface. It can also involve a virtual impression of closeness, or a perceived distance from one’s co-performers.

Lastly, the third design dimension, *autonomy* design, is represented by the z-axis of the model. This axis describes the extent to which non-human agents are integrated into a system, and, more specifically, the extent of design development of the autonomous contributions within a system. This axis implements Eigenfeldt et al.’s seven-level model



**Figure 4.** Visualization of the proposed model for IMPS design.

for musical artificial creativity (Eigenfeldt et al. 2013, section ‘A Taxonomy of Musical Metacreation’).

### 4 Case studies

This section investigates nine cases of IMPS, and categorizes them with the help of the proposed model described above. The selection of cases was influenced by the author’s review process, access to and experience with IMPS, and does not claim to encompass the many systems out there. Instead, the selection reflects a diversity of systems in terms of historical genesis, technology, and cultural performance context. It is intended to be a prototypical test for the model which could be extended in future work to offer more validity to the model. Table 2 presents an overview of the selected IMPS cases.

**Table 2.** Overview of analyzed cases

No.	System	Year	Creator	Embodiment	Participation	Autonomy
1	EIS	1950s	Oliveros, Pauline	0	0	0
2	Voyager	1986	Lewis, George E.	0	2	2
3	Cypher	1992	Rowe, Robert	0	0	2
4	Reactable	2005	Jordà, Sergi	4	5	2
5	GREIS	2010	Van Nort, Doug	3	0	3
6	FILTER	2013	Van Nort, Doug	3	0	4
7	Echobo	2013	Lee, Sang W. & Freeman, Jason	0	5	2
8	DYCI2	2018	Nika, Jérôme et al.	2	2	5
9	MHWA	2020	Turchet, Luca et al.	5	3	2

#### 4.1 EIS, GREIS and FILTER

The Expanded Instrument System (EIS) was developed by Pauline Oliveros, starting in the late 1950s, via the creative use of tape delay (the latency between the playback and record head of a tape recorder), and continuing through her lifetime, as technology advanced (Oliveros 2008). The first versions of EIS do not utilize digital technology so cannot be considered DMIs as such. In 1967, instruments were introduced to the system as sound sources, and in 1983, the tape delay was replaced by digital delay. This design was further developed in the 1990s and 2000s, with the help of more refined digital spatialization models (Van Nort et al. 2013, 306). From the early 2000s onwards, the system, which Oliveros often used with a foot pedal, allowed for the real-time control of up to 40 delay times, ranging from milliseconds to one minute, acting as a ‘time machine’ and an ‘improvising partner, with a history of pushing the performer to make instantaneous decisions and to deal with the consequences musically which long predates the rich work in interactive performance systems being done today (...) [T]he result is a fluctuating mix of live instrumental input with machine feedback that has the feel of human partnering, mirroring his/her own sounds interactively’ (Van Nort et al. 2013, 306). The development of EIS should be framed in tandem with Oliveros’s compositional and performative concept of *deep listening*, wherein individuals enter a meditative state and closely listen to musical events of both human and machine nature (Oliveros 1979, 2005).

In the later 2000s, EIS inspired the Granular-Feedback Expanded Instrument System (GREIS), which was developed by Doug Van Nort, one of Oliveros’s musical co-performers in the trio Triple Point. While Oliveros used EIS mainly to achieve longer delay lines in the range of 20 to 60 seconds via a foot pedal while playing the accordion, Van Nort performed GREIS via hand-gestures with a stylus on a tablet to achieve time delays in the range of milliseconds to a few seconds (Van Nort 2013, 306). These gestures impacted the timbre of the trio’s samples by breaking them down into grains, allowing the performers to use their voices as navigators of the timbre space (Van Nort 2013). Through the performers’ interactions with GREIS, the system builds what Van Nort calls *episodic* and *semantic memory* – the former relating to audio and control parameters, the latter to gestures – which then can be used to control other parts of the system (Van Nort 2013, 311). A full description of GREIS can be found in Van Nort (2010).

While EIS and GREIS are both used by a performer, the FILTER system serves as an actively improvising co-performer or partner rather than an instrument, in this way expanding Oliveros and Van Nort’s trio into a quartet (Triple Point + FILTER, Van Nort 2013, 322). FILTER stands for Freely Improvising, Learning and Transforming Evolutionary Recombination, and its major contributions reside in the field of machine listening – it is, in short, an attempt to implement Oliveros’s deep listening in a non-human performer with the help of advanced computational methods. FILTER is intended to move towards ‘a system that learns “action” – in the sense of intention – that is embedded in the low-level structure of the audio stream of its improvising partner’ (Van Nort 2013, 314). This listening agent uses two different modes: gestural listening and textural listening. Gestural listening happens mainly on three different time scales, from the note level to the phrase level, and in the absence of recognized gestures, the system turns towards textural listening (Van Nort 2013, 317). The results of these listening processes inform the system’s semantic and episodic memories; it then uses a genetic algorithm to generate its active behavior (Van Nort 2013, 319).

Comparing EIS, GREIS and FILTER, it is clear that EIS is the least technologically developed system – it is, after all, restricted to effect generation and does not demonstrate intelligent behavior. Although it is often used with a foot pedal, the system itself does not

implement embodied components per se, nor is it designed for interactions among performers or audience members. This is why EIS appears in the lower areas of the model along all three axes (see Figure 5). GREIS, however, introduces notions of semantic memory generated by hand and voice gestures that navigate timbral qualities of the system's output, which is why it is seen to demonstrate a higher level of embodied design. Finally, the FILTER system uses sophisticated analytical machine-listening models, as well as genetic algorithms to generate its improvisational memory, so it is categorized as level 4 with regard to its autonomy.

## 4.2 Cypher

The Cypher system was developed by Robert Rowe, and described as a 'real-time system with two major components: a player and a listener' (Rowe 1992, 43). Its listening module analyses musical features such as speed, density, dynamic, harmony and rhythm. These analytical capacities are considered single agents in a multi-agent architecture which then report to the player module, which uses 'various algorithmic techniques to produce musical output' (Rowe 1992, 43). The input for the system is a single stream of MIDI data, and the techniques used are mainly rule-based. The output of the player module is then fed into another system stage, called the 'compositional critic', which is there to 'monitor the output of the Cypher player and apply modification to the material generated by the composition methods before they are actually sent to the sound-making devices' (Rowe 1992, 46). Alongside its use with real-time MIDI data as its input, Cypher also represents a stand-alone generative system.

In their typology of metacreative music systems, Tatar and Pasquier describe Cypher as a symbolic, rule-based reactive musical agent (Tatar and Pasquier 2019, 70). Yet Rowe bases his analytical model for harmonies, which he calls 'Chordnet', on a neural network initially presented by Scarborough, Miller and Jones (1989, as cited in Rowe 1992, 51), implying that this system is *not* purely rule-based but implements some early notions of artificial intelligence. When discussing representation in interactive music systems in a recent article, Rowe concludes that the most promising approach to this design might reside in a combination of symbolic representation and the exposure of raw data to AI systems. He then cites Marcus for support: 'The right move today may be to integrate deep learning, which excels at perceptual classification, with symbolic systems, which excel at inference and abstraction' (Marcus 2018 as cited in Rowe 2021, 7).

## 4.3 Voyager

The Voyager system was developed by trombone player George Lewis in the 1990s, and is described in his article, 'Too Many Notes: Computers, Complexity and Culture in "Voyager"' (Lewis 2000). It was constructed to be a 'virtual improvising orchestra' (Lewis 2000, 33) that reacts to MIDI input data. However, Lewis used the system mainly with real-time microphone signals from his analogue trombone and other instruments which were then transferred to MIDI data. Voyager has 64 voices which react to the input data in a rule-based manner which follows the notion of 'Multidominance', which Lewis encountered in Douglas's analysis of Afro-American music. Douglas describes Multidominance in contrast to Western classical music as an aesthetic which involves 'the multiple use of colors in intense degrees, or the multiple use of textures, design patterns, or shapes' (Douglas 1991 as cited in Lewis 2000, 33). Lewis refracted this notion through his Voyager system so that it would return many parallel voices at the same time, often without the clear center of attention one expects in Western classical compositional practice. Lewis's cultural perspective is also clear in his 1999 article 'Interacting with Latter-Day Musical Automata', where he explores the historical

and aesthetic notions of performance with non-human agents. In the light of this article, it is clear that Lewis intended *Voyager* to be as autonomous as possible within the technological constraints of the late 1990s. Still, it is now considered a single-user, rule-based, and reactive multi-agent system which does not demonstrate advanced techniques from the field of artificial intelligence; see, for example, the analysis of Tatar and Pasquier (2019, 57).

#### 4.4 Reactable

The Reactable was developed as a multi-user instrument by Sergi Jordà, to explore shared collective control (Jordà 2005). It is a tabletop instrument with wooden dice-like objects which represent different musical threads, and their movement and rotation inform musical functions such as filters or generators. The instrument can be played by one to six people who are free to pick up and adjust the objects of their choice interchangeably. It is also possible to use multiple Reactables in different locations, resulting in a distributed, virtual live performance (*ibid.*). Xambó explored the affordances of tabletop interfaces in a number of studies with both expert players and beginners and concludes that this instrument group can facilitate the users' 'understanding of contemporary representations of music and provide them with tools and techniques for collaborative music creation' (Xambó 2014, 256).

The tabletop instrument implies both an embodied and a participatory approach. Players use their hands to replace and turn the objects, and they interact with the other players standing next to them in both their intimate and personal zones. In addition, the instrument accommodates both social and public zones, given that it is possible to connect more than one Reactable in the same space or even in separate venues via a digital network. However, the system does not create its own gestures but rather responds with predefined ones, which places it at the level of compositionality (level 2) on the autonomy scale.

#### 4.5 Echobo

Echobo was developed by Sang Won Lee and Jason Freeman as a performance system which invites audience members to participate via their personal smartphones. The instrument features two interfaces, one for a so-called master musician who controls chord progressions and one for audience members which provides a keyboard whose tones change based on the progressions selected in the master interface (Lee & Freeman 2013, 452). Echobo's performance concept also involves an acoustic instrumentalist who improvises while the performance is underway. Because the individual contributions of the audience members are all played at the same time, the result of this system's work is a sequence of advancing chord textures accompanied, in turn, by the instrumentalist. From a design point of view, the system is highly interactive and participatory, and it also takes an embodied approach, since all of its interactions are directly tied to audience members tapping on their smartphones. Still, the system only demonstrates a low level of autonomy, given that all the interactions go directly to a sound-producing module, and the system itself is only used to communicate these interactions.

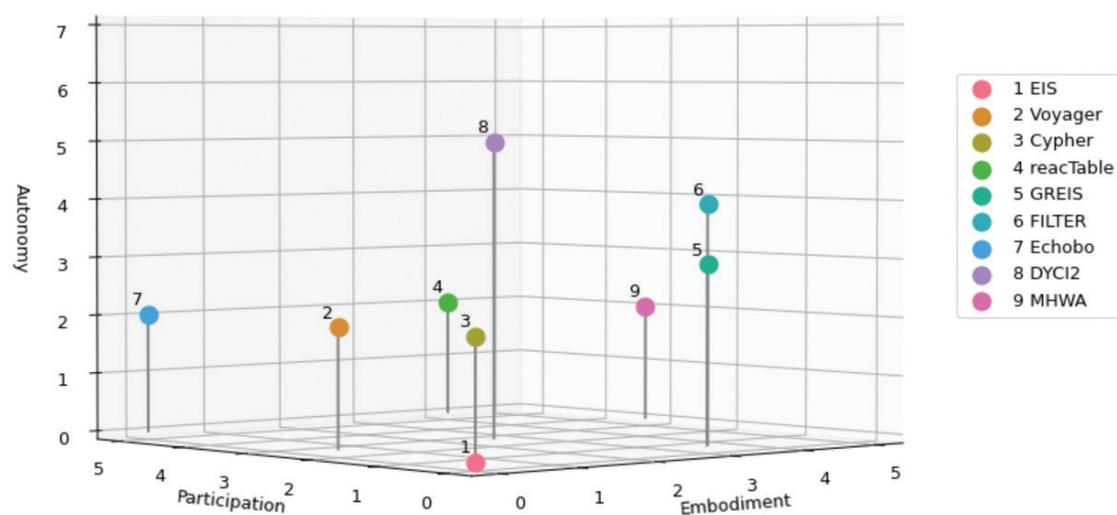
#### 4.6 DYCI2

DYCI2 (Creative Dynamics of Improvised Interaction) is a collaborative research and development project, funded by the French National Research Agency, which explores the creative dynamics of improvised interactions between human and artificial agents to inform the 'design of listening mechanisms, learning schemes and music generation processes of the creative agents' (Nika et al. 2017, paragraph 2). As a real-time co-improvising agent, DYCI2 takes a twofold approach by proposing 'reactive listening' and 'planning' processes which

are balanced according to the musical context and goals (Nika et al. 2017). The reactive listening is built on Somax, an offspring of OMax, which has been widely used for short-term listening and generation tasks in improvisational player agents (Bonnasse-Gahot 2014). In terms of the planning process, the system can be trained on stylistic responses with the use of musical materials in both audio and MIDI formats. The system has been extensively explored through co-play with expert musical performers, and has been able to initiate its own meaningful musical gestures in real-time improvisational scenarios, so it is considered a highly autonomous system. However, it does not use embodied gestures as inputs or use any input other than representational music data. It is also intended not as an interactive or participatory system but as a reactive companion, usually in use with one or a few acoustic human instrumentalists.

#### 4.7 MHWA

Turchet et al. have developed a system for musical haptic wearables for audiences (MHWA), alongside Turchet's work on tactile interaction technology and the use of wearable devices as additional displays for the audience (Turchet et al. 2020). The system has been used for two performances: an electronic live music performance with two DJs, and a solo mandolin performance. In both cases, audience members wore haptic devices around the upper torso which carry multiple vibro-tactile motors to synthesize the musical gestures of the performers into haptic stimuli via a rule-based methodology. Results of their studies show that participants were interested and reported that the augmentation of the musical experience via the haptic devices slightly enhanced their felt connection to the performers and their musical engagement, but, as the authors state, the system is prototypical in nature and requires further validation before a more meaningful implementation (Turchet et al. 2020). In addition, the system does not engage autonomous parts when displaying tactile stimuli to the listeners, but instead executes a set of hard-coded gesture-mappings. Still, the mandolin used in one of the cases introduces a certain level of autonomy when generating sound via the embedded sonic generative methods connected to the sensor placed on this acoustic instrument, which is why MHWA is assigned autonomy level 2. Lastly, the system is not participatory in the sense of allowing one player's gestures to impact another's, so it fails to afford any social interaction or co-creation beyond those of a traditional music performance in front of an audience. Ultimately, MHWA expands the listening experience for



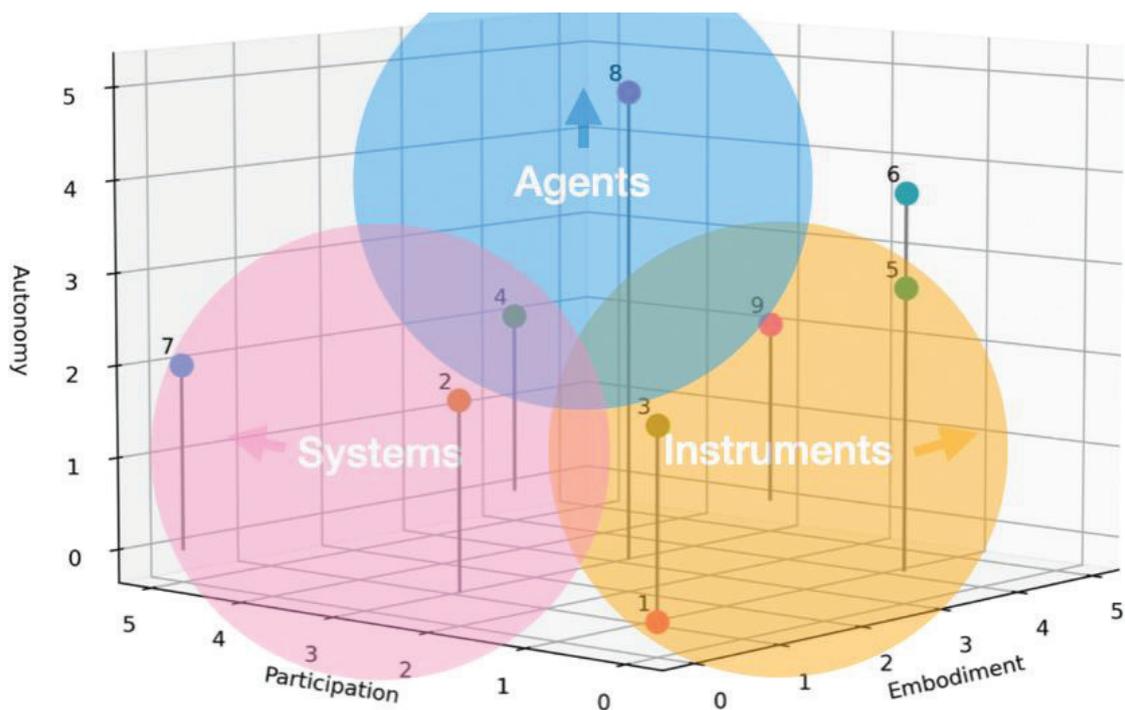
**Figure 5.** Analyzed cases in the IMPS design space.

audience members to the haptic domain and by doing so enables a more embodied experience. Certainly, the system impacts the intimate domain of the listeners, sometimes to an unpleasant degree, as some audience members reported (Turchet et al. 2020).

## 5 IMPS design as speculative design

This analysis of nine IMPS indicates a clear distribution of qualities across the IMPS model space. Earlier systems appear at the lower ends of the axes, as would be expected given that the technology was less advanced, and most IMPS are highly dependent on the technological affordances of their time. On the proxemic plane, a wide spread of combinations appears, but only a few systems (including Reactable and MHWa) offer both a high degree of embodied design and a high degree of interpersonal co-creation affordance. The highest degree of autonomy appeared in DYCI2, which implements highly refined models for machine listening and reactive behavior. Apparently, levels 6 and 7 of the taxonomy of Eigenfeldt et al. remain out of reach among the currently published IMPS. In all, the analysis of IMPS indicates that most of the cases were positioned towards the end of at least one of the three axes. This points to the fact that most systems focus their implementation on one of these three aspects and tend towards one of these categories: *instruments*, *systems* or *agents* (see Figure 6).

The proposed model and analysis are both limited in several ways. First, the model itself is somewhat reductive, given that it does not consider aspects of IMPS such as architecture, public availability, or system-code details, among others. In addition, the model's proxemic zones tend to confuse the issue of literal versus virtual 'space'. The design categories *embodiment* and *participation* refer to how people might interact with the systems both individually and together, and the model is intended to frame qualities of the systems which relate to these dimensions. The numerical quantization of these axes into six steps (0–5) is arranged to enable comparison with the autonomy axis, which uses eight levels (though only six of them are currently possible). Moreover, the proposed taxonomy for autonomous musical

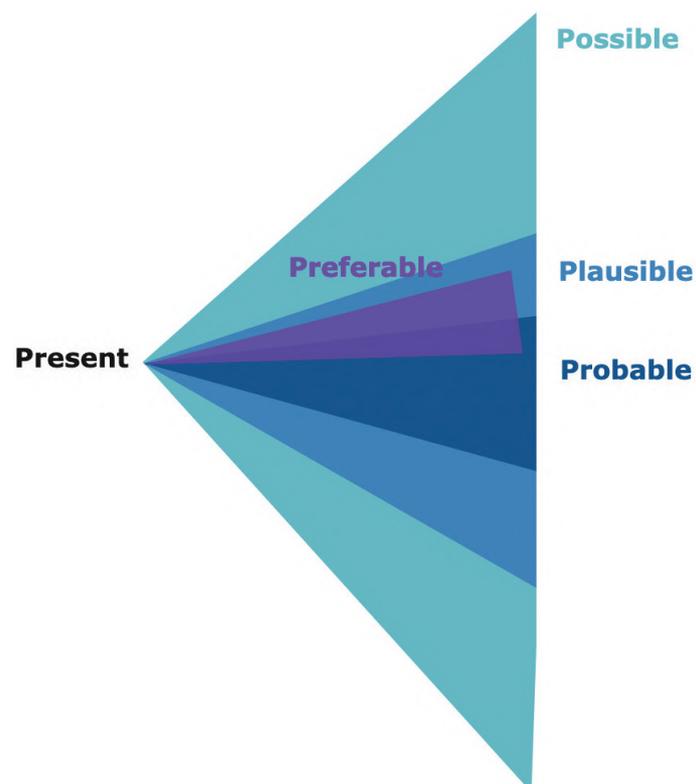


**Figure 6.** Main tendencies of IMPS shown in the model space.

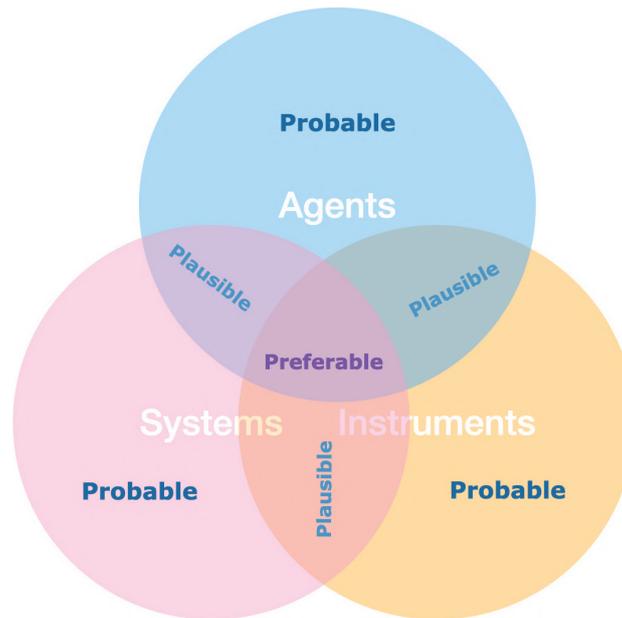
agents is limited by Eigenfeldt et al.'s model. Lastly, the analysis itself is confined to only a small number of IMPS known and accessible to the author as a means of testing the model.

IMPS design is an exploding field of development usefully viewed through the lens of speculative design. Auger suggests that 'speculative design serves two distinct purposes: first, to enable us to think about the future; second, to critique current practice' (Auger 2013, 11). He continues, 'to shift the discussion on technology beyond the fields of experts to a broad popular audience, the choice of "speculative" is preferable as it suggests a direct correlation between "here and now" and existence of the design concept' (Auger 2013, 12). This approach to the field of IMPS is particularly relevant regarding users of the systems who are not the creators, including general audiences in IMPS. Auger also introduces the concept of a 'perceptual bridge', which he describes as a 'careful management of the speculation; if it strays too far into the future to present implausible concepts or alien technological habitats, the audience will not relate to the proposal, resulting in a lack of engagement or connection. In effect, a design speculation requires a bridge to exist between the audience's perception of their world and the fictional element of the concept' (Auger 2013, 12). Lastly, his pledge to develop an 'ecological approach to speculative design' (Auger 2013, 13) resonates with the affordances of music technology, in turn recalling Gibson's ideas of perception (Gibson 1979).

Elsden takes the idea of speculative design a step further and suggests 'speculation as a form of Research through Design', arguing that 'Speculative Enactments constitute an effort to meaningfully enact elements of possible futures with participants' (Elsden et al. 2017, 5387). Grunwald draws upon a similar framework when discussing the prognostic, scenario-based and hermeneutic modes of orientation towards possible futures which emerge under the influence of or are induced by technology (Grunwald 2015, free translation by the author). Finally, Dunne and Raby diagram their notion of speculative design in their book *Speculative Everything*, and Figure 7 is adapted from their work, which was initially based on a lecture by Stuart Candy (as cited in Dunne and Raby 2013). The idea of possible



**Figure 7.** Speculative design, adapted from Dunne and Raby (2013, 5).



**Figure 8.** Speculative design applied to the tendencies of the model for IMPS design.

futures stems from the conviction that many of the challenges we face today are not fixable through design solutions alone but only ‘by changing our values, beliefs, attitudes, and behavior’ (Dunne and Raby 2013, 2). These futures include the possible, the plausible, the probable and the preferable. The probable, or what is likely to happen, is most closely aligned to common practice in design, whereas the plausible, or what could happen, informs current viable alternatives to design practice. The possible, on the other hand, tries to link today’s world to a speculative one by following already understood or at least imaginable paths (Dunne and Raby 2013, 4). Preferable futures are the most desirable outcomes to be reached by a discursive, collective approach (Dunne and Raby 2013, 6).

Possible futures in the context of the IMPS design model appear in the space of the maximum values along the three axes. Referring to the autonomy axis, for example, levels 6 and 7 already transcend current practice and invite design speculation. Probable designs reside towards the ends of the individual axes, since current practice dictates that IMPS are designed as either *instruments*, *systems*, or *agents*. Plausible designs reside at the intersections of these three tendencies, designating areas which are not widely used and therefore present alternatives to current design solutions. Preferable designs, then, might emerge at the center of the model, that is, the intersection of all three dimensions. The ongoing debate in this intersectional space is likely the most fruitful avenue for the design of preferable IMPS futures, and Figure 8 renders the idea visually. In this sense, the possible refers to the area inside of all the circles, and the ‘impossible’ to the white surrounding area, which might be explored by future IMPS designers.

## 6 Conclusion

After delimiting Intelligent Music Performance Systems (IMPS), this article integrated Hall’s theory of proxemics and Eigenfeldt et al.’s taxonomy of musical metacreation into a three-dimensional model of the main design aspects *embodiment*, *participation*, and *autonomy*. Nine cases of IMPS ranging from the 1950s to the present were reviewed with the help of the model to reveal that most systems tend towards one of these categories: *instruments*, *systems*, or *agents*. The article then introduced ideas from Speculative Design to demonstrate that preferable future cases of IMPS will arise at the intersection of these three categories.

## References

- Auger, James. 2013. 'Speculative Design: Crafting the Speculation'. *Digital Creativity* 24, no. 1: 11–35. <https://doi.org/10.1080/14626268.2013.767276>
- Birnbaum, David, Rebecca Fiebrink, Joseph Malloch and Marcelo M. Wanderley. 2005. 'Towards a Dimension Space for Musical Devices'. *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME05)*. Vancouver, BC, Canada. [https://doi.org/10.1007/978-3-319-47214-0\\_14](https://doi.org/10.1007/978-3-319-47214-0_14)
- Bongers, Bert. 2007. 'Electronic Musical Instruments: Experiences of a New Luthier'. *Leonardo Music Journal* 17: 9–16. <https://doi.org/10.1162/lmj.2007.17.9>
- Bonnasse-Gahot, Laurent. 2014. *An Update on the SOMax Project*. IRCAM – STMS, Tech. Rep., Post-doctoral report. <http://architexte.ircam.fr/textes/BonnasseGahot14a/index.pdf>
- Bourriaud, Nicolas. 2002. *Relational Aesthetics*. Les presses du réel.
- Bown, Oliver, and Andrew R. Brown. 2018. 'Interaction Design for Metacreative Systems'. In *New Directions in Third Wave Human–Computer Interaction: Volume 1 – Technologies*, edited by Michael Filimowicz and Veronika Tzankova, 67–87. Springer. [https://doi.org/10.1007/978-3-319-73356-2\\_5](https://doi.org/10.1007/978-3-319-73356-2_5)
- Collins, Karen 2008. *Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design*. MIT Press. <https://doi.org/10.7551/mitpress/7909.001.0001>
- Collins, Karen, and Ruth Dockwray. 2015. 'Sonic Proxemics and the Art of Persuasion: An Analytical Framework'. *Leonardo Music Journal* 25: 53–56. [https://doi.org/10.1162/lmj\\_a\\_00935](https://doi.org/10.1162/lmj_a_00935)
- Cook, Perry. 2001. 'Principles for Designing Computer Music Controllers'. *Proceedings of CHI 2001*, Extended Abstracts: 491–492. ACM.
- Dahlstedt, Palle. 2021. 'Musicking with Algorithms: Thoughts on Artificial Intelligence, Creativity, and Agency'. In *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*, edited by Eduardo R. Miranda, 873–914. Springer.
- Dockwray, Ruth. 2016. 'Proxemic Interaction in Popular Music Recordings'. In *Mixing Music*, edited by Russ Hepworth-Sawyer and Jay Hodgson, 73–81. Routledge. <https://doi.org/10.4324/9781315646602-14>
- Douglas, Robert L. 1991. 'Formalizing an Afro-American Aesthetic'. *New Art Examiner* (June/Summer): 18–24.
- Drummond, Jon. 2009. 'Understanding Interactive Systems'. *Organised Sound* 14, no. 2: 124–133. <https://doi.org/10.1017/s1355771809000235>
- Dunne, Anthony, and Fiona Raby. 2013. *Speculative Everything*. MIT Press.
- Eigenfeldt, Arne, Oliver Bown, Philippe Pasquier, and Aengus Martin. 2013. 'Towards a Taxonomy of Musical Metacreation: Reflections on the First Musical Metacreation Weekend'. *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*. Sydney, Australia. <https://doi.org/10.1609/aiide.v9i5.12647>
- Elsden, Chris, David Chatting, Abigail C. Durrant, Andrew Garbett, Bettina Nissen, John Vines and David S. Kirk. 2017. 'On Speculative Enactments'. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 5386–5399. <https://doi.org/10.1145/3025453.3025503>
- Freeman, J. 2005. 'Glimmer: For Chamber Orchestra and Audience'. PhD diss., Columbia University.
- Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates.
- Gifford, Toby, Shelly Knotts, Jon McCormack, Stefano Kalonaris, Mathew Yee-King and Mark d'Inverno. 2018. 'Computational Systems for Music Improvisation'. *Digital Creativity* 29, no. 1: 19–36. <https://doi.org/10.1080/14626268.2018.1426613>
- Glăveanu, Vlad P. 2014. *Distributed Creativity: Thinking Outside the Box of the Creative Individual*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-05434-6>
- Grunwald, Armin. 2015. 'Die hermeneutische Erweiterung der Technikfolgenabschätzung. Technikfolgenabschätzung'. *Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis* 24, no. 2: 65–69. <https://doi.org/10.14512/tatup.24.2.65>
- Hall, Edward T. 1966. *The Hidden Dimension*. Doubleday.
- Hoedl, Oliver, Christoph Bartmann, Fares Kayali, Christoph Loew, and Peter Purgathofer. 2020. 'Large-Scale Audience Participation in Live Music Using Smartphones'. *Journal of New Music Research*. <https://doi.org/10.1080/09298215.2020.1722181>

- Hoedl, Oliver, Geraldine Fitzpatrick, Fares Kayali and Simon Holland 2017. 'Design Implications for Technology-Mediated Audience Participation in Live Music.' *14th Sound and Music Computing Conference SMC 2017*.
- Kantosalo, Anna, and Tapio Takala. 2020. 'Five C's for Human-Computer Co-Creativity: An Update on Classical Creativity Perspectives.' *Proceedings of the 11th International Conference on Computational Creativity*, 17–24. Association for Computational Creativity.
- King, Elaine, and Anthony Gritten. 2017. 'Dialogue and Beyond'. In *Musicians in the Making: Pathways to Creative Performance*, edited by J. S. Rink, H. Gaunt and A. Williamon, 306–321. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199346677.003.0022>
- Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford University Press.
- Lee, Sang Won, and Jason Freeman. 2013. 'Echobo: A Mobile Music Instrument Designed for Audience to Play'. *Proceedings of the International Conference on New Interfaces for Musical Expressio*: 450–455.
- Leman, Marc. 2008. *Embodied Music Cognition and Mediation Technology*. MIT Press. <https://doi.org/10.7551/mitpress/7476.001.0001>
- Levin, Golan, Scott Gibbons, Gregory Shakar, Yasmin Sohrawardy, Joris Gruber, Joerg Lehner, Gunther Schmid, and Erich Smlak. 2001. *Dialtones (A Telesymphony): Final Report*. [http://www.flong.com/storage/pdf/reports/dialtones\\_report.pdf](http://www.flong.com/storage/pdf/reports/dialtones_report.pdf)
- Lewis, George E. 2000. 'Too Many Notes: Computers, Complexity and Culture in Voyager'. *Leonardo Music Journal* 10: 33–39. <https://doi.org/10.1162/096112100570585>
- Llano, Maria T., Mark d'Inverno, Matthew Yee-King, Jon McCormack, Alon Ilsar, Alison Pease and Simon Colton. 2020. 'Explainable Computational Creativity'. *Proceedings of the Eleventh International Conference on Computational Creativity*. ICCG.
- Lorway, Norah, Edward Powley, and Arthur Wilson. 2021. 'Autopia: An AI Collaborator for Live Networked Computer Music Performance'. *Proceedings of the 2nd Conference on AI Music Creativity*. <http://repository.falmouth.ac.uk/4307>
- Marcus, Gary. 2018. *Deep Learning: A Critical Appraisal*. arXiv 2018, arXiv:1801.00631.
- Matuszewski, Benjamin, Norbert Schnell, and Frederic Bevilacqua. 2019. 'Interaction Topologies in Mobile-Based Situated Networked Music Systems'. *Wireless Communications and Mobile Computing*: 1–9. <https://doi.org/10.1155/2019/9142490>
- Mazzanti, Dario, Victor Zappi, Darwin G. Caldwell, and Andrea Brogni. 2014. 'Augmented Stage for Participatory Performances'. *Proceedings of the International Conference on New Interfaces for Musical Expression*: 29–34.
- McCormack, Jon, Patrick Hutchings, Toby Gifford, Matthew Yee-King, Maria T. Llano, and Mark d'Inverno. 2020. 'Design Considerations for Real-Time Collaboration with Creative Artificial Intelligence'. *Organised Sound* 25, no. 1: 41–52. <https://doi.org/10.1017/s1355771819000451>
- Miranda, Eduardo R., and Marcelo M. Wanderley. 2006. *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*. The Computer Music and Digital Audio Series, vol. 21. A-R Editions.
- Mueller, Florian F., Pedro Lopes, Paul Strohmeier, Wendy Ju, Caitlyn Seim, Martin Weigel, Suranga Nanayakkara, Marianna Obrist, Zhuying Li and Joseph Delfa et al. 2020. 'Next Steps for Human-Computer Integration'. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*: 1–15. ACM. <https://doi.org/10.1145/3313831.3376242>
- Nika, Jérôme, Ken Déguernel, Axel Chemla, Emmanuel Vincent, and Gérard Assayag. 2017. 'DYCI2 Agents: Merging the "Free", "Reactive", and "Scenario-Based" Music Generation Paradigms'. *International Computer Music Conference* (Oct. 2017). hal-01583089.
- Oliveros, Pauline (1979). *Software for People*. Center for Music Experiment and Related Research, University of California at San Diego.
- Oliveros, Pauline (2005). *Deep Listening: A Composer's Sound Practice*. IUiverse.

- Oliveros, Pauline (2008). 'The Expanded Instrument System: Introduction and Brief History'. In *The Future of Creative Technologies*, edited by Andrew Hugill and Tracey Harwood, 21–23. Journal of the Institute of Creative Technologies, De Montfort University.
- Pasquier, Philippe, Arne Eigenfeldt, Oliver Bown and Shlomo Dubnov. 2016. 'An Introduction to Musical Metacreation'. *Computers in Entertainment* 14, no. 2: 1–14. <https://doi.org/10.1145/2930672>
- Pressing, Jeff. 1990. 'Cybernetic Issues in Interactive Performance Systems'. *Computer Music Journal* 14, no. 1, New Performance Interfaces 1 (Spring): 12–25. <https://doi.org/10.2307/3680113>
- Renwick, Robin C. 2017. 'Topologies for Network Music'. PhD diss., Queen's University Belfast.
- Rowe, Robert. 1992. 'Machine Listening and Composing with Cypher'. *Computer Music Journal* 16, no. 1: 43–63. <https://doi.org/10.2307/3680494>
- Rowe, Robert. 1993. *Interactive Music Systems: Machine Listening and Composing*. MIT Press.
- Rowe, Robert. 2021. 'Representations, Affordances, and Interactive Systems'. In *Multimodal Technologies and Interaction* 5, no. 23. <https://doi.org/10.3390/mti5050023>
- Scarborough, Don L., Ben O. Miller and Jacqueline A. Jones. 1989. 'Connectionist Models for Tonal Analysis'. *Computer Music Journal* 13, no. 3: 49–55. <https://doi.org/10.2307/3680011>
- Tatar, Kıvanç, and Philippe Pasquier. 2019. 'Musical Agents: A Typology and State of the Art Towards Musical Metacreation'. *Journal of New Music Research* 48, no. 1: 56–105. Routledge. <https://doi.org/10.1080/09298215.2018.1511736>
- Turchet, Luca, Travis West, and Marcelo M. Wanderley. 2020. 'Touching the Audience: Musical Haptic Wearables for Augmented and Participatory Live Music Performances'. *Personal and Ubiquitous Computing* 25, no. 4: 749–769. <https://doi.org/10.1007/s00779-020-01395-2>
- Van Nort, Doug, Pauline Oliveros and Jonas Braasch. 2013. 'Electro/Acoustic Improvisation and Deeply Listening Machines'. *Journal of New Music Research* 42, no. 4: 303–324. <https://doi.org/10.1080/09298215.2013.860465>
- Visell, Yon, Roderick Murray-Smith, Stephen A. Brewster, and John Williamson. 2013. 'Continuous Auditory and Tactile Interaction Design'. In *Sonic Interaction Design*, edited by Karmen Franinović and Stefania Serafin. MIT Press. <https://doi.org/10.7551/mitpress/8555.001.0001>
- Weinberg, Gil. 2005. 'Interconnected Musical Networks: Toward a Theoretical Framework'. *Computer Music Journal* 29, no. 2: 23–39. <https://doi.org/10.1162/0148926054094350>
- West, Travis, Baptiste Caramiaux, Stéphane Huot and Marcelo M. Wanderley. 2021. 'Making Mappings: Design Criteria for Live Performance'. *Proceedings of the International Conference on New Interfaces for Musical Expression*. <https://doi.org/10.21428/92fbeb44.04f0fc35>
- Winkler, Todd. 1998. *Composing Interactive Music: Techniques and Ideas Using Max*. MIT Press.
- Wooldridge, Michael. 2009. *An Introduction to Multiagent Systems*. John Wiley and Sons.
- Wu, Yongmeng, Leshao Zhang, Nick Bryan-Kinns and Mathieu Barthet. 2017. 'Open Symphony: Creative Participation for Audiences of Live Music Performances'. *IEEE MultiMedia* 24, no. 1: 48–62. <https://doi.org/10.1109/mmul.2017.19>
- Xambó, Anna. 2014. 'Tabletop Tangible Interfaces for Music Performance: Design and Evaluation'. PhD diss., The Open University, UK.
- Xambó, Anna, and Gerard Roma. 2020. 'Performing Audiences: Composition Strategies for Network Music using Mobile Phones'. *Proceedings of the International Conference on New Interfaces for Musical Expression*: 55–60. Birmingham City University. <http://doi.org/10.5281/zenodo.4813192>
- York, Adrian. 2019. 'Transforming Musical Performance: The Audience as Performer'. In *Innovation in Music: Performance, Production, Technology, and Business*, edited by R. Hepworth-Sawyer, Jay Hodgson, Justin Paterson, and Rob Toulson, 3–23. Routledge. <https://doi.org/10.4324/9781351016711-1>

## Article 4 - Beyond Mutation: How Can We Acknowledge Symbiogenesis in Evolutionary Music Coding?

### Abstract

*This paper proposes to embrace alternative concepts for genetic coding that focus on the idea of symbioses and symbiogenesis as an addition to and refinement of currently used genetic algorithms for music creation. Those typically use Darwinist principles of natural selection implemented in a fitness function as a central principle to generate new individuals. It is suggested to consider symbiotic concepts for the context of music generation with a focus on participatory performance systems. The paper finally suggest four early-stage ideas as starting points for a discussion about potential implementation of these symbiotic realities in evolutionary-inspired music coding.*

Keywords: Genetic Programming, Symbiosis, Symbiogenesis

Version: published article

# Beyond Mutation: How Can We Acknowledge Symbiogenesis in Evolutionary Music Coding?

Corresponding author address: matthias.jung@uia.no

Submitted in June 2023

## Abstract

This paper proposes to embrace alternative concepts for genetic coding that focus on the idea of symbioses and symbiogenesis as an addition to and refinement of currently used genetic algorithms for music creation. Those typically use Darwinist principles of natural selection implemented in a fitness function as a central principle to generate new individuals. It is suggested to consider symbiotic concepts for the context of music generation with a focus on participatory performance systems. The paper finally suggest four early-stage ideas as starting points for a discussion about potential implementation of these symbiotic realities in evolutionary-inspired music coding.

## 1 Introduction

### 1.1 Genetic algorithms and evolutionary computation

Genetic algorithms (GA) have been used for solving optimization problems since the 1960s and are mainly inspired by Darwin's evolution theory that is based on natural selection processes [Darwin \(1859\)](#). Typically, they start with a set of initial solutions, that are recombined via a crossover operator and then mutated. The best solutions are then kept for future iterations via a fitness function, very much like Darwin's idea of the survival of the fittest. Kramer argues that today's variants of biology-inspired algorithms evolved along four main historical threads which explains the difference of genetic algorithms ([Kramer, 2017](#), p. 6). In his view, only one of those variant can still be distinguished from the remaining three, which is genetic programming (GP), which evolves machine learning techniques and differs in terms of their representation ([Kramer, 2017](#), p. 7).

### 1.2 Evolutionary computation in art and music

During the last three decades, genetic algorithms have faced a more frequent use for generating art and music. Starting in the 1980s, evolutionary art has been used for generating pictures [Romero and Machado \(2014\)](#). For the field of music, Goldberg is considered one of the first to present a genetic algorithm for composition tasks in the early 1990s [Horner and Goldberg \(1991\)](#). The field emerged throughout the 2000s and the diversity of advances can be observed in Miranda's anthology *Evolutionary Computer Music* [Miranda and Al Biles \(2007\)](#). More recently, Loughran and O'Neil reviewed applications of evolutionary computation to music composition tasks and categorized 47 genetically inspired approaches from the years 1998 to 2018 [Loughran and O'Neill \(2020\)](#). Specifically, they observed the musical problem, the type of evolutionary algorithm, the musical representation, and the fitness function used in each of the projects ([Loughran and O'Neill, 2020](#), p. 62). In the context of music, fitness functions are often created as interactive interfaces, that let humans evaluate the musical output in real-time.

### 1.3 Symbiotic relationships and symbiogenesis

In her book *Symbiotic planet*, Lynn Margulis summarized her arguments for a symbiotic view on evolution alternative to the selection-focused, Darwinist theories, and attempted to connect her symbiotic theories to concepts of planetary homeostasis. This so-called Gaia-theory proposes

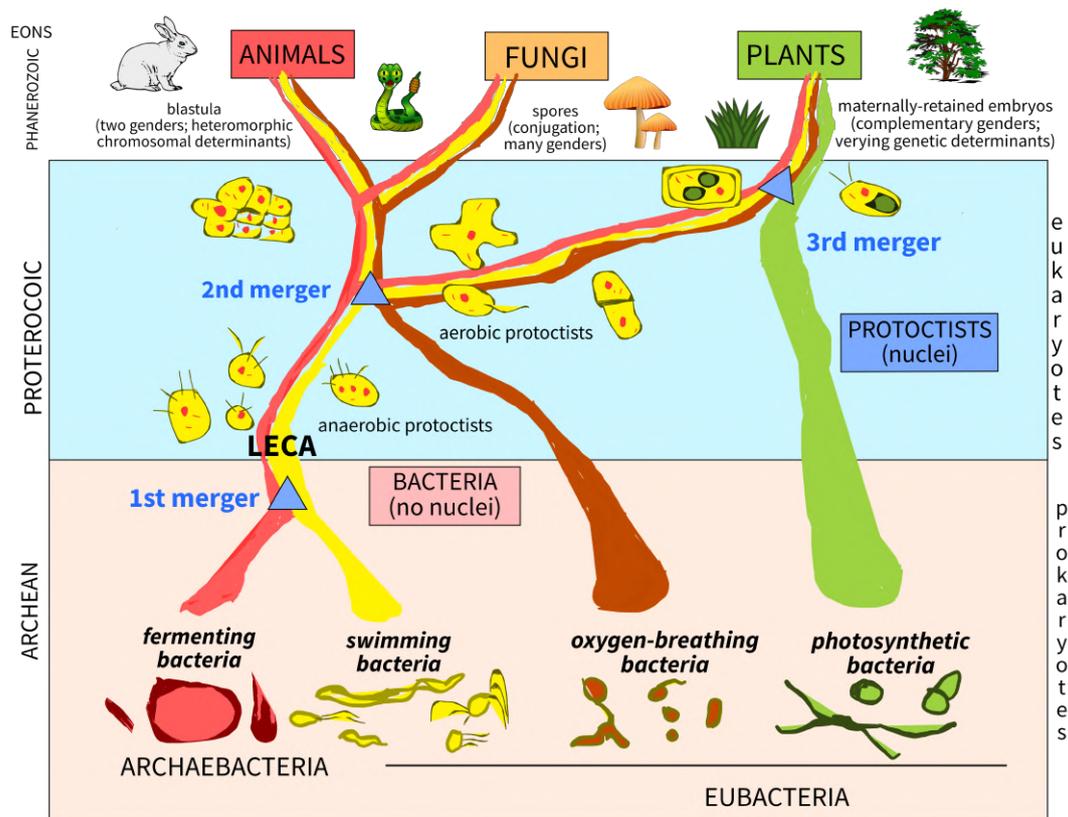


Figure 1: Visualization of symbiogenesis adapted from [Kozo-Polyansky \(2010\)](#), created with sketch.io

that life on the planet is not only feeding from planetary resources, but a necessity for upholding the earthly environment that in turn creates the basis for life [Margulis \(1998\)](#). Margulis argued her lifetime for symbiosis being a core evolutionary driver tightly intertwined with environmental adaption and natural selection processes, and her serial endosymbiosis theory is today widely accepted. As Sapp puts it in the Introduction to his book *Evolution by association: a history of symbiosis*, our evolution "is not due solely to the accumulation of gene changes within species. In fact, we evolved from, and are comprised of, a merger of two or more different kinds of organisms living together. Symbioses is at the very root of our being" [Sapp \(1994\)](#).

The role that symbioses played during evolution remains a matter of academic debate, as a discussion between Margulis and evolutionary biologist Dawkins at Oxford showed. Responding to Dawkins critique of symbiotic relationships in the process of selection pressure "(...) why on earth would you want to drag in symbiogenesis when it's (...) unparsimonious, uneconomical?", Margulis responded "Because it's there." ([MacAllister, 2009](#), p. 38). An in-depth discussion of alternative evolutionary concepts can also be found in [Ryan \(2002\)](#). Ultimately, it seems the role that symbiosis has played and continues to play in evolution is a matter of focus on processes that both have been proven to be apparent: natural selection and symbiogenesis both generating evolutionary novelty. Fig. 1 shows a visualization of the concept of symbiogenesis.

Three different forms of symbioses are commonly differentiated: parasitism, mutualism, and commensalism. Commensal symbiotic associations refer to animals of different species that share the food caught by the other animal. The term is used in a broader sense where the "benefit to one of the symbionts may be nutritional or protective" ([Paracer and Ahmadjian, 2000](#), p. 6). In a mutual symbiosis, both partners benefit from the relationship and there is a reciprocal exchange of nutrients or protection ([Paracer and Ahmadjian, 2000](#), p. 6). Parasitism, however, is "a symbiosis in which one of the symbionts benefits at the expense of the other" ([Paracer and Ahmadjian, 2000](#), p. 7).

Reflecting the theory of symbiogenesis, the question arises whether these realities are acknowledged in current genetic programming for music generation, and what could be suitable concepts to im-

plement those symbiotic principles? Is there a beneficial way to simulate short and longer-term symbiotic relationships that are helpful for musical applications? How can we combine interdependent populations in a way that helps the creative outcome? And what are the problems that an interspecies-level concept can solve that intraspecies concepts cannot? Finally, how can it be helpful to integrate environmentally inspired concepts such as the Gaia theory that seem to be also connected to symbiotic life?

To start addressing those questions, the paper will first review some related work, before four suggestions are made for applying the biological concept of symbiosis for music generation, specifically for live performance context and with audience participation in mind. These suggestions are at an early idea stage, which is why they are presented together with further questions. Finally, these concepts will be further discussed and the paper concludes with some limitations and points to further work planned around the implementation of symbiosis in genetic coding for music.

## 2 Related work

Genetic algorithms can be seen in a multitude of variations, and they are closely related to the concept of multi-agent systems for music generation [Wooldridge \(2009\)](#). Tatar and Pasquier reviewed 78 systems, many of which use evolutionary computing as their architecture [Tatar and Pasquier \(2019\)](#). However, multi-agent systems typically use individual agents in the sense of simulating (human) musical behaviour and interspecies symbiosis has not been observed in those simulations. The same accounts for the use of swarm intelligence that are built on individuals that co-evolve and affect each other rather than focussing on their interconnection via 2nd order species in symbiosis [Borgo \(2006\)](#) [Bonabeau et al. \(1999\)](#) .

Dahlstedt and Nordahl created a world of co-evolving creatures that recreate based on a genetic algorithm reproducing single children. The creatures have listening capacities and can be seen as a multi-agent system, that creates music as a result of their individuals that work together, trigger each other, and create a small part each ([Dahlstedt and Nordahl, 2001](#), p. 247). Strictly speaking this concept does not involve symbiotic relationships, even if the environment balances all consumed and produced energy levels, that the creatures take in and digest ([Dahlstedt and Nordahl, 2001](#), p. 244), and that is due to the fact that the single creatures do neither share food or benefit in other symbiotic ways.

With their system *Locus Diffuse*, Hoy and Van Nort created a multi-user instrument that incorporates the food-searching behavior of slime mold [Hoy and Van Nort \(2021\)](#). The work is inspired by both computational simulation of slime mold behavior [Adamatzky \(2010\)](#), as well as real-life application of using slime mold as a sound-generating source [Miranda et al. \(2011\)](#). The system is an interesting approach, since it is situated "at the crossroads of sonic ecosystem design, agent-based musical systems, multi-user instruments, and networked performance" ([Hoy and Van Nort, 2021](#), p. 21).

Implementing the idea of symbiotic relationships, Hirasawa et al. found that their Genetic Symbiosis Algorithm (GSA), that took into consideration various mutual symbiotic relations between individuals was able to find more flexible solutions compared to conventional methods [Hirasawa et al. \(2000\)](#). They introduced a symbiotic parameter to their genetic algorithm, that was used after the crossover and mutation stages before the fitness function. Their approach however is limited to intraspecies symbiotic mutualism and did not consider other forms of interspecies symbioses such as parasitism or commensalism or interspecies mutualism.

On a rather analytical level, Moriaty uses the concept of symbiosis for theorizing the interaction between musical performers and dancers. She maps the different kinds of symbiotic relationships to the ones that are in place between dancers and musicians when performing together on a stage [Moriaty \(2020\)](#). The three different symbiotic forms (parasitism, mutualism, and commensalism) are used as taxa describing interactive relationships between dancers and musicians. This biological metaphor offers a valuable new perspective upon the performative relationships, however are they not sonified or used in other generative ways ([Moriaty, 2020](#), p. 552).

Lastly, Lorway's participatory performance system *Autopia* shall be mentioned, which uses an interactive fitness function that is controlled by audience members [Lorway et al. \(2021\)](#). Her melody-focussed live coding approach controls an initial population of sine waves that are consecutively shaped by a gamified audience voting rationale.

### 3 Suggestions for introducing symbiosis in genetic algorithms for music

Moving forward, some early-stage ideas for strategies to implement ideas from symbiogenesis to evolutionary coding are attempted. This is specifically done with a performance-based scenario in mind, in which the audience is recognized as a participating entity.

#### **Suggestion 1: Representing quantities of audience participation as mitochondria**

Mitochondria have evolved as the energy-generating symbiotic bacteria in eukaryotes during the evolutionary process [Sapp \(1994\)](#). They are equipped with their own DNA and could be conceptualized as symbionts representing the participatory energy of the users of a multi-user instrument. Instead of representing musical elements such as tones, melodies, rhythm, or harmony, the quantities of participation generated from the server data in networked instruments could be seen as a measure for the overall musical activity represented by a separate, but connected system of mitochondria present in all individuals of a population in the genetic algorithm. This idea could be extended to other holobionts (hosts and their microbiota) [Guerrero et al. \(2013\)](#), in the sense that the microbiota affect the host collective on an inter- and supra-individual level, which affects the evolution of a population over time.

#### **Suggestion 2: Symbiotic collaboration as part of the mutation operator**

Inspired by mathematical models for cellular interaction as described in [Lindenmayer \(1968\)](#), symbiotic collaboration could also be implemented as part of the mutation function. Returning to the example of mitochondria, which are passed on maternally with the egg, evolving microbiota in eukaryotes could be implemented as a separate mutation operator that reflects symbiotic relationships and evolves alongside the separate genome of the nuclei. Those symbioses could also vary in their stability over time and represent both short and long-term symbioses, that represent more stable collaborations between species.

#### **Suggestion 3: The eukaryotic cell as a metaphor for musical performance**

Rather than simulating symbiogenetic processes, the eukaryotic cell could itself be taken as a metaphor for simulating interactive concert situations. The nucleus of such cells could represent the staged musical performers, that also hold the DNA of such an event. Audience members in such a scenario could be seen as (different) microbiota, living symbiotically within these cells. The cell constituting the closed environment such as a concert hall or other type of venue shielded from the outside world, could represent the cell membrane.

#### **Suggestion 4: Appreciate environmental conditioning (Gaia) within the fitness function**

Following Margulis' attempt to relate symbiogenesis to the planetary environment [Margulis \(1998\)](#), the evolution of populations could also influence the surrounding resources that in turn enable and influence its genome. In that sense, the symbiotic relationship affects the homeostatic of the environment creating a circular loop between evolutionary processes of the population and their habitats. A further exploration of the Gaia hypothesis as suggested by Lovelock and Margulis [Lovelock and Margulis \(1974\)](#) is considered beneficial.

## 4 Conclusion

The paper suggested to review symbiotic associations as a biological metaphor and approach to extending currently used genetic algorithms and evolutionary programming strategies that are often inspired by Darwinist evolution biology. It reviewed related work in the field of multi-agent systems, swarm algorithms and symbiosis-inspired approaches, and found that there many of the suggested works are in their core based on individualistic and selection-based concepts. It was

argued that an attempt to surpass these concepts in the sense that genetic programming for music generation might take advantage of symbiotic concepts in a more profound manner. Symbiogenesis in evolutionary computation is neither acknowledged by refined mutation operators, nor is it a form of collective intelligence typically represented by swarms or as part of multi-agent systems. Symbioses are interspecies relationships and collaborations that can spread across individuals and form mergers of more or less stable durability. This is why it is thought to be beneficial to discuss how we might implement this type of biological reality within musical systems.

## References

- Adamatzky, A. (2010). *Physarum machines: computers from slime mould*, volume 74. World Scientific.
- Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999). *Swarm intelligence: from natural to artificial systems*. Number 1. Oxford University Press.
- Borgo, D. (2006). Sync or swarm: Musical improvisation and the complex dynamics of group creativity. In *Algebra, Meaning, and Computation: Essays Dedicated to Joseph A. Goguen on the Occasion of His 65th Birthday*, pages 1–24. Springer.
- Dahlstedt, P. and Nordahl, M. G. (2001). Living Melodies: Coevolution of Sonic Communication. *Leonardo*, 34(3):243–248.
- Darwin, C. (1859). *On the Origin of Species*. John Murray, London.
- Guerrero, R., Margulis, L., Berlanga, M., et al. (2013). Symbiogenesis: The holobiont as a unit of evolution. *International Microbiology*, 16(3):133–143.
- Hirasawa, K., Ishikawa, Y., Hu, J., Murata, J., and Mao, J. (2000). Genetic symbiosis algorithm. In *Proceedings of the 2000 Congress on Evolutionary Computation. CEC00 (Cat. No.00TH8512)*, volume 2, pages 1377–1384 vol.2.
- Horner, A. and Goldberg, D. E. (1991). Genetic algorithms and computer-assisted music composition. *Urbana (Caracas, Venezuela)*, 51.
- Hoy, R. and Van Nort, D. (2021). Locus diffuse: An agent-based sonic ecosystem for collaborative musical play. *CMMR*, page 21.
- Kozo-Polyansky, B. M. (2010). *Symbiogenesis: a new principle of evolution*. Harvard University Press.
- Kramer, O. (2017). *Genetic Algorithms*, pages 11–19. Springer International Publishing, Cham.
- Lindenmayer, A. (1968). Mathematical models for cellular interactions in development: I. filaments with one-sided inputs. *Journal of Theoretical Biology*, 18(3):280–299.
- Lorway, N., Powley, E., and Wilson, A. (2021). Autopia: An AI collaborator for live networked computer music performance. *Proceedings of the 2nd Conference on AI Music Creativity (MuMe + CSMC)*. accessed May 15, 2023.
- Loughran, R. and O’Neill, M. (2020). Evolutionary Music: Applying Evolutionary Computation to the Art of Creating Music. *Genetic Programming and Evolvable Machines*, 21:55–85.
- Lovelock, J. E. and Margulis, L. (1974). Atmospheric homeostasis by and for the biosphere: The gaia hypothesis. *Tellus*, 26(1-2):2–10.
- MacAllister, J. D. (2009). A commentary on homage to Darwin debate.
- Margulis, L. (1998). *Symbiotic Planet: A New Look at Evolution*. Basic Books.
- Miranda, E. R., Adamatzky, A., and Jones, J. (2011). Sounds synthesis with slime mould of physarum polycephalum. *Journal of Bionic Engineering*, 8(2):107–113.
- Miranda, E. R. and Al Biles, J. (2007). *Evolutionary Computer Music*. Springer.

- Moriarty, M. (2020). Symbiosis: a biological taxonomy for modes of interaction in dance-music collaborations. In *Proceedings of the International Conference on New Interfaces for Musical Expression, Computer Music Journal*, pages 551–556.
- Paracer, S. and Ahmadjian, V. (2000). *Symbiosis: An introduction to biological associations*. Oxford University Press.
- Romero, J. and Machado, P. (2014). *The Art of Artificial Evolution: A Handbook on Evolutionary Art and Music*. Natural Computing Series. Springer Berlin Heidelberg.
- Ryan, F. (2002). *Darwin's blind spot: Evolution beyond natural selection*. Houghton Mifflin Harcourt.
- Sapp, J. (1994). *Evolution by association: A history of symbiosis*. Oxford University Press.
- Tatar, K. and Pasquier, P. (2019). Musical Agents: A Typology and State of the Art Towards Musical Metacreation. *Journal of New Music Research*, 48(1):56–105.
- Wooldridge, M. (2009). *An Introduction to Multiagent Systems*. John Wiley & Sons.

# Appended Questionnaires and Interview Guides

### Audience Questionnaire for Participatory Concert

*Thank you for taking part in our audience participation experiment and this related self-report - we appreciate it! Answering these questions will take about 10-15mins time and we encourage you to be as honest and precise as possible. The answers will be anonymised and treated with confidentiality as stated in the information letter.*

Please fill in the anonymous participation number that you used for the experiment: \_\_\_\_\_

#### Personal background

1. What is your gender?

- a) Woman
- b) Man
- c) Diverse
- d) Choose not to identify

2. What is your age? \_\_\_\_\_

3. What is your nationality? \_\_\_\_\_

4. What is the highest formal degree you received during your education so far?

- a) Primary or lower secondary school
- b) Upper secondary school
- c) Professional degree
- d) College without degree
- e) Bachelor's degree or similar
- f) Master's degree or similar
- g) Doctoral degree
- h) Choose not to say

5. What kind of musical training do you have?

*Formal training refers to institutional learning (e.g. in music-schools or universities), non-formal refers to learning outside of institutions.*

- a) No formal or informal musical training
- b) Most of my musical training is self-taught or informal
- c) Less than 1 year of formal musical training
- d) Between 1-5 years of formal musical training
- f) More than 5 years of formal musical training
- g) Choose not to say

6. Which title describes you best?

- a) Tone-deaf
- b) Tone-deaf music-loving non-musician
- c) Tone-deaf amateur or professional musician
- d) Non-musician
- e) Music-loving non-musician
- f) Amateur musician
- g) Serious amateur musicians
- h) Semiprofessional musician
- i) Professional musician
- j) Choose not to say

7. Please check all boxes that apply to you.

- a) I sing or make music just for myself (e.g. sing in the shower or practice individually)
- b) I sing or make music together with others (e.g. sing in the choir or play in a band)
- c) I make music digitally (e.g. create music or instruments with software or engage in live-coding)
- d) Choose not to say

Personality traits

The following questions use a scale from 1 to 7 that relates to how much you agree with each statement. Please respond to the questions individually, that is as if they weren't connected. Don't worry about not being consistent in your answers. If you can't or don't want to answer to one of the questions, please leave it blank.

- |                        |                       |                       |                       |                       |                       |                       |
|------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| <input type="radio"/>  | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 1-strongly<br>disagree | 2-disagree            | 3-slightly<br>agree   | 4-neutral             | 5-slightly<br>agree   | 6-agree               | 7-strongly<br>agree   |

8. I see myself as a confident person among strangers
9. I enjoy being in public spaces.
10. I enjoy getting in touch with new people.
11. I am generally open to new experiences.
12. I would describe myself as a critical person.
13. I consider myself as experienced in creative practices (not limited to music or art).

Concert and listening habits

If you haven't been able to go to concerts lately due to the Covid19 pandemic, please refer to a period where your behaviour wasn't compromised.

14. If possible I enjoy going to live concerts.
15. I usually go to concerts with a group of friends.
16. Music makes me bond with other people.
17. At a concert I generally feel connected to the musicians and performers on stage.
18. At a concert I feel connected to other people in the audience.
19. If I enjoy the music at a concert and the situation allows I like to move and dance to it.
20. Generally I listen to music a lot.
21. I often feel emotional response when listening to music.
22. I use music to control my mood or change my emotional condition.

Digital literacy, genre familiarity and valence

23. I would consider myself as experienced with digital technology.
24. I generally enjoy using my smartphone.
25. I enjoy playing games (not limited to computer games).
26. I enjoy the competitive aspect of playing games.
27. I am familiar with music similar to the one played at the concert.
28. I generally like the style of music played at the concert.

Participatory experience

*Please rate the statements for the concert as a whole. For some items, we will also ask you to rate each participatory part separately.*

*VALENCE*

29. I enjoyed the overall experiment performance.
30. I enjoyed the musician's performance overall.
31. I enjoyed the audience participation parts.
32. Which audience participation part did you enjoy the most?
- Participation 1 (controlling background sounds)
  - Participation 2 (controlling the melody in the foreground)
  - Participation 3 (controlling the instrument of the drummer)
  - None of the above

*AROUSAL*

33. I felt excited during the audience participation parts.
34. Which participation part did make you feel most excited?
- Participation 1 (controlling background sounds)
  - Participation 2 (controlling the melody in the foreground)
  - Participation 3 (controlling the instrument of the drummer)
  - None of the above

*HUMAN-COMPUTER-INTERACTION DESIGN*

35. The interaction software worked well for me.
36. The interaction software was easy to use.
37. I enjoyed using the interaction software.
38. Interacting with my smartphone held me back from dancing when I wanted to.
39. The participatory parts enhanced my musical experience.
40. Which participation part enhanced your musical experience the most?
- Participation 1 (controlling background sounds)
  - Participation 2 (controlling the melody in the foreground)
  - Participation 3 (controlling the instrument of the drummer)
  - None of the above

*EFFECTIVENESS AND MOTIVATION*

41. I felt that my interactions with the software had an effect on the performance.
42. I understood what part of the music was created by the audience.
43. I felt engaged to participate when it was possible.
44. Which participation part encouraged you to participate the most?
- Participation 1 (controlling background sounds)
  - Participation 2 (controlling the melody in the foreground)
  - Participation 3 (controlling the instrument of the drummer)
  - None of the above
45. I feel the audience has some level of ownership of the creative work.



**Interview guide for the group interview with the audience after the concert***0. Feedback on the self reports*

0. Was there any specific points that you think should be addressed or elaborated on after just having done the self report questionnaire?

*I. VALENCE*

1. What did you enjoy or not enjoy about the musical experience?

2. What did you enjoy or not enjoy about the participatory parts of the concerts and why?

3. Which participatory part did you enjoy the most and why?

4. At what parts of the experience do you think the musical style and the participative setting worked well together?

5. What do you think about the relation of the musical genre and the participatory aspect generally?

*II. AROUSAL*

6. Which situations during the concert led to emotional responses such as pleasant sensations during the participation parts of the concert?

7. In which way were the participatory and the non-participatory parts different in terms of feeling emotional responses to the music?

8. How did the participation affect your excitement for the concert experience?

*III. HUMAN-COMPUTER-INTERACTION DESIGN*

9. How did you experience the functions of the interaction software, was it easy to understand what was happening at a specific moment?

10. Was there anything that you didn't understand how to use or made difficulties during usage?

12. What was it that made the tool hold you back from?

13. How did you experience the affordances the tool?

14. How could you describe the way of exploring the tool, did you try to test its boundaries and limitations or use it differently than intended (hacking)?

*IV. EFFECTIVENESS AND MOTIVATION*

15. At what moments did your personal actions produce the musical results that you envisioned?

16. What led to you not feeling that your action had any effect on the music?

17. In what way did your personal contribution seem to affect the musicians' actions on stage?

18. Did you feel actively motivated to interact with the music and the musicians?

19. What would have to be given in order for you to participate in a concert like this again?

*V. SURPRISE*

20. At what points during the concert were you surprised by what happened musically?
21. What surprised you during the participatory parts of the concert and why?
22. At what points did the music performers on stage play differently than you expected?
23. What surprised you in relation to the people surrounding you during the concert?
24. Can you specify what led to you experiencing moments of surprise as enjoyable or disturbing?

*VI. CONNECTEDNESS AND EMPATHY*

25. At what moments during the participation did you feel an enhanced connection with the musicians on stage?
26. How did the environment affect both the aesthetic and participatory experience?
27. How and when did the participation affect the relations to the other audience members?
28. How would you describe the collaborative experience of the concert?

**Interview guide for the in-depth interviews with the musicians after the concert***I. VALENCE*

1. What did you enjoy or not enjoy about the musical experience?
2. What did you enjoy or not enjoy about the participatory parts of the concerts and why?
3. Which participatory part did you enjoy the most and why?
4. At what parts of the experience do you think the musical style and the participative setting worked well together?
5. What do you think about the relation of the musical genre and the participatory setting?

*II. AROUSAL*

6. Which situations during the concert led to emotional responses such as pleasant sensations during the participation parts of the concert?
7. In which way were the participatory and the non-participatory parts different in terms of feeling emotional responses to the music?
8. How did the participation affect your excitement for the concert experience?

*III. HUMAN-COMPUTER-INTERACTION DESIGN*

9. How did you experience the functions of the interaction software, was it easy to understand what was happening at a specific moment?
10. Was there anything that you didn't understand how to use or made difficulties during the integration of the audience-controlled parts to your musical performance?
12. What was it that made the tool hold you back from?
13. How did you experience the affordances the tool?
14. How could you describe the way of exploring the tool, did you try to test its boundaries and limitations or use it differently than intended (hacking)?

*IV. EFFECTIVENESS AND MOTIVATION*

15. At what moments did your personal actions produce the musical results that you envisioned?
16. What led to you not feeling that the audience's actions have an effect on the music?
17. In what way did your personal contribution seem to affect the audiences' actions?
18. Did you feel actively motivated to interact with the audience?
19. What would have to be given in order for you to participate in a concert like this again?

*V. SURPRISE*

20. At what points during the concert were you surprised by what happened musically?
21. What surprised you during the participatory parts of the concert and why?
22. At what points did the audience act differently than you expected?
23. What surprised you in relation to the your fellow musician during the concert?
24. Can you specify what led to you experiencing moments of surprise as enjoyable or disturbing?

*VI. CONNECTEDNESS AND EMPATHY*

25. At what moments during the participation did you feel an enhanced connection with the audience?
26. How did the environment affect both the aesthetic and participatory experience?
27. How and when did the participation affect the relations your fellow musician?
28. How would you describe the collaborative experience of the concert?



# Acknowledgements

Thank you to the Faculty of Arts at the University of Agder for funding this PhD project and supporting my stay abroad as a visiting researcher. Thank you to the administration for being quick and flexible and always helping in the best way possible. I am very grateful for the scholarship and resources that I was able to use for this project. Thank you to my supervisors, Prof. Michael Rauhut and Prof. Tor Dybo, for always being accessible, and for your dedicated feedback, and thank you to everyone else at the Department of Popular Music who has helped and supported me during my time in Kristiansand.

I am grateful for all the other external support from departments, conferences, and especially colleagues with whom I had the pleasure to interact and from whom I learned so much. Thank you to the colleagues at RITMO at the University of Oslo for the great courses I participated in and the valuable feedback I received about my work. Thank you to Prof. Jason Freeman, Ian Clester, and my other colleagues at Georgia Tech for the privilege of studying and collaborating with them. Thank you to the musical performers Vegard Kummen, Tobias Øymo Solbakk, Kristian Olav Vågen, and lighting designer Mike Holm for supporting the concert experiments. And thank you to Nils Nadeau for copy editing this thesis and providing valuable comments along the way.

Thank you to my fellow PhD students at UiA's Department of Popular Music, Lucija, Vegard, Ingvild, and Jelena - it has been a pleasure to share this experience in all its dimensions with you. Thank you to all the other PhD fellows who inspired the project and enriched my social life in Norway.

Lastly, thank you to everyone in my personal proximity for the holistic support during my time as a PhD fellow: my family, Lukasz, Daniel, Aktzi, Elena, Floyd, Ken and everyone who joined me as part of the journey. Finally, thank you to Prof. Bobi for helping me with the translation to Norwegian and everything else.

Matthias Jung  
Kristiansand  
31.01.2024