Assessing creativity in computer music ensembles: a computational approach

Josep M Comajuncosas

TESI DOCTORAL UPF / 2015

THESIS ADVISOR
Dr. Sergi Jordà
Department of Information and Communication Technologies

UNIVERSITAT POMPEU FABRA

To my beloved parents, Josep and Esther

Acknowledgments

First and foremost I would like to express my deep gratitude to Xavier Serra for accepting me in the Sound and Music Computing Master, and later for supporting me as a Ph.D. candidate.

I owe my Ph.D. advisor, Sergi Jordà, a debt of gratitude for his enthusiasm, support and guidance. This thesis would never have seen the light of day without his energetic optimism and the confidence he placed in this project. He also enthusiastically supported our artistic work with the Barcelona Laptop Orchestra and was an inspiration to me in many projects to list here.

I also want to express my gratitude to my colleague, friend and co-director of the BLO, Àlex Barrachina, a great creative mind behind many of our artistic projects and a precious combination of inspiration, toughness and practicality. My artistic research is truly indebted to him. I am also greatly thankful to Victor Sanahuja and Álvaro Sarasúa, for their continued commitment to the Laptop Orchestra, taking responsibility for rehearsals whenever was required. I extend my gratitude to Sonia Espí and Andrés Lewin too, for their support in our artistic endeavors.

My gratitude to John Bischoff and Chris Brown for their inspiring discussions around their work with The League and The Hub. They have been an invaluable source of inspiration both for my artistic and scientific research, and really helped me put things in perspective. Thanks as well to Kazjon Grace and Katherine Brady for their support in the implementation of my creativity metric.

I would like to thank Graham Colleman for his valuable discussions, and Glen Fraser and Roger Pibernat for their active support in the implementation of much of the BLO repertoire. Thanks to Ariadna Alsina and Urbez Capablo as well, two great local composers which collaborated with the Laptop Orchestra. It was a real pleasure working hand in hand with their projects.

I want to extend my gratitude and thankfulness to all active and former participants in the Laptop Orchestra, to Mathieu Bosi, Andrés Bucci, Angel Cataño, Kainan Chen, Miquel Cuxart, Regina Domingo, Aleix Fabra, Antonio Garzón, William Goutfreind, Enric Guaus, Alessandro Inguglia, Nadine Kroher, Quim Llimona, Stefano Marvulli, Emilio Molina, Víctor Núñez, John O'Connell, Carlos Vaquero, Luis Vélez Pedrosa, Alex Rodríguez, Carlos Gustavo Román, Tim Schmele, Oriol Tió, Felix Tutzer, Jan Valls and Nelson Vera. I wish to thank too all the participants of the conducted experiments.

Special thanks go to the people from the MTG administration. Thanks as well to my colleagues at the ESMUC and the Music School of Igualada for their help and understanding. They were always there for support and inspiration.

Last but not least, thanks go to my family and friends for their loving support, understanding and patience.

Abstract

Over the last decade Laptop Orchestras and Mobile Ensembles have proliferated. As a result, a large body of research has arisen on infrastructure, evaluation, design principles and compositional methodologies for Computer Music Ensembles (CME).

However, little has been addressed and very little is known about the challenges and opportunities provided by CMEs for creativity in musical performance. Therefore, one of the most common issues CMEs have to deal with is the lack of a systematic approach to handle the implications of the performative paradigms they seek to explore, in terms of their creative constraints and affordances.

This is the challenging goal this thesis addresses, and for attaining so it first seeks to find a common ground in the strategies developed for assessing creativity in different performative setups, for later proposing an informed pathway for performative engagement in CMEs.

Our research combines an exploratory stage and an experimental stage. The exploratory stage was informed by out artistic praxis with our own CME, the Barcelona Laptop Orchestra. Through the study of the multi-user instruments developed over the past years, we identified the creative constraints and affordances provided by different performative paradigms. Informed by the findings provided by our artistic research, the experimental stage addressed the study of musical creativity through the performance analysis on specifically designed multi-user instruments. For such purpose we proposed a novel computational methodology to evaluate the creative content of a musical performance.

Two experiments were conducted to incorporate our computational methodology into ecologically valid scenarios, aimed at a better understanding of the relationship between topologies of interdependence and creative outcome. For both experiments, we captured performance data from ensemble improvisations, from where the creativity metrics were then computed. As a preliminary step, we investigated the performative engagement and sharing of musical ideations in an ensemble scenario. In a further step, we computed the creativity attributes to comparatively evaluate performances under different scenarios.

The findings provided quantitative evidence of the differences between musical creativity in individual, ensemble and interdependent scenarios. Additionally, the findings point out what strategies performers adopt to best keep their own musical voice in interdependence scenarios, and what novel creative behaviors may be promoted through new topologies of interdependence. Our findings shed light on the nature of performers' creative behavior with interdependent multi-user instruments, and show that the introduced methodology can have applications in the broader context of analysis of creativity in musical performance.

Resum

Durant la darrera dècada les *Laptop Orchestras* i els *Mobile Ensembles* han proliferat arreu. Com a conseqüència d'això, ha aparegut un volum considerable de recerca al voltant de la infraestructura, l'avaluació, els principis de disseny i les metodologies composicionals per ensembles de computadors (CMEs).

Tanmateix, poc coneixem dels reptes i oportunitats que els CMEs ens ofereixen respecte de la creativitat en la pràctica musical. En conseqüència, un dels reptes que la majoria de CMEs han d'encarar és la manca d'una estratègia sistemàtica per preveure i abordar les implicacions dels paradigmes performatius a explorar, respecte de les seves limitacions i possibilitats creatives.

Aquest és el repte que adrecem en aquesta tesi, i per assolir-lo primer tractem d'establir un denominador comú en les estratègies desenvolupades per avaluar la creativitat en diversos entorns performatius, per després proposar un itinerari que permeti assolir una adequada involucració creativa en els CMEs.

La nostra recerca ha combinat una fase exploratòria i una d'experimental. La fase exploratòria s'ha fonamentat en la praxi artística duta a terme en el nostre propi CME, la Barcelona Laptop Orchestra. A través de l'estudi dels instruments multi-usuari desenvolupats durant els darrers anys, hem identificat les potencialitats i restriccions presents en diversos paradigmes performatius. Basats en els resultats de la nostra recerca artística, la fase experimental s'ha centrat en l'estudi de la creativitat musical a través de l'anàlisi interpretatiu en instruments multi-usuari desenvolupats a tal efecte. A tal fi, hem proposat una nova metodologia computacional per avaluar el contingut creatiu d'una execució musical.

Hem dut a terme dos experiments incorporant la nostra mètrica en escenaris realistes, a fi de comprendre millor la relació entre topologies d'interdependència i resultat creatiu. Per ambdós experiments, hem recollit informació d'improvisacions en grup, d'on hem calculat les mètriques de creativitat. Com a pas previ, hem investigat en grau d'involucració i la compartició d'ideacions musicals en escenaris col·lectius. Tot seguit, hem calculat els atributs de creativitat per comparar execucions musicals en diferents escenaris.

Els resultats proporcionen una evidència quantitativa de la diferència entre la creativitat musical en escenaris d'execució musical individual, en grup i interdependent. Addicionalment, ens il·lustren quines estratègies adopten els músics per mantenir la seva pròpia individualitat musical en escenaris d'interdependència, i quins nous comportaments creatius podem promoure a través de noves topologies d'interdependència. Els resultats obtinguts aporten nova llum en la natura del comportament creatiu dels músics amb instruments multi-usuari interdependents, i mostren que la metodologia presentada pot tenir aplicacions en el context més ampli de l'anàlisi de la creativitat musical en l'execució musical.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1 Problem Statement

Musical performance is regarded as a fundamental part of human existence [Rink, 2002]. Indeed, musical ensembles are virtually present in all human cultures, flexibly adapting to the roles, rituals and resources of each context and historical period [Fletcher, 2004]. It should not surprise us, therefore, the rapid adoption of the computer as a musical tool in the mid-twentieth [Mathews, 1963], and the subsequent creation of musical ensembles with computers as instruments only a few years later [Bischoff et al., 1978].

As a classically trained musician and computer engineer, it came to me naturally to found a computer band within the Sonology Department of the ESMUC in 1998, joining the flourishing community of Laptop Orchestras worldwide [Trueman et al., 2006]. Our band, the Barcelona Laptop Orchestra [1] was constituted as a networked computer ensemble devoted to the exploration of performative paradigms of network music through improvisational performances.

The years that followed became a thrilling and challenging endeavor, combining an endless fascination and ingenuity around *digital lutherie*[Jorda, 2005] with recurrent theoretical discussions concerning our roles as musicians and performers in a networked ensemble and the continuous re-evaluation of our work after public performances, the true milestones for our projects.

Not all was plain sailing, however. Beyond the long, painful hardware and software debugging sessions - which should be taken for granted in any computer ensemble- what often became more worrying to us was the feeling of being overtaken by our own setups, instead of being active participants who are in charge of the collective process of music making. In other words, we all shared the intuition that a deeper understanding of the implications of the novel performative envi-

[1] `http://www.blo.cat/`

ronments developed was required in order to take full advantage of their creative possibilities.

Most CMEs face similar challenges and indeed, while a large body of research has arisen, rooted in the diverse infrastructural and compositional strategies adopted by each ensemble, little has been addressed and very little is known about the challenges and opportunities provided by CMEs for creativity in musical performance.

There is, therefore, a need to better know the implications of the performative paradigms that CMEs seek to explore, in terms of their creative constraints and affordances.

This is the challenging goal this thesis addresses, and for attaining so it first seeks to find a common ground in the strategies developed for assessing creativity in different performative setups, for later proposing an informed pathway for creative engagement in CMEs.

1.1.1 Improvising with multi-user instruments: beyond group creativity

Musical performance is a creative endeavor. Creativity manifests in all social roles participating in music, from the composed musical work, to the expressiveness imparted to a score-based performance or the ideations arising during an ensemble improvisation, even as emotions and ideations elicited in listeners [Hargreaves and MacDonald, 2012].

In this respect, ensemble performance with CMEs do often resort to free or loosely structure improvisational practices. Indeed, this has been the case with the Barcelona Laptop Orchestra. In free group improvisation, performers seek to achieve a communal experience in which their ego dilutes and even their individual contributions become disintegrated in a sort of shared identity.

A next step towards a truly conjoint and intimate experience in group performance lies in the sharing of the musical devices themselves. By exploring performative contexts where the traditional assumption of *one performer, one action, one sonic result* no longer holds, we reach a point in which the very notions of personal identity and performer's autonomy are no longer guaranteed.

Those shared performative contexts constitute what we call multi-user instruments: environments where the interdependence which is already present in any group performance extends as well into the intimate link between a musician and his instrument. Multi-user instruments are thus shared instrumental entities in which the sonic outcome necessitates the collaboration of a group of performers to be achieved.

Over the past six years, my work with the Barcelona Laptop Orchestra has been focused in exploring network music performative paradigms for co-located computer ensembles. Among those, we payed special attention to multi-user instruments with different topologies of interdependence. Those instruments epitomize the inherently social nature of network music, and are at the same time a paradigmatic instrumental model in networked computer music ensembles.

For the aforementioned reasons, interdependent multi-user instruments will be the performative environments of interest in this thesis.

1.1.2 Challenges and opportunities for creativity in network music

Performing creatively with digital multi-user instruments is doubly challenging. By one hand, performing with computers tends to place severe constraints in interpersonal communication. By the other, the often complex topology of interdependence established through shared resources may put into question our preconceived performative roles.

This thesis focuses in the study of this second challenge.

As multi-user instruments do challenge that traditional instrumental paradigm by allowing several creative processes to be concurrently shaped and mutually influenced by several performers, the new performative context they provide may have a great impact in performer's awareness and agency, both individually and collectively.

We may ask ourselves how musical creativity is affected by such contexts where the very own instrumental identity is challenged. Will them promote a more communal discourse, no so based based on individual contribution but on more collectivized musical ideations, or, on the contrary, the complexity of the interdependences, the competence for shared resources and the higher degree of indeterminacy and sonic emergence might eventually inhibit creativity?

Some authors claim that such interdependence in multi-user instruments is a good prospect for intense and engaging interplay [Weinberg, 2003] [Jordà, 2005b], and experimental evidence of such claim is provided in the studies on mutual engagement and its relationship to group creativity by Bryan-Kinns, who carried out extensive studies with collaborative musical interfaces [Bryan-Kinns and Hamilton, 2012] [Bryan-Kinns, 2013].

Our experimental research will follow a similar approach in order to answer the aforementioned questions, by resorting to the analysis of performer's behavior and the creative content of their contribution. For this purpose, we will carry out comparative evaluations of distinct performative scenarios with interdependent multi-user instruments. Its main contribution, in this respect, will be bringing back musical creativity as a criteria for the evaluation of the performative environments of interest.

Ultimately, we expect that the additional insight into the relationship between topologies of interdependence and creative outcome will provide us a better understanding of the challenges and opportunities for creativity in CMEs.

1.1.3 Towards the computational assessment of creativity in music performance

Our experimental approach lies at the intersection of several computer and music related fields, namely Human Computer Interaction (HCI) and more specifically New Interfaces for Musical Expression (NIME), Musicology, and in particular the study of Music performance, and finally Computational Creativity, a field concerned with theoretical and practical issues in the study of creativity. We will briefly put in context those disciplines and highlight the research gaps identified through our literature review.

Research in **NIME** has addressed the relevant fields of multi-user and collaborative interfaces, their taxonomies and criteria for evaluation. The emerging scene of Laptop and Mobile ensembles contributed with an emphasis on often particularized approaches to sound, network and programming infrastructure, as well as in the evaluation, design principles and compositional methodologies.

While in cannot be denied the role played by scientific and technical research in the genesis of a new *ensemble identity* for CMEs, there is a need to address the *creative implications* of the new performative environments in a more systematic way. This evidence asks for further quantitative, user-experience based evaluation methodologies focused on the study of such environments, and the use of state-of-the-art tools for the computational assessment of performative behavior.

Research in **Music Performance** is addressed from different perspectives according to the musical context of study. When surveying the literature on music performance analysis, a clear distinction can be drawn between the analysis of score-based performance and the analysis of improvisational music genres. Over the last decades, a large body of research addressed the analysis of expressivity in score-based performance, a broad term which encompasses the creative resources utilized by an interpreter when performing a composed work. Most quantitative research on expressivity relies on performance measurements, greatly facilitated by the adoption of the MIDI protocol in keyboard instruments.

By contrast, analysis on improvisational practices, in which the creative contribution of performers is far more explicit, tend to be more qualitative and focused on motive development and adherence to the stylistic conventions of the musical genre of study, and, in the case of ensemble performance, on the study of coordination and negotiation of roles between performers.

We might conclude that the less adherence to clearly stated goals in a performance - be them score instructions or stylistic rules - the less prevalent are the quantitative analysis of the performance. At the same time, while studies of individual performance focus in the analysis of the musical outcome, studies of group performance tend to be more focused in the analysis of the behavior and interaction between the members of the ensemble.

We may therefore conclude that there is a lack of quantitative, outcome based analysis in musical ensemble performance based on improvisational practices.

This is the case in the study in the analysis of music performance in computer mediated environments such as CMEs, which as we mentioned mainly resort to free improvisational paradigms: the majority of research conducted consist of qualitative, ethnographic studies [Booth and Gurevich, 2012] [Reinecke, 2010], only occasionally combined with quantitative performance measurements [Troyer et al., 2012]. Nevertheless, the evaluation of the actual musical outcome is far less relevant that the study of performers' interaction and their strategies for collaboration.

With respect to **Creativity**, we may distinguish two approaches in the research literature, related to our focus of interest: whether *process-based*, aimed at the understanding of the creative process from a cognitive or psycho-social perspective, or *product-based*, which aims at establishing uniform criteria to evaluate the creativity of artifacts, products and ideations.

As we want to assess creativity from performance analysis, our approach will be product-based. Several *computational models* have been proposed, halfway between the disciplines of Artificial Intelligence and Design Science, which identify the key attributes of a creative product and propose a methodology to quantitatively evaluate them.

When comparing such computational models of creativity to the theoretical models of musical improvisation proposed in the literature, we may observe striking similarities. However, computational creativity models for the assessment of musical creativity in musical performance are yet to be fully explored.

Once again, we conclude that there is a need to incorporate state-of-the art methodologies for computational creativity assessment into the study of musical creativity.

Our approach will therefore address the gaps detected in the literature on NIME research for CMEs, performance analysis and creativity research, by providing a **quantitative, outcome based metric for musical performance analysis based on a new computational model of musical creativity** which will be fully developed and incorporated in our experimental research with CMEs.

Finally, while we focus our study of musical creativity on performance with CMEs, we believe the results may shed light on the nature of creativity in musical performance in a broader sense. We will briefly justify this statement.

The admittedly specific case of multi-user instruments in computer music ensembles places an extremely restricted scenario which may subvert most of the assumptions taken for granted in a musical ensemble. By analyzing their impact on musical creativity and determining what strategies performers develop to face the challenges inherent in networked computer ensembles, we may better know their relative importance in the creative process as it takes place in a musical ensemble.

In short by seek to assess musical creativity in constrained and potentially challenging, but ecologically valid performative scenarios such as the most paradigmatic ones in networked computer ensembles, which are interdependent multi-user instruments. From their evaluation we expect to gain greater insight into the prerequisites, motivating and inhibiting factors that contribute to the creative practice in musical groups.

A strong reason to assess creativity in computer music ensembles is their flexible nature. They constitute a performative environment more suitable for controlled experiments than their acoustic counterparts, allowing for a tight control over the experimental variables. Musical instruments and their interdependences are easily reconfigurable by software, user interaction may be regulated by computer mediated feedback, the dimensionality of controlling gestures and sonic outcomes are up to the designer and finally such data may be easily captured for posterior analysis. In short, computer music ensembles constitute an environment suitable for comparative studies on musical performance creativity assessment.

1.2 Research methodology

Our research methodology is based in the aforementioned assumption that Computer Music Ensembles provide an environment well suited both for artistic research and for experimental research on musical creativity:

- Because of their flexibility, they facilitate the exploration of different performative scenarios to comparatively evaluate them.

- It is entirely feasible to design controlled experiments which are ecologically valid, as long as they adhere to the performative practices of existing CMEs.

- Compared to acoustic music ensembles, CMEs greatly facilitate the capture of performance data by means of their own design.

Our research will therefore incorporate both elements of artistic and experimental research, by combining an exploratory stage and an experimental stage.

As an exploratory stage, this research incorporates ethnographic and autoethnografic methodologies to determine what creative constraints and affordances are provided by the multi-user instruments developed throughout these years of artistic practice in the context of the Barcelona Laptop Orchestra.

Informed by these findings, we will seek to quantitatively study musical creativity on multi-user instruments through a second experimental stage.

For such purpose, we will propose a novel computational metric for creativity assessment, based on existing literature on product-based creativity metrics, to assess creativity in musical performances with computer ensembles. We will then conduct two experiments in which we will incorporate such metric to evaluate musical creativity in free ensemble improvisations on specifically designed multi-user instruments.

We expect our experimental findings will support and provide further evidence on the nature of performer's creative behavior in multi-user instruments.

1.3 Aim of this thesis

This thesis addresses the following research question

What are the challenges and opportunities provided by digital multi-user instruments for creative performance in computer music ensembles?

This question will be addressed from two complementary and interrelated perspectives:

- Through our research into the artistic process, we will identify the creative constraints and affordances provided by the performative environments developed.

- Through experimental research, we will quantitatively assess creativity in musical performance in order to gain further insight into the requirements, challenges and opportunities provided by such performative paradigms.

Our experimental approach will seek to investigate how performers address musical performance on multi-user instruments with different topologies of interdependence. Through the research into such specific context, we will attempt to shed light on the nature of musical creativity in multi-user instruments as well as more generic issues concerning creativity in musical performance.

In this respect, this research aligns with the recent trends in the research in Computer Supported Collaborative Musical Interfaces, which combine user-experience evaluation with quantitative measurements of ensemble interaction and collaboration - see for example [Fencott and Bryan-Kinns, 2012], [Bryan-Kinns, 2013]

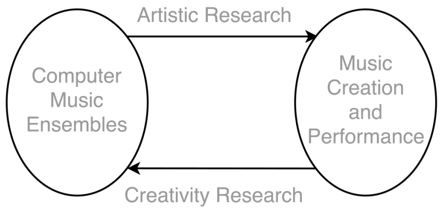


Figure 1.1: A schematics of the interrelation between our two research approaches

and [Xambó, 2015]). In our case, our quantitative measurement goes one step beyond, bringing back a relevant attribute of the musical performance as is its creative content into the evaluation of the user experience within multi-instrument based performative environments.

Refining the main research question in light of the aforementioned observations, the following subsidiary questions will be addressed to provide as comprehensive a vision as possible of the nature of musical creativity as manifested in our performative contexts:

1. What is the nature of performers' creative behavior with interdependent multi-user instruments?

2. What are the commonalities and differences, in terms of musical creativity, between a free improvisation with individual instruments and a performance with a multi-user instrument, in the context of a network ensemble?

In order to address these questions, two studies are conducted, focusing on understanding how interdependence and shared resources influence creativity. The first experiment explores the continuum between independence and full interdependence, while the second experiment focus on the study of the effect of control and sound ownership. The two experiments aim to shed light on the two subsidiary research questions just mentioned.

1.4 Summary of the Contributions

Following this agenda, the novelties, achievements and main contributions of this thesis are several, which I summarize below:

1. A novel *Interdependence metric* for the classification of interdependent networks, accounting for (i) the number and typology of resources shared and

(ii) the amount of interdependence in such resources, has been proposed, based on Booth's matrix representation for interdependent systems.

2. In the domain of *Network Music repertoire*, some novel and unique performative paradigms have been explored and developed through the artistic practice documented in this thesis, such as (i) in-depth exploration of a co-located asynchronous environment for ensemble performance, (ii) mapping techniques to facilitate target-based gestural navigation in concatenative synthesis, (iii) the use of machine listening and machine learning tools to assist in the conduction of multi-user instruments, and (iv) the paradigms of time-interleaved multi-threaded collective sequencer and homogeneous environment developed for Experiment 1.

3. The concept of *Creativity metrics* based on the assumption that a musical improvisation can be characterized as a sequence of gestural ideations, (i) is new, as it is (ii) the use of dimensionality reduction techniques (PCA) to build the needed conceptual spaces, and (iii) the use of GMM based novelty assessment in the context of creativity studies.

1.5 Thesis RoadMap

This research project is multidisciplinary as it addresses diverse topics of music performance, creativity and music creation with computer ensembles. Figure 1.2 schematizes the interrelation of the three disciplines: **Computer Music Ensembles**, which are our research environment and **Music Performance**, which is our context of study, both addressed in our artistic praxis and in our experimental research. Finally, through measurements of performance we will quantitatively evaluate **Musical Creativity**, as it manifests in our musical practice with computer music ensembles.

We will ultimately propose a **computational model for music creativity** which, through performance analysis in controlled experiments with computer ensembles, will be employed to comparatively evaluate a number of scenarios in interdependent multi-user instruments, to provide additional evidence and further insight into the research questions aforementioned.

The subsequent chapters of this thesis, finally, are structured as follows:

- **Chapter 2** Introduces the social and aesthetic implications of musical composition, performance and improvisation. It discusses the role of the musical score and the conductor in the creative processes of a performance. Finally, is what opportunities for creativity arise in musical improvisation through different cultures and cultural contexts.

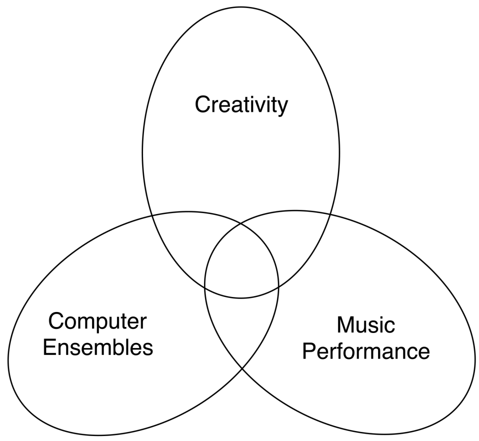


Figure 1.2: Research disciplines involved in this thesis

- **Chapter 3** Reviews the literature on network music performance. It contextualizes historically the phenomenon of Network Music, characterizes multi-user instruments and reviews the main taxonomies and evaluation techniques for digital musical interfaces and computer supported collaborative music interfaces.

  Finally, it surveys the literature on Laptop Orchestras, identifies the prevalent ensemble models and reviews the diverse compositional and notational approaches in the repertoire for computer music ensembles.

- **Chapter 4** Illustrates the aforementioned performative and compositional paradigms for computer ensembles with an in-depth repertoire analysis of the Barcelona Laptop Orchestra. We focus in the analysis of the creative constraints and affordances provided by the multi-user instruments explored.

- **Chapter 5** Reviews the literature on Music Performance analysis, from research based on performance measurement to studies on subjective evaluation, from research focused on individual performers to the analysis of group performance and improvisation. It finishes with a survey of the research in performance evaluation for Computer-Supported Collaborative Musical Interfaces and Computer Music Ensembles.

- **Chapter 6** Reviews the scientific literature on Creativity, the more relevant theoretical frameworks and computational metrics to assess creativity and surveys the current state of the art in music creativity research and the proposed computational models of musical performance, improvisation and creativity.

- **Chapter 7** Proposes a metric to assess Creativity particularly suited to evaluate creative musical performance in the context of network music.
- **Chapter 8** Presents two studies on creativity assessment for multi-user instruments within network music ensembles.
- **Chapter 9** Summarizes the contributions of this thesis and draws directions for future research.

Chapter 2

COMPOSITION, PERFORMANCE AND IMPROVISATION: THREE FACETS OF MUSIC CREATION

2.1 Introduction

This introductory chapter aims at providing the philosophical and aesthetic grounds of musical creation. The following sections will therefore review, from a conceptual point of view, the relationship between music creation, composition and performance. The key concept here is *musical creativity*, though it is presented here as it is commonly and informally accepted when we refer to musical creation. It should be emphasized that no formal definition, neither evaluation of this concept will be attempted in this chapter.

As a starting point, we discuss the relationship between two manifestations of musical creation traditionally considered antagonistic, at least in Western music tradition: the composed work and its relationship with the musical score, and musical improvisation.

The next section, devoted to expressiveness in (score-based) music performance, further discusses the relationship between musical composition and the performance of a composed work, the socially mediated role of the performer and his opportunities and strategies for *being creative through expressiveness*.

The following section examines the other facet of performer's creativity: music improvisation. We analyze some of its defining traits through comparison both with musical composition and with spoken conversation - a recurrent analogy when defining what musical improvisation is.

As ensemble improvisation is particularly relevant in the context of this thesis, an extensive survey of the diverse improvisational practices through musical cultures is provided.

Finally, we examine the role of a musical conductor, if present, in the musical performance, both in score-based performances and in improvisational ensembles.

2.2 Composition: from the musical score to the improvisational act

2.2.1 Composition and the score-centric perspective

From the traditional, modernist classical Western music perspective, musical works are autonomous entities of an ideal quality, which once created have a timeless, immutable existence [Husserl et al., 1973]. Such works, instead of popular music, are not transmitted orally but through notation, and are ultimately preserved and transmitted by means of their interpretation in a music performance.

Musical works share in this respect similar attributes to other artistic creations which remain somehow immutable beyond endless repetitions, such as plays, which are continuously performed or novels, which are continuously read.

This static perspective inevitably brings to the ossification of the repertoire, kept by written transmission and object of veneration and rejecting any significant variation. Assuming that a Musical work is a sort of *ideality* which transcends any particular interpretation, musical performances are mere particular instances of it, the more valuable the more closely they follow the score. Indeed, complete compliance with the score is both the necessary and sufficient condition for genuineness and the definitive guarantee that the musical work will be preserved. In Goodman's words, [Goodman, 1968] *if we allow the least deviation, all assurance of work-preservation is lost*.

Historical and sociological evidences put this score-centric approach into question, though. No matter how much efforts are put in preserving the work, once a music piece is made publicly available it will acquire an autonomous existence and will be subject of evolution by means of *continuous recontextualization* [Benson, 2003]. It suffices to listen to historical recordings to realize how musical canon is a social construct, and how the schematic, inherent ambiguity of musical scores place a creativity opportunity (and a requirement) for performers.

A Notation-centric musical culture prioritizes some skills (such as sight reading) with to some authors may see as inhibiting improvisational skills. To Bailey, however, possibly the most positive effect of the removal of the music score is the fact that it gets rid of the composer bailey1993improvisation]. Conversely, it is not possible to transcribe improvisation [Bailey, 1993], and even if one might transcribe into common music notation an improvised performance (such as jazz solos) it misses the whole point of the improvisational act.

Nowadays the very concept of musical score as the only documentation of compositional work is into question. By one hand, since the second half of the XX century, avantgarde composers questioning conventional musical notation with graphical scores which incorporated from indeterminate or aleatoric passages up to visually tantalizing graphic creations, eventually advocating for absolute interpretative freedom (such as in Cardew's Treatise[Cardew, 1971]). Conversely, com other composers utilized more conventional notational resources in non-conventional contexts, incorporating elements of chance music in their performance (such as in Stockhausen's *Klavierstück XI* or in Earle Brown's *Available forms II*).

Many contemporary genres do make extensive use of alternative digital formats to encode relevant performance instructions. Some widely used formats, like MIDI, have their own idiosyncrasies and limitations as well. Similarly, many works do require additional materials to be performed, such as sound samples and computer code, which despite their apparent reliance and unambiguity have shown to suffer from the obsolescence of storage supports and software.

Actually, the role and scope of musical scores in a technologically mediated context may change dramatically, from scores as self-sufficient encoded performances for humans or computer programs to scores which act as instructions for generative processes, to, ultimately, the concept of computer code as musical score.

2.2.2 Composition and performance

In cultures where music is transmitted orally, there is hardly any difference between composers and performers: music is just created and played. Only with the incorporation of music notation both roles gradually separated, although most composers in the Western classical tradition performed their own music as well.

However, the prevalent belief in contexts where both roles are clearly established, such as in Western classical music, is that composers create while performers preserve their artistic legacy. Under such assumptions, only the composer is the true creator and the performer must adhere to the ideal of *Werktreue* [Goehr, 1992], as a faithful servant of the score. Thus a performer shows his faithfulness to the composer by accurately following what's in the score (*Texttreue*, fidelity to the written score, otherwise his performance may be questioned as *inappropriate*.

The very role of a composer is imbued with an aura of mistery and mysticism since the Romantic period. Contemplating musical works out of context and minimizing the vital, aesthetic and historical influences exerted on their creation just exacerbates this view. In contrast, composers belonging to the established, sacred canon of Western classical music were themselves much more prosaic regarding their own work. Take as an example the well known quote from Bach - *My music is better because I work harder. Anyone who works as hard as me will write music that is just as good*.

However Bach himself carefully annotated very precise embellishments in his instrumental pieces: both a testimony of his improvisational techniques and a clear statement against excessive extemporization by performers [Benson, 2008].

The paradox is that this clear division between creativity in composition and performance was not the case with some of the most venerated composers, especially prior to the Romantic period. As he composer was more often than not the performer and music was primarily functional, musical compositions where short-lived, flexibly adapted and performance oriented.

Possibly the most rigid approaches to written music performance came from composers from the first half of the twentieth-century. Take as an example Stravinsky, notorious champion of objective performance, who asked performers for minimal creative interpretation, which he himself identified with terms such as criminal assault or betrayal (!)[Stravinsky, 1970].

This perspective reached a climax with the Post-WWII, Darmstadt school composers who advocated for the utmost strict accuracy in the rendition of their works. The musical score was just regarded as an exhaustive set of instructions to performers and an absolute, transparent relationship between notation, realization and perception was expected. For integral serialists in the 50s and 60s, performers were just a passive mediators between the composer and the audience (see, for example, [Stockhausen and Maconie, 1989] and [Babbitt and Peles, 2003]).

In a process parallel to the development of more open art forms, this strict role separation and hierarchization was progressively called into question within the new musical avantgardes. Even with the extreme challenges found in scores from composers from the so-called *New Complexity*, epitomized by B. Ferneyhough, a more bi-directional dialogue between the performer and the composer's intent is expected: in a way, the sense of intensity conveyed by the effort in attaining absolute accuracy and the unavoidable failure when a passage enters the unperformable realms are both to be regarded as equally positive [Duncan, 2010].

2.2.3 Improvisation and composition

Improvisation and Composition are not as clearly delineated creative activities as is conventionally assumed. Dobrian cites[Dobrian, 1991] three common distinctions made between both (Written vs non written, Real-time vs non real-time and Group vs individual activity) and shows how blurry such distinction can be, and how is it often informed by social cliches inherited from a Romantic archetype of a composer.

Indeed, the compositional act has been traditionally seen as qualitatively differentiated from improvisation. To Wolterstorff[Wolterstorff, 1980], for example, an improviser is not properly composing even while improvising, as his activity lacks a reflective process which serves to determine requirements for correctness as found in a composed score. It would seem that composition and improvisation exhibit two distinct kinds of creativity, which implicitly differ in their social value.

Not surprisingly, the very term improvisation has negative connotations, such as *ad-hoc*, chaotic or unprepared music, up to the extent that some improvisers avoid this term to define their art [Menezes, 2010]. The fact that improvisation is still poorly understood and defies the analytic approach employed to study composed music contributes to such traditional lack of acknowledgment:

> Defined in any one of a series of catchphrases ranging from *making it up as he goes along* to *instant composition*, improvisation is generally viewed as a musical conjuring trick, a doubtful expedient, or even a vulgar habit.
>
> Derek Bailey, in *Improvisation: Its nature and practice in music*[Bailey, 1993]

Conversely, composition within the European tradition still retains an aura or superior craftsmanship, an attribute that clearly differentiates it from irreflexive and contingent improvisation. As Stravinsty states, composition is "the fruit of study, reasoning and calculation that imply exactly the converse of improvisation" [Stravinsky, 1970]. At most, Stravinsky, who was a proficient improviser and composed itself by playing on the piano, regarded composition as *filtered improvisation*.

Benson [Benson, 2003] argues against those either pejorative or mystical definitions of improvisation which assume that it is creation from nothing. Instead, he defends that all creative activities involve reworking of something that already exists. And conversely, it is equally questionable that composers create *ex nihilo*. To illustrate this, the author provides extensive examples of improvisational activities which settle between composition and performance, such as ornamenting, realizing a *continuo*, orquestrating or arranging a score for example. Needless to say, Benson understands the term improvisation in a broad sense there.

From a different perspective, some composers see improvisation as the driving force in the compositional process, which ultimately obeys an irresistible creative urge which, by its immediacy, bears striking resemblances to a live improvisation. As paradoxical as it may seem, for example, Schoenberg was known for writing extremely fast [MacDonald, 2008], indeed he believed that *Composing is* *a slowed-down improvisation; often one cannot write fast enough to keep up with the stream of ideas* [Schoenberg and Stein, 1975].

If, despite the claims for clearly differentiated generative processes in composition and performance, they are creative acts which greatly overlap both procedurally and conceptually, a similar observation can be done from listener's viewpoint. In an empirical study based on listening experiments, Lehman [Lehmann and Kopiez, 2010] noticed that listeners cannot easily discern one generative process from the other. Interestingly, such perceptual discrimination was higher in a pair of composed and improvised excerpts from a free jazz improvisation. The author hypothesizes that listeners search for cues related to the performative character to infer the underlying generative processes: constraints and "togetherness" of an ensemble may indicate composition, and a higher degree of entropy could signal improvisation.

As a conclusion, the limits between composition and improvisation are not so well delimited and mutually exclusive as it is commonly accepted [Benson, 2008]. argues for a new perspective: that both composers and performers are improvisers in a broad, rather philosophical sense. Ultimately, to the author, music making manifests as a sustained process of creation and recreation, put simply improvisation. Indeed neither the limits between composition and improvisation as traditionally understood, nor between composition and performance are clearly delimited and such "messiness" is what actual musical practice reflects.

2.3 Performance *as* musical creation

Music performance is regarded as a highly creative activity. In some musical genres such as jazz or contemporary open forms, creative improvisation skills are highly regarded, both at personal and collective level. Yet creativity is present in the performance of Western classical score-based repertoire as well, as long as performers do constantly re-interpret the repertoire according to novel perspectives, often unexpected but nonetheless highly valued by the audience. Creativity is not assumed to be inextricably associated to music performance however, indeed, in certain contexts it may be even avoided, such as in ritualistic or therapeutic music [Clarke, 2005b].

The traditional Romantic aura of creativity was a powerful drive towards defining attributes of musical creativity such as novelty and uniqueness. But creativity in Western classical music performance has been as well promoted as a sign of distinctiveness and authenticity to attract audiences[Williamon et al., 2006].

In the particular context of Western classical music, performer's creativity mainly manifests as what is broadly known as *musical expression*. Musical expressiveness has been traditionally understood as irregularity, deviation from the score or departure from the performative canon [Clarke, 2005b]. This is arguably most important source of musical creativity in a performance, the interpretative nuances carefully crafted (and highly automated) through intensive practice.

But a live performance may exhibit another kind of creativity, that related to the adaptive behavior required in a performance. Chaffin [Chaffin et al., 2006] finds evidence that performance cues set up in rehearsals define points of conscious control and intervention where performers adapt to the unique demands and opportunities of each performance. This source of creativity imparts spontaneity and and is one of the reasons why live performance are still highly valued nowadays.

2.3.1 Creativity and ritualism in music performance

Ritualism is the opposite of creativity, or more precisely, the attitude which most strongly inhibits creative behavior. As much as the social function of rituals, music associated to them has conversely a ritualistic nature and avoids novelty at all costs. Clarke [Clarke, 2012], exemplifies the need for such ritualistic approaches to music performance in three distinct social contexts: in the preservation of cultural heritage, as a metaphor for ritualistic work (as a tool to help in coordinating physical work for example) or in music therapy, where songs have a well defined purpose and their precise replications reinforces their effectiveness. Much more contexts may benefit from such a ritualistic use of music, such as in religious, sportive or political events, or in commercials jungles or theme tunes in radio and TV programs, to mention just a few. It suffices to resort to a live performance instead of a recording in one of such events to realize what disturbing comicality may entail any kind of creative reinterpretation of the ritualized musical excerpt.

This process of musical fossilization is shared among different musical cultures, both Western and not Western, both considered high art or not. Whenever a social group focuses on the preservation of musical corpus, similar social patterns are to be observed. For example, the pioneering revivalist movement in the 60s and 70s and the subsequent trend towards Historically informed performances shared similar trends towards the claim for an authority, preservation of the tradition and an even sacred approach to music making [Bithell and Hill, 2014].

Dobrian [Dobrian, 1991] points out social factors which might discourage spontaneity in music performance. In a (legitimate or not) desire to achieve greater intellectual validity and respect, musicians may pretend their performance carefully avoids arbitrariness and is otherwise the product of careful and conscious planning. As an example, consider a composer who wants to exhibit his intellectual skills through a performance of his work, or a performance which is focused on the exhibition of the skills of execution (a competition, for example). In those situations it may be clearly preferable that the performance conveys the impression that everything was carefully planned up to the last detail.

From an ethnomusicological perspective, Juniper Hill [Hill, 2011] analyzes the creative practices in six different music-cultures, in order to elicit what cultural beliefs and conventions and what performance practices encourage or inhibit creativity in music performance. As inhibiting factors we could mention a more pronounced division of labor (between listeners and performers, or between performers and composers), mystical beliefs on the origin of music which may make musical creativity appear to be inaccessible, perceiving that only geniuses or shamans are able to be musically creative, acknowledging authority to be creative only to legitimized masters, or having a sacralized view of music, for example giving higher social value to preservation of musical heritage and/or striving for historically authentic reproductions, an attitude often enforced through community censorship

Conversely, creativity will be greatly encouraged in a more egalitarian music-making environment which holds the belief that musical creativity is actually achievable for anyone and which provides authority and opportunity for creative practices, actively engaging members of all the community both in music performance and creation, accepting as stylistically appropriate musical variations, a context in which musicians learn from the traditions but then express them in their very own individual way, where the ideal of authenticity resides not on replication but on embodying the creative process, where creation is a collective process which enhances group solidarity.

2.3.2 Conducted performances and conducted improvisation

A brief history of musical conducting

The conduction of musical performances dates back to the very creation of stable music ensembles. Hieroglyphics and stone etchings do illustrate the presence of conductors providing pitch and rhythmic instructions to players. In early vocal music in the Middle Ages a conductor was responsible of synchronizing among singers music which was rhythmically flexible (text based and not meter based) and ornamented. Such technique, consisting of hand gestures indicating the melodic line is known as *cheironomy*, and is still of use in the singing of liturgical music in catholic churches and in synagogues.

Early instrumental ensembles could successfully coordinate their performance without a conductor. Usually one of the performers was in charge: a continuo player or a concertmaster could give the required visual cues to the other performers.

It is in the nineteenth when the role of professional conductor is well established. As symphonic orchestras evolved and became larger and more complex, similarly the role of conductors gradually became more central. The romantic fascination with personal expression and individuality naturally extended to the emerging figure the *virtuoso composer*, with Berlioz being considered the first reference.

Historical video recordings indeed show us how individualistic such conductors were. Not until the end of the twentieth century the technique of orchestral conducting was standardized enough, both allowing for a more efficient use of rehearsals and making possible for itinerant conductors to easily communicate with new orchestras.

Conductorless ensembles and conducted improvisation

In instrumental music genres there exist well assumed contexts for conductors to be accepted as necessary for a successful performance. Negotiation through auditory and visual feedback (in chamber music) or informal turn-taking and hierarchic schemes (in jazz) are typical examples of coordination without a conductor. However, it is accepted that a larger ensemble benefits from having one.

Though the presence of a conductor might suggest that performers must face higher constraints in their opportunities for creative (ie expressive) performance, evidence just proves the opposite. Conductorless orchestras, which do exist mostly as a political statement of new horizontal, democratic forms of orchestra management, are often criticized for a lack of coherence and long term coherence in their performances. A plausible reason for it would be that performers must invest additional energy in negotiating synchronization between them -rhythmic, *incipits* and endings, dynamic timing and so on- through non verbal cues, thus requiring a higher use of communicative motivated gestures. All those facts may actually inhibit, or at least put occasionally on a second term performer's focus on his own expressive contributions.

Two paradigmatic and rather idiosyncratic examples of conducted improvisation are the *Conduction*®, by Butch Morris, and *Soundpainting*, by Walter Thomson. Another example of conducted improvisation, close in spirit to the practices of performance art, was Frank Zappa's conducted improvisations with the Mothers of Invention, which often involved audience participation. All conducted improvisation practices rely on a lexicon of conductor signs, a musical sign language which have a broader scope than classical conduction.

Butch Morris developed his sign system inspired by Los Angeles new jazz scene and the creative emergence in the Bay Area of San Francisco of the late seventies. He actually introduced his concept in 1985. Later in New York, Morris actively collaborated with writers, dancers, visual artists and performers of varied musical styles on multi-disciplinary performances. Throughout his live, Morris offered around 200 Conduction® events around the World.

Walter Thomson, by the other hand, started working with his sign language around 1974, while working with his teacher Anthony Braxton. He called it Soundpainting claiming it was a universal composing language for composers and artists of all disciplines and abilities. His language is widely used in music education and an active community of Soundpainters continuously expand the corpus of signs in the language.

Both methods therefore are aimed at multidisciplinary improvisation and exhibit a range of signs which are purposefully ambiguous [Scholar and Zanter, 2013] [Duby, 2006], allowing for a broad margin of freedom for performers to translate and express conducted signs into their own vocabulary.

A positive aspect of a conductor shaping the work is its balancing role, ensuring that performers keep a proper weight of their own contribution in relation with the whole, even in the most virtuosic passages [Scholar and Zanter, 2013].

Little research has been carried regarding the creative behavior of performers under a conducted improvisation scenario, even if conducted improvisation is widely used to teach improvisation and performance attitudes that it conveys, such as quick reaction to conductor commands, flexibility to translate them to one's own instrument and collective awareness.

Faria [Pi and Larsen, 2015] explores how classical performers face musical indeterminacy as found in Soundpainting. The author argues that, by proper selection of the repertoire of signs being learned and used, classical musicians can gradually build confidence by feeling able to manage an expanding improvisational framework. Aimed at the study of the pedagogical usage of the practice of conducted improvisation, Pi [Pi and Larsen, 2015] made similar remarks, though warning that there is a learning phase which cannot be omitted to achieve a fluid internalization of the sign language.

2.4 Improvisation

In improvisational contexts, creativity may manifest in the generation of musical materials as well. Several cognitive models have been hypothesized to support the unpredictable outcome in improvisational performances. From such outcome based perspective, Gibbs [Gibbs, 2010] relates creativity to inventiveness and originality, while Cohen [Cohen, 2012] sees creativity as a relational practice, which makes extensive use of reworking and derivation of preexisting materials which are appropriated and recontextualized.

But once again, the mere analysis of the procedures involved in the generation of musical materials in an improvisation won't tell us the whole picture of the creative process. Of comparable, if not higher importance is the physicality of the improvisational act itself, understood as a continuum from ergonomic to choreographic informed decisions [Clarke, 2005b].

For example, in their study on jazz performance, Davidson [Davidson and Coulam, 2006] observed that creativity in performance can be modulated though manipulation of the socio-cultural elements involved, thus shaping co-performers and audience intercommunication and perception of the musical process. The author found that performer's ratings on the expressivity and creativity of a musical performance show direct correlation with the amount of illustrative and adaptive gestures, while technical regulators and display gestures had an inverse effect. They conclude that illustrating musical ideas and intimate states in a fluent and cohesive way lead to higher appreciation of the quality and inventiveness of the performance.

Such social interaction is of paramount importance within the ensemble as well, to sustain creative flow in improvisational scenarios. Berliner [Berliner, 2009] describes the conversation-like nature of jazz improvisation, as a performance which moves forward as performers give and take new musical ideas, elaborating or modifying them thus establishing a creative musical dialogue.

2.4.1 Improvisation as conversation

Musical improvisation is a complex process which can be viewed from different perspectives. By extending the analogy of music as language, free improvisation may be considered as a form of conversation. Seeing music as a language for communication would assume that music is meaningful. What is less clear is what music actually means, because the negotiated meaning of spoken words in a spoken conversation greatly differs from, say, the meaning of musical motifs played in a musical conversation. Swain [Swain, 1997] defends that music is self-referential, it does not refer to objects in the world as words do, and accordingly is has a less defined and precise meaning compared to words. Possibly the most agreed meaning in music is emotion. From the perspective of cognitive psychology, Slodoba provided evidence of significant agreement in the emotional content of certain musical features and events [Sloboda, 1992].

However, while language and music communicate different things, there are indeed shared elements in the communicative act itself. Sutton [Sutton, 2001] resorts to conversation analysis methodologies to reveal interpersonal and interactive processes taking part in a musical improvisation, analyzing the patterns of interaction from existing audio recordings of duets played by established free improvisers. The author defends that free improvisation is like conversation, highlighting how musicians negotiate musical beginnings, endings, turn takings and silences throughout an improvisation.

2.4.2 Improvisational practices though music cultures

Improvisation in ancient and classical Western music

Improvised practices played an important role in Medieval liturgical music, alongside written repertoire. With the advent of polyphony, performers learned specific techniques to layer improvised counterpoint to *cantus firmus*. This practice is found in popular dance music as well.

In the Renaissance and the early Baroque period some popular musical forms were based on improvised melodies over repetitive chord sequences, such as the *Chaconne* or the *Passacaglia*. Such harmonic structure was provided by the *Basso Continuo*, a flexible arrangement of polyphonic instrumentalists who rendered extemporaneously the chord sequence by following a schematic figured bass notation. The art of improvised accompaniment survived in organ church music, indeed, some of the greatest improvisers of all times were accomplished organists, notably Buxtehude and J.S.Bach. Another form where the performer was expected to improvise was the *Da Capo Aria*, as seen in operas and oratories. In a Da capo aria, of ternary, reexpositive structure, singer's ability to improvise ornamented variations in the repetition of the first section was highly were highly regarded. Baroque music was not a period of fossilized forms but of extremely active search and innovation. This attitude translates within the musical work itself, variations, embellishments, *fiorituras* have been considered the musical counterpart to the Improvised vocal and instrumental ornamentation on melodic lines where so commonplace that they were gradually incorporated by composers as annotated embellishments, making use of precise abbreviation symbols. However, improvisation was also present and deeply integrated in the performance practices. As Couperin summarizes, "what we write is different from what we play" [Rink, 2002].

Contrary to popular belief, improvisation was prevalent in the Classical and Romantic periods as well. It is well known that Mozart and Beethoven impressed their audience with improvised variations. Some of their published work might even be seen as written transcriptions of improvised performances. Another well established passage for performer's improvisation was the Cadenza in classical and romantic concertos, a section near the end of the concerto in which the soloist was free to recall and ornament previous materials and display his virtuosistic skills at will. Even if originally improvised, Cadenzas were gradually written out in full by composers during the XVIII and especially the XIX century.

Improvisation and virtuosic skills were definitely popular in the Romantic period. Some of the most eminent composers and pianists of their time were acknowledged as virtuoso improvisers, like Mendelssohn, Liszt or Paganini to mention a few.

On the contrary, classical Western music on the first half of the XX century markedly rejected improvisational practices, with notable exceptions in the United States, such as we see in some works by Charles Ives or notably Henry Cowell, a pioneer of open form techniques, as seen in his *Mosaic Quartet* (1935).

Several factors may explain the decline of improvisation practices within the classical Western music context. Composers during the 19th and 20th centuries gradually increased the amount of annotated performance details in their scores. And conversely, prominent performers advocated for a strict execution of the former repertoire in order to be as much faithful as possible to the composer intentionality - and also as a reaction to *former creative licenses of questionable taste* (Maria Calls, *come scritto*).

Non Western music improvisational practices

Improvisation is not only observed in musical cultures around the world, it is, compared to Western classical music practices, notably commonplace and socially highly valued. Improvisatory music is deeply rooted to the social context of a culture, and serves often a well defined social purpose in the community. It therefore must adhere to social conventions and procedures in order to be well accepted. Musically speaking, such conventions may pose stylistic and formal constraints which otherwise guarantee that the improvised music adheres to the idioms and norms of that particular style.

India Indian music, particularly Hindustany, puts heavy emphasis on improvisation. Because of its spiritual nature and its traditional oral transmission from master to apprentice, the tradition of Indian musical performance was less tied to theoretical monitoring than Western music [Bailey, 1993]

The improvisational frameworks in Indian classical music are generally known as raga, and exhibit a large variability. Each raga, which is commonly associated to seasons or times of the day, has distinct rules both for composition and improvisation. Such rules concern the salience of notes, registers, intervalic rules and ornamentation techniques in a melody. In a musical performance, an improviser begins first by introducing the defining characteristics of a raga to establish its particular mood, usually unmetered and over a drone, after that a more metrically structured section gradually builds up and eventually takes over the melody. Such melodic improvisation is known as Alapa (from Sanskrit “conversation”).

Having their origins in Northern India, the Romani people exerted a notable influence in many musical cultures around Europe and even the Middle East. Gipsy music, or music genres strongly indebted to the musical heritage of Romani people, display a strong improvisational character.

Flamenco The assimilation and cultural blending of gipsy music and the rich folklore in Andalucia and neighboring regions gave rise to a genre which incorporates improvisation in a combined setting of dancers, singers and instrumentalists. Its framework provides flexible resources for improvisation based on *palos*, flamenco styles that encompass specific modal and rhythmic traits, as well as defining melodic motifs and stylistic cliches. The ineffable, trance-like state which flamenco performers seek to achieve is commonly called *duende*. Reports by performers clearly state their awareness of the creative process in a flamenco performance. Creativeness, to flamenco players, is an essential purpose and lets them feel completely involved and identify themselves with the music being played [Bailey, 1993].

Javanese Gamelan Gamelan is an Indonesian ensemble of struck metalophones and drums. The very design of the instrumental ensemble is oriented towards collective playing [Susilo, 2001], and performers have a considerable amount of freedom, though within the strong constraints of the style. There are a number of Javanese terms to refer to improvisation in a Gamelan performance, such as *kembangan* (adding beauty), *isen-isen* (filling gaps) or *sambang* (joining a fellow performer who was lost in the performance). A general schemata allows performers to join even with no prior joint rehearsal, and very long performances assume that most of the time performers are actually keeping a stable set of patterns. As the space for performer's creativity is highly constrained and, indeed, seasoned performers do really prefer not to risk with too much spontaneity, some authors consider that Gamelan performance, as a whole, is hardly improvisational, even if players do make extensive use of subtle improvisational techniques [Sutton, 1998]

West African percussion The music from Central and West African regions is markedly rhythmic, centered on accompanied songs and deeply integrated into the social life of the community, qualities it shares with most of the sub-Saharian cultures. The stereotypical ensembles include percussion instruments with antiphonal choral singing, and dancers. In their performances extensive improvisation is commonplace, though constrained by stylistic boundaries [Locke, 1980].

Notably, West African polyrhythmic dance drumming shows an interesting hierarchical structure for improvisation. The leading drum is usually encouraged to display the highest level of creativity, while a stable, almost never modified pattern by the bell sets a common meter. The remaining supporting instruments improvise on their patterns through embellishment, syncopation, repetition and variation. The music focuses in those subtle, flowing variations on an stable, entrancing on-going beat.

Improvisation in Jazz

Jazz arguably brought back improvisation, or more precisely the social prestige to improvisation to the Western musical practices in the twentieth century.

To many, and despite the fact that improvisation was, as previously explained, still commonplace in certain social contexts, Jazz revolutionized the musical scene, eventually raising interest among high-art western composers and opening a period of rich cross-fertilization between such different musical traditions.

Dixieland and collective improvisation One of the first available jazz recordings is from the Original Dixieland Jazz Band, a New Orleans based quintet. Their first recording dates from 1917. Their style and sound, which was soon to be widely imitated by other bands, was innovative in many aspects: musical use of non-musical instruments (drum sets full of farmyard related sounds), novel uses of musical instruments (screaming, guttural and animal-like sounds) and a restricted use of improvisation. It contributed to establish the new genre and composed or popularized many tunes which are considered as jazz standards nowadays.

The most common sources of tunes in the early jazz bands were ragtimes, blues, military, religious or work-related music and original tunes as well. Classic New Orleans jazz established a typical instrumentation and instrumental roles, and was characterized by collective improvisation over simple tunes and harmonies. Except for the solos, everybody played and improvised together. Improvisational resources were mainly melodic and timbre-related.

Swing and soloist improvisers Jazz was incorporated as a popular danceable genre at the end of the 20s, with the development of Swing. The widespread of radio and recording highly popularized this new genre the next decade, epitomized by Benny Goodman and his orchestra.

Swing is characteristically orchestral, fast paced and oriented towards dance. The new genre run in parallel to the development of Big Bands, medium to large-sized instrumental ensembles consisting of woodwinds, brasswinds and percussion. During the *Swing Era*, as is known the 35-45 period, hundreds of such bands and specifically tailored performance venues spread across the States. Most of them reflected the individuality of their bandleader or lead arranger. Besides, big bands employed typical orchestral resources like instrumental sections and highly detailed score written arrangements.

In some aspects swing was less improvised than the earlier Dixie music: larger ensembles asked for a more homophonic style, with greater reliance on written arrangements, dance-oriented music posed higher constraints to improvisation, and incidentally, many brass players which joined Big Bands came from military bands and were more seasoned music readers than improvisers. On the opposite, the new genre favored a higher level or instrumental proficiency in terms of speed, agility and control. Improvisation was mainly carried out by soloists, who had allocated sections for their solos accompanied by rhythmic sections or background figures.

With the swing era, the Saxo become the most iconic instrument in jazz, with the saxophonist Coleman Hawkins as one of the most prominent swing soloists. He approached improvisation with a more harmonic perspective: instead of horizontal, thematic oriented improvisations, he advocated for a greater awareness of the underlying harmony by resorting to a wider use of up and down arpeggios. Ultimately, Hawkins is considered an important antecedent of Bebop.

Bebop and jam improvisation As swing was quickly established as a somewhat rigid, highly standardized mainstream genre, a growing number of jazz musicians felt it was too artistically constraining. Swing just didn't devote enough space for improvisation and both harmonies and rhythms had become too cliche.

As a reaction, over the 40s musicians began meeting in small clubs to play in informal, unrehearsed performances which were known as *Jam sessions*, giving rise to a new genre which was to be known as Bebop. Ultimately, Swing would decline with the Second World War.

It opposed Swing in many aspects, for it was focused on listening instead of dancing and had a more artistic ambition, compared to the entertainment focus of swing. It also favored smaller groups, such as a quintet, and showed a higher refinement and complexity both rhythmically and harmonically, with extensive use of virtuosic tempi and highly cromatic resources.

Bebop is targeted to improvisation, and thus the solos are the primary focus of a performance.

It was both consequence and a cause of removing jazz from the popular mainstream music. To the audience, the new genre was considered complex and furious, revolutionary and somehow addressed to *connoisseurs*. Bebop provided the standard vocabulary to jazz. In a way it sets the common practices and idioms to the genre. Figures like Thelonious Monk, Charlie Parker or Dizzy Gillespie were among the most influential bebop artists.

Free jazz Ultimately, if Swing ended up being too cliche, bebop reached a similar artistic dead end because of excessive sophistication.

Free jazz originated in the fifties as an attempt to break down the formal and harmonic conventions of jazz and bring back its essentially primitive, untamed improvisational character which was considered as largely lost within the Bebop aesthetics and practices. It must be noted that free jazz, however, retained always enough characteristics of jazz as not to completely depart from that genre.

At that time, jazz performance adhered to strict and predetermined, formulaic practices: both song structure, time allowed for improvisation and harmonic patterns were clearly delimited. Although such structures serve as a structural and stylistic framework for improvisation, eventually jazz reached an aesthetic *cul-de-sac* in the United States [Menezes, 2010]. Musicians searched for new musical paths and explored resources and techniques from XXth century avantgarde music, such as atonality, microtonality or aleatoric processes.

The term Free Jazz was quickly adopted by the African-American musical community. It conveyed as well The pioneer of free jazz is considered to be Ornette Coleman, who was at his time truly revolutionized the jazz scene with melodic based (instead of harmonic based) improvisations abruptness and aggressive textures which eventually became idioms of the new genre.

A notable contribution of the free jazz scene was bringing back the interest for collective improvisation in small groups, in sharp contrast with rigidly allocated improvised solos so characteristic in classic jazz. Take as an example the revolutionary album *Free Jazz by the Ornette Coleman Double Quartet* (1960). The performance primarily depends on collective improvisation: as it lacks a predefined structure, players must learn to react to the stimulus from other players, moving jazz away from the solo performer to a sort of collective conversation. As Ornette himself points out

> The most important thing was for us to play together, all at the same time, without getting in each others' way, and also to have enough room for each player to ad lib alone (...)
>
> Free Jazz by the Ornette Coleman Double Quartet (1960), liner notes

As free jazz players refused the structural constraints of the rigid jazz forms and harmonic progressions, some put greater emphasis on extended instrumental techniques as well (as in Pharoah Sanders or John Coltrane). In the search for a more primordial and authentic voice, free jazz was characterized by a growing interest in the earliest jazz styles, in African music and in primitive or exotic musical instruments and traditions.

Among some characteristics which free jazz retained from traditional jazz we should mention the notion of swing, at least as a flowing, meterless pulsation, and idiomatic melodic gestures reminiscent of blues. But beyond such loose references, performers embraced a broad, often heterogeneous range of innovative techniques, from the primitivism and physicality of Cecil Taylor, who incorporated novel harmonic resources to the piano improvisations, to the extended techniques and microtonal improvisation practices of Albert Ayler or the adaptation of Indian aesthetics and techniques by John Coltrane, just to mention two of the most relevant figures from the free jazz scene.

Free improvisation

The European, particularly the british free jazz scene quickly approached free improvisation in the mid 60s, incorporating developments from avantgarde classical composers and moving even further from the jazz tradition. Many of those performers coming from free jazz gradually approached more abstract realms. Though deeply influenced by their jazz background, free improvisers perceived functional harmony, tonality or swinging rhythm as constraints that had to be overcome.

Free improvisation, if partially, differs from free jazz in its aesthetic aims. A typical free jazz collective improvisation may consist of the collective attempt to establish a narrative out of the individual voices of performers [Menezes, 2010]. On the contrary, free improvisation not only departs from stylistic and cultural narratives but from the ego itself, with players ceasing to function as individual actors. A free improvisation is much more deconstructed and less linear than a jazz improvisation, with each musician's part being uncomplete by itself, though distinctively recognizable (as in the SME example) or, in some extreme cases, ultimately submerged in the collective sound (as in the AMM)[Scott, 2014].

Seated in London, the Spontaneous Music Ensemble, is regarded by many as one of the quintessential examples of free improvisation. The SME, which around 1967 comprised most of the leading figures in the London free improvisation scene (Wheeler, Rutherford, Watts, Parker, Derek Bailey, Barry Guy, Stevens), put great emphasis on collective awareness and interaction in extremely open and leaderless performances. Not unintentionally, their fast but quiet music, full of subtle and quick nuances is sometimes reminiscent of Webern.

John Stevens, founding member of the SME, documented his methods in his Search and Reflect book, which provides exercises and musical games intended to facilitate collective creativity. Maybe no one better than himself to summarize the essence of free improvisation

> Being in tune, as close as possible, with all the people that are around you and at the same time contributing within that and never contributing to the extent that you couldn't hear what the other people were saying. So nothing you had to say was more important than an awareness of the whole.
>
> John Stevens, in interview with Richard Scott, August 1987 [Fadnes, 2015]

Ensembles such as AMM adopted techniques from minimalism and chance music as well, combining extended instrumental techniques with electroacoustic textures. The band rejected rehearsals or planning at all, advocating for extremely spontaneous long-form improvisations. This preference for extremely spontaneous improvisations was epitomized by the Company Week festivals organized and supported by Derek Bailey since 1979, promoting the joint performance of musicians in *ad-hoc*, unplanned ensembles.

Free improvisation is also known as non-idiomatic improvisation, as an statement of an explicit desire to avoid references to existing genres and musical styles, particularly jazz. It is difficult to argue, however, to what extent free improvisation is actually free of idioms and truly outside of any other historical or contemporary genres [Berlin, 2014].

Improvisation and Live Electronics Nowadays a multiplicity of improvisational scenes flourish, resorting to all sort of novel electronic musical instruments or plain electronic devices which are used in a musical context. The broad term of live electronics encompasses a wide range of sub-genres and aesthetics, but nonetheless they typically incorporate free improvisation practices.

Electroacoustic improvisation deeply focuses on extended techniques and timbral exploration. It is indebted to the pioneering electroacoustic improvisations by Cage , David Tudor and others, and by the collective improvisations of groups like the aforementioned AMMA, which incorporated electroacoustic devices in their performances.

Noise music has its roots in the Futurist movement , as well as the experiments and theorisations of *Musique Contrète*. Cage, interested in the musical use of extra-musical materials, pioneered the creative use of devices such as radio receivers, cartridges or amplified everyday objects in music performance. From the 60s several collectives incorporated all kind of electronic noises in their works and performances, usually with a strong improvisational and exploratory approach, as in Musica Elettronica Viva or in the Fluxus collective.

Deejaying has evolved as a well established genre encompassing performance, production and improvisation in a number of popular electronic music styles such as hip hop and techno. Beyond the secondary, functional role of mixing music for a party, a Disk Jokey can be a performer himself. He may embellish pre-recorded tracks with gestural uses of filters and crossfades or scratching techniques, rework them in a live performance, looping fragments and incorporating additional drum machines and samples to the mix. Ultimately, some DJs are true virtuoso performers of their turntables, attract huge audiences and even DJ competitions take place regularly. Because of this, it has evolved as a highly stylized genre.

Circuit Bending advocates for creative musical use of all sort of customized electronic devices, and it was pioneered by Reed Ghazala on the 60s. The Circuit Bending community thus inherits practices and aesthetics from chance music, performance and noise music, but with a strong emphasis in the Do-It-Yourself culture. Electronic toys and electronic toy instruments are a particularly preferred device suitable for bending. Practitioners are encouraged to modify the circuitry of such devices by trial and error and with no prior engineering knowledge, in order to discover unexpected sonic behaviors. This work, often done in collective workshops, results in unique musical instruments which are subsequently used in collective improvisatory performances.

Live coding is an improvisational practice well established in the field of *intermedia art*. Commonly utilized by a number of computer ensembles as well, it is based in on-the-fly, interactive programming of musical processes. It combines algorithmic composition to generate low level musical structures with collective compositional and improvisational practices such as concurrent programming and code sharing. The code as it is being written is almost always projected as a visual feedback -at least for computer literated audience. Live coding performances usually incorporate algorithmically generated visuals as well. The focus in the performance is thus not only in the result but especially in the musical process of writing and tweaking generative algorithms. As opposed to some computer music performance practices, the audience can clearly see what and how performers shape musical processes up to the very last detail.

2.4.3 Conclusion

In this chapter we introduced the central topic of musical creation as it manifests in composition and performance, both in score-based performances and in improvisational contexts. We discussed the responsibilities and creative opportunities for performers, and how their roles may be mediated by a musical score of a conductor.

We have seen that the challenges in the social roles of composers and performers in Western high-art music are just a consequence of the fossilization and ritualization of the repertoire, a phenomenon shared by other musical cultures as well. Regarding score-based performance, we discussed the social and aesthetic constraints that a performer faces when interpreting a musical score, and to what extend a performance may incorporate the creative contribution of the performer.

Such opportunities and constraints are present in improvisational contexts as well. What do different musical traditions teach us regarding improvisational practices? We may observe that all of them are based on flexible frameworks which provide a stylistic reference and a set of resources to the improviser. The improvisation itself may take place in many distinct musical dimensions (modes, rhythms, intervals). We may notice as well the prevalence of ritualized performative setups, they serve as a framework to performers too. A repertory of resources for the performer or improviser constrain and actually guarantee stylistic coherence and social acceptance.

But beyond what is generically called idiomatic improvisation, new musical genres advocated for improvisational paradigms which defy the traditional frames of musical style. We observed how free improvisation let performers get rid of stylistic constraints and unfold their creative resources without resorting to a repertory of musical idioms.

Finally, we analyzed the role of the conductor, highlighting how this role may actually promote a better creative involvement in the ensemble, both in score-based performances and in conducted improvisations.

In the next chapter we will focus on the performative environment which this thesis addresses, computer music ensembles, and we will how the aforementioned aesthetic and social challenges are re-visited and re-contextualized. As we will see, the genesis of computer music ensembles opens new opportunities to redefine the aforementioned roles and responsibilities of composers and performers.

Chapter 3

MUSIC CREATION AND PERFORMANCE WITH COMPUTER ENSEMBLES

3.1 Introduction

While Chapter 2 introduced the development of aesthetic frameworks for musical creation through history and different cultures and musical styles, this Chapter will be framed in the performative context this Thesis is based on: Computer Music Ensembles, and particularly Laptop Orchestras.

A first section provides an historical context for the genesis of Network Music, briefly introducing the earlier developments around electronic and computer music and reviewing the relevant contributions of the pioneers of this genre.

The next sections provide a generic definition of Multi-user instrument and review systems for topological and functional characterization, to elaborate descriptive taxonomies and to provide consistent design practices. Closely related to design criteria, we review the research in evaluation techniques most relevant for Multi-user instruments, both from an usability, Human Computer Interaction perspective and from a user experience perspective.

Finally, we will approach the worldwide phenomenon of Laptop Orchestras and survey the most relevant research on them, as well as reviewing the different ensemble models and compositional strategies which Laptop Orchestras adopted. We will see how this process of reflection and applied research contributed to forge an ensemble identity for Laptop Orchestras.

3.2 Computer Music Ensembles. Overview

This section will provide an overview of the historical landmarks that contributed to the emergence of Computer and Network Music. We will then focus in the definition, contextualization and throughout characterization of multi-user instruments.

3.2.1 Electronic music overview

Early developments and Aesthetical foundations

The first really influential electronic musical instrument, the Telharmonium appeared in 1906, with its broadcasted performances from New York. It was an electronic organ of huge proportions capable of distributing its signal through telephone lines[Collins and D'Escriván, 2007]. Utterly impressed by the possibilities of the Telharmonium, the Italian composer F.Busoni wrote the prophetic essay The Sketch of a New Aesthetic of Music (1907) where he envisaged the potential of new technologies to forge new musical possibilities:

> I almost think that in the new great music, machines will also be necessary and will be assigned a share in it

The origins of electronic music are indebted to the technological advances in sound generation, recording and reproduction as much as to the novel aesthetic perspectives brought up by the new creative medium. Busoni advocacy for the new music lead him to support the artistic Italian collective known as the Futurists. Another key referent in the new aesthetics, which eventually permeated through to concert music in the 20s, was Luigi Russolo's Futurist manifesto The Art Of Noises (1913), the first attempt to categorize all sounds and actually treat them as potential music[Collins and D'Escriván, 2007].

Simon Crab, in his exhaustive historical review of the development of electronic musical instruments [Crab, 2004] highlights three distinct driving forces in the early era of electronic music, notoriously prevalent from the end of the XIXth century to the Seconds World War. The first was the technological utopianism widely embraced at the turn of the XIX century, which favored a view of science and technology as a liberating force. This radical optimism translated in the musical domain to an interest towards the study and development of new tunings, new timbres and new musical instruments which could liberate the artist from the conventionalisms and limitations of tradition. A second related force originated within the musical community itself. For some composers, incorporating technological means into the compositional process was, as it allowed for ultimate precision and strictest compliance to the score[Roads, 1996]. Technology might let composers eventually get rid of the performer itself, a role -as we seen in Chapter 1 - specially questioned in the post-romantic era. Finally, Crab mentions the often underestimated role of the radical ideological and political changes in the early XXth century, from National Socialism to Soviet Communism, in the promotion of avant-garde and futuristic ideas, within the music domain as well - even if both regimes eventually rejected it.

With the introduction of the vacuum tube (1906) many other electronic instruments, of smaller scale, were developed over the next two decades. Instruments such as the Theremin (1919-20), Ondes-Martenot (1928), the Trautonium (1928) or the Hammond Organ (1929), just to mention a few, brought new timbrical possibilities and novel modes of interaction. Despite how much innovative such instruments and technologies were, however, it took a time to shift from its recreational use to the realization of the visionary ideas of Busoni and the Italian Furutist movement. As Collins points out

> During the early years of the twentieth century, the musical spirit that would ultimately form into 'electronic music' found itself stultified and frustrated by the limitations of both the technology and the conventions of the day [Collins and D'Escriván, 2007]

We may exemplify this point with the rather traditional use of the aforementioned instruments by European composers, notably Ondes-Martenot, incorporated into some compositions by Messiaen, and the Trautonium, for which Paul Hindemith wrote several short trios. Varèse quickly incorporated those instruments into his orchestral works as well. Indeed Varèse, deeply influenced by Busoni's aesthetic beliefs and in close collaboration with the electronic instrument inventors of his time, exemplifies such tensions between the utopian promises of Futurism and the limitations of his time

> Organized sound seems better to take in the dual aspect of music as an art-science, with all the recent laboratory discoveries which permit us to hope for the unconditional liberation of music, as well as covering , without dispute, my own music in progress and its requirements
>
> Edgar Varèse (1940), as quoted in Thom Holmes' Electronic and Experimental Music: Technology, Music, and Culture [Holmes, 2012]

Post-war developments and landmarks

Technologies for sound amplification and sound recording, also developed through the first half of the XXth century, had a major influence in the electronic music field. Acquainted with the technology of radio broadcasting, the french composer

Pierre Schaeffer founded the *Groupe de Recherche Musicale de Musique Concrète* (later the GRM) in Paris in 1951 to explore the musical possibilities of the new sound recording and editing techniques. Schaeffer termed his compositional procedures Musique Concrète, referring to the musical use of non-musical, everyday sounds abstracted into a composition through elaborated tape editing techniques. Schaeffer was a prolific writer as well, and his essays on musical objects are still influential.

At the same time, Meyer-Eppler founded the Electronic Music Studio of WDR in Cologne. It was equipped with primitive sound generation devices and some precursors of the synthesizer, such as the Monochord and the Melochord, and Meyer-Eppler in his thesis conceived the idea of sound entirely realized from electronic signals recorded to tape, a procedure which the composer termed *Elektronische Musik*. Stockhausen, who attended Schaeffer's GRM studio in 1952, offered the first concert of tape music in Cologne in 1954, including *Studie I* and *Studie II*, credited as the first electronic music works entirely composed for sine waves, combining integral serialism, electronic tone generators and tape editing techniques. *Studie II*, on the other hand, is considered a landmark in the history of electronic music, being the first notated electronic score.

Electronic music studios proliferated in Japan and in the States as well, notably the Columbia-Princeton Electronic Music Center administered by Vladimir Ussachevsky. A remarkable contribution of the CMC was the development of the RCA Mark II Sound Synthesizer (1958), the first digitally controlled analog synthesizer, which allowed composers to sequence an score, thus liberating composers from laborious manual tape editing procedures.

Film music contributed to the popularization of electronic musical instruments, and electronic soundtracks share the aforementioned gradual adoption of a new aesthetics consistent with the sonic palette and musical possibilities offered by the new medium. We must mention here Louis and Bebe Barron's soundtrack for Forbidden Planet (1956), credited with being the first completely electronic film score.

Computer Music

The rapid development of computer technology after the Second World War fostered their musical use. Indeed, simple musical recreations were already carried out with the first-generation valve-driven computers in the 50s, such as Australian's CSIRAC and Manchester's Ferranti Mark 1 computers.

But composers soon realized their potential as a tool to aid in the formalization of musical processes, an attractive prospect for post-serialist composers whose compositional techniques could benefit of automated computation [Roads, 1996]. Lejaren Hiller, in collaboration with Leonard Issaacson, wrote in 1957 the Illiac

Suite, credited as the first computer generated musical score [Isaacson, 1959].

Simultaneously, a different approach was carried out: the use of computers for sound synthesis purposes. Working at Bell Laboratories, Max Mathews inaugurated the era of digital sound synthesis with the development of the influential MUSIC software in 1957, the first computer program which could generate digital audio through direct synthesis. MUSIC had a number of descendants, encompassed by the name MUSIC-N, which followed the same modular paradigm and shared similar approaches in the division between audio and control computation rates and sound synthesis and musical score specifications. Notably, most of the current software tools employed nowadays in computer ensembles are indebted to the MUSIC-N heritage [Boulanger, 2000].

Mathews summarizes some of the strong points of computer music: precision, exact reproducibility, and the ability ho handle extreme complex specifications with sufficient flexibility [Mathews, 1963]. It should be noted, however, that both computer aided composition and especially sound synthesis over the first decades working with computer mainframes was a lengthy process, requiring hours of costly computation and eventually migrating the data to another computer to perform the digital to analog audio conversion [Roads, 1996].

3.2.2 Network Music

In the beginning of the second half of the XX century, the San Francisco Bay Area became a particular microcosm of research clusters, actively promoted by the Stanford university. They were closely tied to the emergent technology industries such as semiconductor manufacturing, computers and telecommunications.

The nickname Silicon Valley, entitled in 1971,was quickly adopted to designate the southern part of the San Francisco Bay Area. Just three years later, in 1974, Intel released Intel 8080, one of the first really widespread microprocessors, Apple I went on sale in July 1976 and IBM, who had created the first transistor-based mainframe computer in 1958, introduced the Personal Computer on 1981.

Telecommunications developed in parallel. Thus, ARPANET was conceived and implemented in the late 60s. In the early 1970s, professor Vinton Cerf's networking research group at Stanford developed the TCP/IP protocols. Additional protocols followed, such as the electronic mail (1972) and the FTP protocol (1973).

As for computer music is concerned, some of the major centers devoted to the research and creation with computer related technologies were founded in that area. We must mention here the Center for Computer Research in Music and Acoustics (CCRMA), founded by John Chowning in 1975, or the University of California at San Diego. Around those academic institutions composers and engineers worked together in languages for sound synthesis, in computer assisted music composition techniques and in development of new Dsp hardware.

The development of personal computers encouraged an active and collaborative musical scene. Composers helped each other to learn about programming microcomputers, an attitude which built a strong feeling of community. Those pioneers embraced with enthusiasm new approaches in interactive algorithmics and live computer performance, by incorporating the open creative practices from the experimental music scene from the 60s and the informal, communitary DIY approaches which characterized the emerging Silicon Valley garage-era.

In 1977 Rich Gold and Jim Horton offered at Mills College the first public performance of network music, by linking two KIM microcomputers, each with distinct generative algorithms, which exchanged musical data through parallel ports or direct interrupt lines. Lately John Bischoff and David Behrman joined and created The League of Automatic Music Composers, what is considered the first computer band. The band performed regularly as a trio with Tim Perkins, who joined them in 1980. Their performances were a community event with an openly didactic purpose, because of the common interest in electronic instrument building. The League kept its activity till 1984.

The band started with unattended interactive performances but soon its members joined the computer setup to tweak and adjust parameters in realtime. Their concerts didn't properly consist of pieces but of different network setups which were often build in a down-top approach, from very idiomatic compositions for a single microcomputer which were interconnected after discussing together the possible interrelations and links between their otherwise autonomous processes.

As the musicians and the audience noted, the band somehow resembled a very idiosyncratic free improvisation.

> (The League) sounded like a band of improvising musicians. You could hear the communication between the machines as they would start, stop, and change musical direction. Each program had its own way of playing. I hadn't heard much computer music at the time, but every piece I had heard was either for tape or for tape and people, and of course none of them sounded anything like this. I felt like playing, too, to see whether I could understand what these machines were saying.
>
> George Lewis quoted in "Composers and the Computer", p. 79, by Curtis Roads, William Kaufman pub. 1985

> It just knocked me out. It was electronic, but it had this feeling of improvised music, that everyone was talking at the same time and listening at the same time, and I thought, "How did they do this?"

> Chris Brown, after hearing The League of Automatic Composers playing on KPFA Radio, 1982 in [Chadabe, 1997].

This extremely open an non hierarchical approach in computer ensemble performance embraced the communication protocols as well. Performers were free to interpret musical control data received from other computers, even directly converting them to audio signals.

The same late members of the League Chris Brown and Jim Horton joined efforts again in 1986 during the Network Muse Festival, which took place at The Lab, in San Francisco, and which was devoted to small networked computer bands. The very name of the Hub emphasized a new approach to handle interconnection by means of a central computer which acted as a mailbox, a shared storage to allow for an easier way to setup and join in the performances.

The new infrastructure provided means for easy expandability. In 1987 the Hub offered in New York the very first remote performance linking two identical Hub setups between through a modem link. Despite the public notoriety of their remote performance, the Hub kept their interest in co-located, tightly interdependent setups which explored the emergent properties of interconnected musical networks, more than keeping up with the latest technologies and sound synthesis techniques.

> I see the aesthetic informing this work as perhaps counter to other trends in computer music: instead of attempting to gain more complete control over every aspect of the music, we seek more surprise through the lively and unpredictable response of these systems, and hope to encourage an active response to surprise in the playing. And instead of trying to eliminate the imperfect human performer, we try to use the electronic tools available to enhance the social aspect of music making.
>
> Tim Perkis' liner notes to The Hub's first CD (1989 Artifact Recordings 1002)

Eventually, at 1990 the Hub switched to a MIDI patchbay to route control messages between computers. Instead of a shared memory space, a MIDI patchbay only allows for message routing without possibilities for storage. Also, it was up to the sender instead of the receiver, the decision to place a certain message to a given computer. As MIDI is stream based, lost messages could not be recovered, but the hardware allowed for higher data rates and the protocol allowed for flexible one-to-one or broadcast messaging.

The pioneering work and prophetic views of the League anticipated much of the practices which are currently commonplace in computer-based music ensembles. Their setups were environments in which humans are computers were sharing musical responsibilities, and where a complex network of interdependences between them all contributed to shape the musical result. They explored a multiplicity of network topologies and performative paradigms, facing and reflecting on much of the challenges and opportunities that the new genre of computer music provided.

3.2.3 Multi-user instruments

This section will provide an in-depth review of possibly the most paradigmatic contribution in the field of network music : multi-user instruments.

First, we will first provide a generic definition of the term and put it in historical context, citing some relevant analog precursors.

Next, we will introduce the basic architecture of a Digital Musical Instrument to follow with a review of the taxonomies for multi-user digital musical instruments proposed in the literature on the subject.

We will finish this section with a survey on the design and evaluation criteria for musical interfaces, with a specific focus on the closely related contexts of collaborative interfaces and multi-user instruments.

Definition

A Multi-user Instrument is a device, or collection of devices perceived as a unity, which is conceived to be operated by multiple performers and which is capable of producing sound in a musical context.

Multi-user Instruments are particular use cases of musical instruments. A Musical instrument is, in a broad sense, any device capable of producing sound. We acknowledge such devices as musical instruments, though, when we perceive them as specifically tailored to musical performance. Traditionally such musical instruments are designed to require a single performer to be played, though some of them may be occasionally played by more than a single musician, such as in four-hands piano or guitar pieces. There are also a few instruments which are traditionally played by two or more performers. A compelling example could be the Txalaparta, a traditional basque percussion instrument in which two performers play in a tightly coordinated way on the same instrument.

Even historically, though, there are some examples of musical instruments requiring coordination between two or more ”players” to actually work. An extreme example would be Halberstadt organ (ss XIV), a large pipe organ installation which required a performer on the keyboard and up to ten men blowing its twenty elbows.

But monolithic, single-entity devices are not the only ones which may be considered as musical instrument entities. If metaphorically, an orchestra conductor is said to have the orchestra as the instrument at his hands. Indeed, his commands turn into performative actions, though mediated by performers who individually address their own musical instruments, such that a whole orchestra can be effectively considered a Multi-user Instrument from the conductor's perspective.

Analog Precursors

Cage's *Imaginary Landscape No. 4* for 12 radios, 24 performers and conductor (1951) is considered one of the earliest examples of interdependent networked music instruments. The score consisted of precisely timed indications for frequency dials and volumes which were operated by pairs of performers. On one hand it was one of the first examples of chance music, where the actual sounds coming out of the radio speakers were unpredictable and dependent on time and location. On the other it was a collective exploration of a sonic space, albeit strictly guided by a score. And finally, there was an interdependence of performers as the sound was simultaneously modified by two players at the same time, and one player could, say, raise the volume while the other was reaching a silent radio frequency and vice-versa, thus both controls were simultaneously altering timbrical parameters of the same instrument.

Stockhausen's *Mikrophonie I* for tamtam, 2 microphones, 2 filters, and controllers (1964) is considered a pioneering work in the genre of live electronics, and one of the first compositions utilizing microphones as musical instruments. The instrumental setup consists of two trios of percussionist, microphone player and filter operator. The trios usually alternate except for the tutti sections where all play together. While the percussionist plays the tamtam, another player captures and amplifies subtle noises through a hand-held directional microphone, while another player filters the sound coming out of the microphone and distributes it into a quadraphonic system. Each trio of percussionist, microphone and filter constitutes an interdependent unit, serially connected. As the tamtam is eventually played by the two percussionists at the same time and its sound is picked up by the microphones, we could consider the whole setup constitutes a complex multi-user instrument with serial dependencies.

A primer on Digital musical instruments

After defining what a multi-user instrument is and having illustrated the term with two relevant analog precursors, we should present now the Digital Instrument model, still on an individual basis, while the next section will be devoted to multi-user digital instruments.

This section therefore will introduce the anatomy and principles of Digital instruments, before delving into the new paradigms of multi-user and collaborative digital musical interfaces that naturally emerged as a consequence of the historical achievements of the aforementioned computer and network music pioneers.

Generically speaking, musical instruments whose sound is computer generated are called digital musical instruments (henceforth, DMI). In order to drive the musical parameters of such a digital sound generator, the instrument must incorporate a *gestural controller* or *control interface* to allow performers to interact with it. The most basic DMI model therefore assumes a bi-modular architecture which consists of a Gestural Controller module and a Sound Production module. The control interface may be a gestural controller, a control surface or whatever device with which the performer interacts physically. By means of a mapping layer, the acquired human actions are mapped into parameters which drive the sound generator device. This basic structure is illustrated in Figure 3.1

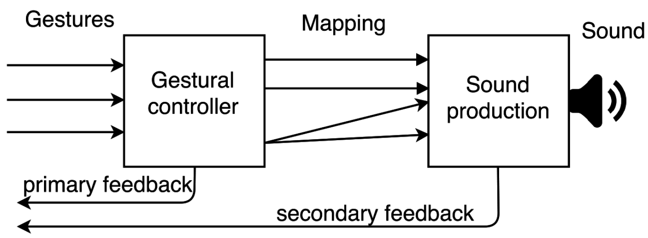


Figure 3.1: Structure of a Digital Musical Instrument

In this respect, a DMI takes to the extreme the decoupling between performative gesture and sound generation which we already observed in early electronic musical instruments. Most of the current research in DMI design and evaluation in the NIME community is closely related to the implications both for the performer and for the audiences of such inherent decoupling.

The design of a DMI is usually addressed according to the aforementioned bi-modular model: first, we decide what gestures will be used, and which sensing strategies will optimally capture them, then we must select the appropriate sound synthesis algorithms and choose the most suitable implementation, once both modules are functional we must implement the mapping strategies to link the gesture capture to the sound generation and, finally, we must decide on the eventual additional feedback modalities that our DMI might provide -apart from the sound [Bryan-Kinns, 2004].

Both the gesture acquisition and the sound generation modules may be designed to mimic existing instruments, or may try to extend or augment them or may ultimately be based on completely new performative and timbrical paradigms.

Digital replicas of acoustic instrument may benefit from the acquired knowledge of their instrumental identity (both for the performer and for the audience point of view), while alternate designs ask for new performance abilities but open new paths to musical expression without the constrains of mechanic construction, convention and tradition. Typically, therefore, DMI may be classified as augmented musical instruments, instrument-like gestural controllers, instrument-inspired gestural controllers and alternate gestural controllers [Miranda and Wanderley, 2006].

The Mapping Layer While an in-depth discussion of the components of a DMI is out of the scope of this Dissertation, it is necessary to briefly discuss the mapping layer as it takes a key role in the formalization of our creativity metric, discussed in Chapter 7, and in the design of the performative environments developed in Chapter 8.

It is accepted that, due to the separation between gesture input and sound output, the choice of mapping strategies lie at the essence of digital musical instruments[Rovan et al., 1997]. As a consequence, a great deal of research has been invested in studying the implications of mapping and determining the optimal mapping strategies from different perspectives.

Mapping strategies may deeply impact both in musical behavior and performer's attitude towards an instrument. In an extensive experimental research aimed at investigating the importance of the mapping layer, Hunt et al. -see [Hunt and Wanderley, 2003] and [Hunt et al., 2003]- concluded that complex mappings are most suitable for all but the most simple musical tasks, both in terms of expressive potential and performative engagement. One-to-one mappings should be, according to the authors, restricted to a higher semantic level, more abstract and perceptually oriented mappings.

For example, we could map quantity of movement (itself, a complex mapping of discrete qualities of gesture) to sound brightness, which in its turn would be mapped to several synthesis parameters at once). This perspective lead researchers to propose two- or even three- layered mapping schemes for a DMI (see Figure 3.2). Wessel already introduced the concept of high level timbre as a control structure in [Wessel, 1979], this approach was followed by Wanderley [Wanderley et al., 1998], one of the key researchers in the systematic study of mapping in DMIs.

This multi-layered mappings provides better independence from the actual gestural controllers and sound synthesis algorithms involved [Hunt et al., 2003].

As for the actual mapping typologies, several categorizations are possible [Arfib et al., 2002]. In short, a mapping layer should be studied from different perspectives, summarized below:

- **Cardinality** Mappings may be one to one, convergent (many to one), diver-

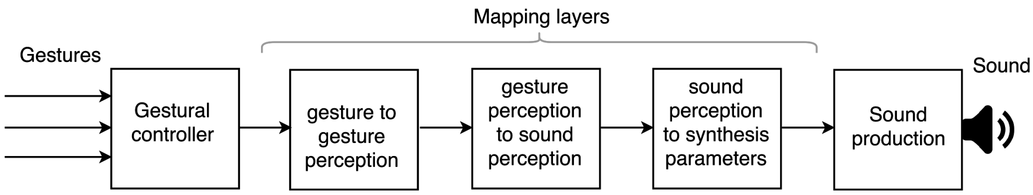


Figure 3.2: Structure of a Digital Musical Instrument implementing a three-layer mapping model.

gent (one to many)[Rovan et al., 1997] or more complex (many to many).

- **Explicit/implicit**. Mappings may be explicitly formalized or stated through generative rules or databases, or may be implicitly encoded within generative algorithms such as neural networks.

- **Static/dynamic**. A mapping may be static or dynamic. An example of the former are adaptive mappings.

- **Linearity** Mappings may exhibit a nonlinear behavior and unpredictability. Such non-linearity may be incorporated to emulate the characteristics of acoustic instruments or be the result of higher-level processes within the gesture mapping layer (like gesture recognition and tracking, or extraction of gestural descriptors).[Jordà, 2005a]

From DMIs to Multi-user DMIs The last decade, the convergence of computer-based digital musical instruments and their native capabilities for interconnectivity greatly facilitated the extension of DMIs to an ensemble context, up to an extend that many of the aforementioned practices of network music pioneers became commonplace.

Nowadays, computer music ensembles in which digital instruments interact through networked communication do allow us to design computer-mediated social environments which can settle anywhere between the continuum from purely isolated musical instruments to interdependent networks or completely shared spaces for music performance.

The next section will provide an overview of the taxonomies proposed to categorize the broad range and variety of multi-user instruments which have been being developed over the last decades.

Functional and descriptive taxonomies for multi-user DMIs

Multi-user instruments for computer networked ensembles are usually addressed by researchers as particular cases of the broader fields of interconnected musical devices or environments for collaborative music making, both not specifically oriented to live music performance.

Weinberg analyses the social implications of the structural topologies in *Interconnected Musical Networks* [Weinberg, 2003]. He proposes a taxonomy based on the central goal the musical network tries to accomplish, dividing them into process-centered networks, and structure (or outcome)-centered networks. The first ones offer less direct control by providing an evolving context where rules and/or performers determine the musical outcome, whereas the second ones are designed to optimally achieve a preestablished score or plan.

He further classifies such networks according to the main motivations for interaction and the player's experience the network is centered on, and highlights the challenging balance required when designing a network with both goals in mind, process and outcome. Regarding the level of centralization in the control and the amount of equality between peers, the author states that hierarchic networks with a fixed leadership would be more effective for structure-centered motivations than democratic networks.

By the other hand, the author analyses the network design from the perspective of its social organization, and observes that synchronous networks may better support truly immersible interactive environments but may be more challenging to maintain individual autonomy, whereas sequential, turn-based networks might challenge the group cohesiveness, being the balance between the two architectures, the topology, direction and amount of interaction and idiosyncrasy of roles of paramount importance in the design of the social experience. For the specific case of small, co-located ensembles, which Weinberg calls *Small-Scale collaborative musical networks* [Weinberg, 2005], a higher and subtler lever of interactivity is expected to promote more intricate musical collaborations.

Barbosa [Barbosa, 2003] suggests a time/space axis to classify collaborative music systems which is inspired in the Computer Supported Cooperative Work literature. He makes a clear distinction between four different collaborative music systems taking into account how actors interact: either co-located or remotely, and either synchronously or asynchronously: Local interactive music systems, shared sound environments, music composition support systems and remote music performance systems.

More from a HCI perspective, Jorda [Jordà, 2005b] proposes a number of distinctive features in multi-user instruments from the point of view of complexity and freedom in shared collective control: namely the number and flexibility of number of players, the assignment of roles (number of roles per player, whether they are dynamic, even throughout the performance, or not, the possibility to duplicate or leave roles unattended) and the topology of interdependences between performers (kinds of mutual interaction, hierarchical/horizontal interdependences, individual or shared *threads*). He also questions the term *multi-user* if no mutual interaction takes place, as it equals just the sum of the individual contributions.

Booth [Booth, 2010] identifies two key factors to evaluate interdependent systems: *inclusivity* (or degree of cross-platform interconnetivity via open protocols) and *mutability* (or easiness to reconfigure the network topology, such is the nodes and their interconnectivity, even dynamically). The author proposes a matrix representation to visually identify archetypal models of interconnection in parameter sharing scenarios, in which each cell $w_{i,j}$ summarizes the degree of influence of player i over player j. The three tables 3.1 provide illustrative examples for, respectively, an ensemble of independent instruments (A), an ensemble driven by a single performer (B) and an interdependent, equally balanced shared environment (C).

Table 3.1: G.Booth's Archetypal Models of Interconnection for multi-user instruments, and their corresponding interdependence metric.

(a) Independent model. $im = 0$

1	0	0
0	1	0
0	0	1

(b) Dominant model. $im = 2/3$

1	1	1
0	0	0
0	0	0

(c) Shared model. $im = 2/3$

.33	.33	.33
.33	.33	.33
.33	.33	.33

As shown in the former figure, we proposed as well a simple metric to evaluate the mean level of interdependence in multi-user instruments, which may be readily computed from such matrix representations: the ratio of parameters (and amount of them) controlled by a performer versus the sum of all parameters in control. Formally

$$m = 1 - \frac{\sum_i w_{ii}}{\sum_i \sum_j w_{ij}} \tag{3.1}$$

While the proposed *interdependence metric* does not tell us how directed or unequal such dependency is, it may be useful to summarize in a single metric the amount of control being given to other players, being 0 its value for independent instruments and $1/n$ its value for a balanced, fully shared environment for n players. We will employ this metric to comparatively evaluate the multi-user instruments employed in our experiments, in Chapter 8.

Another systematic analysis of collaborative musical performance systems is proposed by [Hattwick and Wanderley, 2012], offering a multidimensional space representation derived from previous work by the authors [Birnbaum et al., 2005]. The proposed dimension space (see fig 3.3) may serve to elicit valuable clues on agency, hierarchy, mutual awareness and interdependence topology to quantitatively evaluate and compare the architectures of collaborative systems.

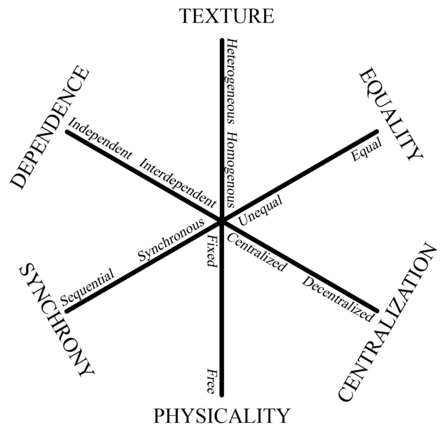


Figure 3.3: The Collaborative Dimension Space diagram for DMIs

Table 3.2 summarizes the reviewed proposals for multi-user interface classification criteria.

Radar charts do provide a useful visualization tool to quickly assess and compare relevant features of multi-user digital instruments. Finally, we must cite two additional multidimensional space representations which prove valuable insight from different stakeholder's perspectives: from the audience and from the conductor's point of view respectively.

Barbosa [Barbosa et al., 2012] employs chart plots to summarize the level of *audience* comprehension concerning a digital musical instrument (see fig 3.4a). The data is obtained through questionnaires and evaluates how successfully the audience understand causes and effects in a performance.

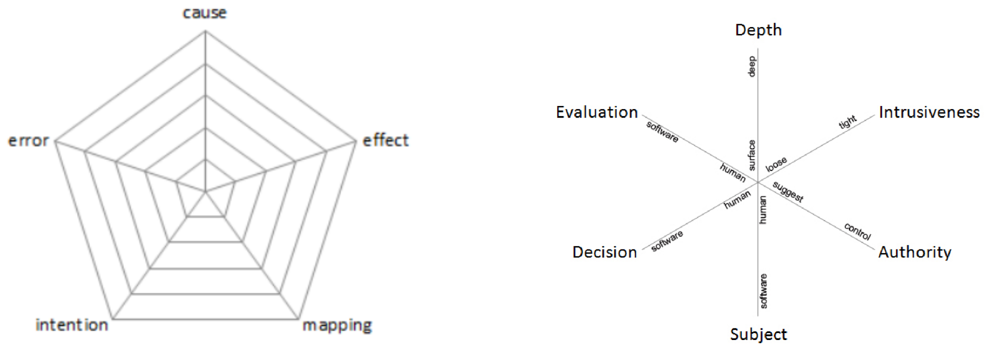
Comajuncosas [Comajuncosas and Guaus, 2014] suggests a *Conductor Dimension Space* to characterize roles and responsibilities in conducted multi-user instruments(see fig 3.4b).

Design criteria for collaborative interfaces

Compared to the plethora of often highly refined electronic instruments designed for single performers, the number of electronic instruments (or instrumental sets) designed for social collaborative sonic creation is significantly more reduced. A similar observation could be drawn on design criteria, though much design issues

Table 3.2: Classification criteria for Multi-user musical interfaces

Criteria	**Weinberg (2003)**	**Blaine and Fels (2003)**	**Barbosa(2004)**	**Jordà (2005)**	**Booth(2010)**	**Hattwick and Wanderley(2012)**
Target	novices, experts	focus (players, audience)				
		learning curve (slow,med,fast)				
		musical genre				
Ensemble	small scale, large scale	scale				
						texture (discernibility, homogenity)
Motivations	process, product					
Creative focus	exploratory, collaborative, competitive					
Spatial topology	centralized, decentralized					centralization
Interdependence topology	democratic, autocratic	directed interaction		hierarchical, non hierarchical		
	equal, unequal	same, different interaction		balanced, unbalanced		equality
	unidirectional, bi-directional			parallel, serial	inclusivity	dependence
				role flexibility	mutability	
Temporal topology	synchronous, sequential		synchronous, asynchronous			synchrony
Location		local, local+net, net	co-located, remote			physicality (space)
Physicality		sensors, level of physicality				pysicality (physical interaction)
Media		image,sound, light…				



(a) DMI characterization from the audience perspective

(b) DMI characterization from the conductor's perspective

Figure 3.4: Two complementary views for DMIs characterization

and recommendations for single-user digital musical instruments are still applicable to collective instruments. Indeed, the ultimate goal for both is to achieve optimum performative engagement [Paine, 2004].

Key considerations in digital instrument design may differ depending on the author's perspective, but they are ultimately closely related. For example, Wanderley emphasizes *learnability*,*exploreability* and *controllability* [Wanderley and Orio, 2002], Settel and Lippe focuses in instrument *resolution* and *expressive depth* [Settel and Lippe, 2003] and Jorda [Jordà, 2004c] prescripts *efficiency*, *learnability* and user-role *flexibility*, while Jorda and Mealla [Jordà, 2014], in the context of a recently proposed methodology for teaching NIME design to students, prescript *musicality*,*expressiveness* and *virtuosity*, among other evaluation criteria.

Blaine [Blaine and Fels, 2003] analyzes multi-player interface design, particularly for collaborative interfaces targeted to novices, from which she provides a sample listing. She emphasizes that low complexity is a prerequisite to increase social interaction and thus opportunities for creativity, even at the expense of expressiveness and virtuosity capabilities. In a word, the author advocates for a design of musical interfaces in which the learning curve facilitates flow and a playfulness to novices. A wide range of hierarchical interaction is shown in the interfaces reviewed. In a related research, Blaine suggests incorporating highly directed rule-based and goal-oriented game-like strategies to achieve engaging collective interfaces while, at the same time, encouraging collaboration in free-form improvisation.

Weinberg [Weinberg, 2005] advocates for an holistic design approach, where musical purpose and social experience are mutually determined when designing a network topology, warning against the incoherent experience that may result when mutual control operates on the more salient control parameters (such as pitch contour) and advocating instead for mutual influence granted over ornamental and expressive parameters. He also states the design must take into account two potential fails regarding mutual control: a too large degree of interdependence may lead to uncertain control of roles, while full autonomy with simple one-to-one mappings may, on the other hand, impair the desired immersive interdependent experience, especially for beginners [Weinberg and Gan, 2000].

As noted by [Morreale et al., 2014], many of the existing research on DMI design ultimately resorts to proposing a set of design metrics and heuristics to be applied in the design process, but don't provide a systematic methodology to guide the whole design process. To overcome such limitations, the authors propose MINUET, a experience-oriented framework for musical interface design. MINUET is structured as a two-stage process comprising two stage, goal and specifications. The conceptual model consist of four entities which determine the objectives and constraints of the interface: People, Activities and Context (for the first stage, which serves to define goals) and Technologies (for the second stage, which defines the specifications). Such entities respectively identify the end-users, expected interaction models, environment specifications and design requirements to satisfy the given design goals. The authors prove the versatility of the MINUET framework by showing how previous frameworks are taxonomies may relate to their generic PACT model and demonstrate their use by systematically addressing the conceptual design of a tangible controller.

More recently, Livingstone [Livingstone, 2003] provides an extensive review of design criteria for collaborative sound environments from different perspectives. The author proposes a taxonomy of interaction models and behaviors to characterize collaborative sound environments. Interaction models are defined by the instrument affordances, which promote specific interaction behaviors: direct parameter manipulation (control model), organization of elements, either in a linear on non-linear fashion (sequential and organization models), linkage of elements to build musical relationships (relational model), sustained musical dialogue with the interface itself (conversational model), even by adapting to user behavior (transformative model). Design decisions can be motivated by the user behavior we want to promote. Livingstone defines four distinct performative behaviors: exploratory (oriented to discovery), interpretative (goal oriented actions), transformative (oriented to abstraction of musical materials) and sociable (oriented to adaptive behavior, including software agents).

Except for Livingstone's work, much of the previously reviewed authors preceded the subsequent proliferation of computer ensembles. Design criteria and recommendations for computer based collective instruments is more oriented towards music HCI research methodologies. A review of relevant human-computer interaction studies for computer-mediated collaborative musical interfaces will provided in chapter 5, in the context of performance evaluation studies.

Evaluation of musical interfaces

One of the main goals of HCI evaluation is to provide generic and practical goals for design and evaluation of design of interfaces.

The evaluation of digital musical instruments, however, is a challenging issue from an HCI perspective. Classical quantitative HCI evaluation techniques to assess the productivity of an interface, oriented towards a task-based paradigm and a stimulus-response interaction model, may not suffice, and may even collide with the artistic goals of a musical interface and the multiplicity of stakeholders (manufacturer, performers, audience) involved [O'Modhrain, 2011]. Some well established methodologies suitable for task-oriented performance measurement like GOMS or *cognitive walkthrough* don't fit well in musical interface evaluation because of its continuous nature and lack of proper models of the cognition involved in the music-making process[Stowell et al., 2009].

There is additionally a noticeable lack of formal evaluations in the literature of digital musical instruments. Many interfaces are often introspectively evaluated by their own creators, thus limiting the generalization of their findings. Stowell's survey[Stowell et al., 2009] on oral papers presented at the NIME conference shows a consistently low proportion of proper formal evaluations, with only 37% of papers presenting new instruments described some sort of formal usability testing. Only in recent years design and evaluation methodologies have received greater attention.

Formal evaluation of musical interfaces may inform designers and provide a more systematic approach in interface design. *Framing a design space for musical interfaces*, as noted by [Morreale et al., 2014] can both guide the process of design and serve as a reference point for further studies.

Confirming the aforementioned observations, on a more recent review on NIME proceedings spanning the 2012-2014 period, Barbosa surveys the different approaches to DMI evaluation, according to targets and skateholders, goals set, evaluation criteria and methodology employed [Barbosa et al., 2015].

The most predominant perspective is that of the performer, followed by designer and audience in the last position. Some notable evaluation criteria from a performer's perspective, such as engagement, effectiveness and expressiveness, still lack consensus on their assessment and measurement. Also from the Performer's perspective, evaluation was mostly qualitative and resorted mainly to interviews and questionnaires.

The goals of evaluation were summarized as five non-exclusive categories, showing a wide variety of purposes behind the term "evaluation" on NIME literature. Among them, assessing the suitability and performance according to predefined criteria was the most common goal, followed by comparisons between interfaces and describing behaviors while testing the interface.

While the authors notice a growing interest in incorporating some sort of evaluation of DMIs, they notice that goals, criteria and evaluation methodologies are either absent on a surprising amount of papers, or lack consensus regarding their precise meaning and are too informally addressed. Consequently, they aim for a more systematic and unified approach to NIME evaluation to guarantee validity, meaningfulness and *replicability* of the results.

Usability oriented evaluation Wanderley and Orio [Wanderley and Orio, 2002] conducted a comprehensive review of HCI usability methodologies for digital musical instrument evaluation. Their main motivation was to overcome the limitations inherent in digital instruments which are used in very few circumstances by very few people, which makes versatility and usability notoriously hard to evaluate from actual musical practice.

The authors comment different musical contexts in which controllers may be utilized, such as performer-instrument, score-control, post-production, multimedia installations and some others not primarily musical such as interactive music scoring or sonification for dance or videogames. They point out that standard HCI metaphors have indeed been incorporated as new contexts of interaction in computer music as well, such as drag and drop, scrubbing and navigation. Focusing on the first, performer-instrument context, the authors propose maximally simple musical tasks to perform quantitative evaluation of learnability, explorability, feature controllability and timing controllability. Those tasks require the user to perform discrete/quantized or continuous movements to reach a given target, from single events, to motifs or whole phrases or gestures, challenging rhythmic performance as well. Interestingly, the authors opt for Likert-scale feedback instead of objective accuracy measurements, relying more on the user to assess for the musical qualities of the interface.

Other authors followed a similar approach for quantitative evaluation of musical interfaces.

In [Kiefer et al., 2008] the authors evaluate the musical usability of the Wimote controller. The musical tasks consisted on performing rhythmic patterns, coordinating pitch changes to a metronome, modifying two continuous parameters simultaneously and finally using predefined shapes to gesturally control . Data collection consisted of performance logs and video recordings. Interestingly, the authors report that the most informative results came from analysis of the interview data, as the conclusions reached from the quantitative analysis seemed limited compared to the subtlety of the participants' observations. Their research provided further suggestions regarding the methodology for musical usability assessment. In this respect, the authors highly encourage to carry out a pilot study to calibrate the interface and to determine a minimum of allotted practice time before carrying out the tasks, and taking into account that musical tasks always entail some degree of creativity which must be properly assessed - and constrained- for a successful experiment.

In [Gelineck and Serafin, 2009] the authors carried out usability tests for evaluating hardware based faders and knobs as music controllers. Those are among the simplest low level interfaces for continuous unidimensional control, and are commonly employed in GUI based control interfaces for music creation as well. For the usability tests, users had to perform musical tasks which consisted in imitating reference sounds with the provided interfaces. The similarity was assessed by experts. A questionnaire to measure the perceived difficulty of each task was filled afterwards, rating how successfully each interface provided accurate and intuitive control and whether it was inspiring, frustrating, or predictable among other criteria. Although the authors found no significant differences between the two, their methodology provides a valuable approach to investigate the creative and exploratory affordances of musical interfaces.

The aforementioned approaches are based on quantitative analysis, usually measuring an operational gain of the interface in terms of utility or usability. However, it is often difficult to demonstrate such gain in the context of creative applications, and its relevance is often questionable [Marquez-Borbon et al., 2011].

Some common criticisms to Wanderley's and other authors' methodology are related to its reductionist approach, which neglects the affective and creative aspects of music making and equates controllability with expressiveness [Stowell et al., 2009].

Jorda [Jordà, 2004a] elaborates on the relationship between performer and instrument from the perspectives of instrumental efficiency, apprenticeship and learning curve, by establishing parallelisms with some paradigmatic acoustic instruments. The learning curve is a graphical representation of the relationship between learning time and acquired skills, being better a curve which depicts a higher level of achieved skills in a shorter time period, and the possibility to achieve higher skills by devoting more training time. Therefore a desirable instrument design should balance a provision for a gentle entry point, incremental rewards and space for virtuosity, both avoiding boredom and frustration. The concept of instrumental efficiency from an HCI perspective is an indicator of the relationship between the amount of performance features offered by the instrument and the actual range of musical and expressive possibilities offered. Such efficiency dynamically varies according to a performer's skills and expectations. Of course, a very efficient instrument might entail poor user engagement in case of minimal interaction scenarios, Jordà accordingly introduces the term Performer Freedom to account for the degrees and range of freedom available to the performer in terms of user intervention and musical outcome, and finally summarizes the whole set of relationships with the illustrative formula

$$MusicInstrumentEfficiency = \frac{MusicOutputComplexity \cdot PerformerFreedom}{ControlInputComplexity} \tag{3.2}$$

The same author [Jordà, 2004b] suggests that DMIs, indeed any musical instrument, can be described in terms of their ability to support a performer in realizing diverse musical goals. It therefore advocates for an evaluation centered on actual musical contexts rather than on restricted musical tasks. The author further categorizes such desirable musical diversity in three not mutually exclusive levels:

- **Macro diversity** the ability to adapt to varied musical contexts, styles and genres
- **Mid diversity** the ability to provide resources for distinct performances, making possible very contrasting realizations from the same set of performative actions available, for example in the form of varied repertoire
- **Micro diversity** the ability to support varied performative nuances, being sensitive to subtle forms of playing, and thus encouraging sophisticated levels of virtuosity

Though there have not been, to this author's knowledge, any quantitative evaluations of DMIs using Jordà's approach, it nonetheless offers a different, complementary perspective to the aforementioned methodologies, which might provide feedback on other aspects of DMI designs, as noted by O'Modhrain [O'Modhrain, 2011].

Experience oriented evaluation J. Barbosa's review [Barbosa et al., 2011] on evaluation methodologies for digital musical instruments show a progressive tendency towards ethnographic, qualitative approaches to interface evaluation, with most researchers advocating for performer centered evaluations.

While quantitative usability tests may provide some insight into the design of musical interfaces, assessing user experience usually must resort to qualitative approaches. A formal qualitative evaluation technique which has been occasionally employed in the evaluation of digital musical instruments is Discourse Analysis. Discourse Analysis proposes a systematic methodology to analyze transcribed discourse data looking for patterns and consistencies among the participants, in order to reconstruct and compare their experiences in approaching an interface in a situated context.

Such methodology, drawn from linguistics and social sciences, may be incorporated to the study of the complex musical interactions such as those expected in collaborative musical interfaces. For example Stowel [Stowell et al., 2008] suggests a procedure consisting of video recording solo sessions exploring an interface and subsequently complementing them with group sessions to encourage discussion and further experimentation. The analysis stage consisted of a contextualized reconstruction of each participant "world" arising from the objects (objects, actors, relationships and attitudes) identified in the transcriptions of the recordings. Discourse Analysis analysis can thus yield valuable information on users' preferences, approaches and attitudes, by means of capturing and analyzing subtle nuances in their responses through unstructured dialogue.

Alternatively, content analysis techniques by Grounded Theory [Corbin and Strauss, 2014] my prove useful in a more exploratory analysis, as they allow to identify and categorize emergent concepts arising from participant's verbal descriptions -this methodology has been proved useful in assessing evaluation criteria for music performance as well, as we will discuss in Chapter 5. For example [Ghamsari et al., 2013] carried out an evaluation of a novel musical interface by asking performers to improvise with them and then using content analysis to identify the main categories constructed by participants in the subsequent verbal discussion.

Some researchers combine quantitative measurements in task-oriented experiments with users' feedback. Poepel [Poepel, 2005] presented a methodology for evaluating the potential for musical expressivity. Participants were asked to play and compare distinct string instruments coupled to different FM-based synthesis algorithms, by performing simple musical tasks which served to evaluate accuracy on a number of musical dimensions as relevant factors to assess the potential for musical expressivity of the interfaces. The experiment showed good agreement between the operationalized metric for expressivity potential and performers' feedback in the questionnaires and final interview.

One should not attempt to relate quantitative methods to simple interfaces and qualitative methods to complex ones. Indeed, the diversity and complexity of musical behaviors, even on the simplest interface, may benefit from ethnographic analysis to evaluate a digital musical instrument. For example, Gurevich [Gurevich et al., 2010] designed a one-button single-pitched instrument and asked performers to work on solo performances with it. Through video recordings and questionnaires the authors examined the relationship between such severe constraints and the musical style developed by performers. Even if the limitations eventually discouraged some performers, many others resorted to a diversity of strategies to achieve a significant degree of stylistic diversity. The authors observed that this process took place both in spite of and because of the constrained design. They conclude that there is not a simple causal relationship between the properties of a musical device and its potential for virtuosity.

In a second study, the authors [Marquez-Borbon et al., 2011] evaluated the perception of an spectator regarding the complexity of a digital musical instrument. By incorporating nonlinearities and mode switches in an unfamiliar instrument, participants were easily confused about the non univocal relationship between gesture and sound, either attributing it to performer's skill or to lack of performative control.

A corollary of the study of both interfaces is that, in a real scenario, instruments can be appropriated and perceived in all sort of ways, and such unintended anomalies are valuable indicators of affective and creative aspects of music-making. Therefore they should not be avoided by placing too reductive quantitative studies [Stowell et al., 2009].

O'Modhrain [O'Modhrain, 2011] proposes a framework for the evaluation of digital musical instruments based on the roles and perspectives of the various stakeholders involved in the design process. The development of a new musical interface is an iterative process which relies on feedback from several actors who actively participate in their incremental refinement process. Each actor provides a different perspective and challenges in the design process:

- **Audience** The main issue in the evaluation of DMIs from the audience perspective is the perception of gesture to sound causality. Expressive and emotional intent is often extracted from gestural cues, even more than from aural cues. Additionally, the perception of effort and success assumes an acquired knowledge of the instrument through previous repertoire and well established performative practices. It is therefore relevant to know how, in a context challenging both gestural causality and instrumental identity, an spectator elicits what are the most relevant cues to understand a performer's intentionality.

- **Performer** The main concern here is the generic concept of instrumental efficiency, what factors may compromise it (such as instrumental unreliability, a poorly balanced learning curve, lack of repertoire, or poor design coherence among others). As performers may have a background in acoustic instrument performance, it must be evaluated to what extent building upon an existing instrument metaphors is an efficient approach. Interestingly, performers may adapt, even creatively, to non ideal interfaces, so independent assessment is essential.

- **Designer** From a designer's viewpoint, traditional HCI evaluation methodologies are unsuitable to evaluate interactive systems, because of the inherent realtime coupling between the performer and the system. The main concern here is how to evaluate user experience unobtrusively. An instrument-centered evaluation focuses on playability, while a user-centered evaluation is more concerned about playing experience.

- **Player** The most challenging factor in the evaluation of the player experience is how to assess it without disrupting his engagement in a fully immersive and time-critical context. Evaluation is usually both quantitative and qualitative. Quantitative measurements usually are provided by capturing performance data (accuracy) while qualitative measurements in form of questionnaires (retrospective reflections on his cognitive experience). The authors highlight the fact that both evaluations are not concurrent. Non-intrusive systems to monitor physiological data are thus an open, and yet hardly explored field of research to directly to evaluate the experience of playing a DMI.

3.3 Laptop Orchestras

3.3.1 Historical overview

The native processing power and networking abilities of portable computers in the new millennium made possible to extend the HUB performative paradigms to incorporate live audio (and visual) generative processes which didn't have to resort to custom and often cumbersome setups. This is the case of the first Laptop Orchestra founded in 2002 by Philippe Chatelain in Tokyo, a laptop ensemble closely adhering to the principles of Network Music.

However, the proliferation of Laptop Orchestras really started with the adoption of a larger, more participatory format within the computer music departments and research centers of several Universities from the United States. We must mention here the pioneering work of The Princeton Laptop Orchestra, founded in 2005 by Dan Trueman and Perry Cook and the Stanford Laptop Orchestra (SLOrk) founded in 2008 by Ge Wang. Those bigger ensembles brought a new model of computer ensemble which is now prevalent in the United States, the so-called *LOrk model*.

A more recent development was the creation of computer ensembles employing mobile devices such as mobile phones as their primary instrument. The Stanford Mobile Phone Orchestra (MoPhO), created in 2007, pioneered this novel approach to computer ensembles. Though sharing similitudes with Laptop Orchestras, Mobile Orchestras do offer some new interaction possibilities [Oh et al., 2010]. Some novel features provided by such devices are the possibility to move freely, walking around in the performance space or among the audience, design expressive, gesture-based interfaces which both encourage face-to-face interaction between performers and allow the audience to visually map sounds to movements. Since their creation, they put specific emphasis in building a persistent repertoire, mostly consisting of free-form, structured and conducted improvisations

[Wang et al., 2008].

Figure 3.5 summarizes a restricted survey[1] of computer ensembles ordered by founding data, showing a peak in 2008-2011 period. This survey focuses on well established, mostly institutional Laptop Orchestras, but the actual number of active computer ensembles is much larger.

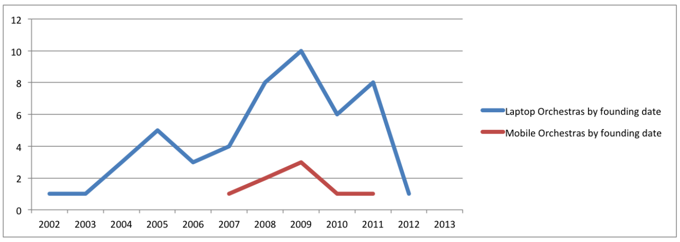


Figure 3.5: Number of laptop and mobile ensembles by year of creation

A recent survey carried out by Shelly Knotts is, to our knowledge, the most comprehensive overview of the developing laptop ensemble scene[Knotts and Collins, 2014]. The survey was undertaken of 160 laptop ensembles and was focused on their organizational and social structure.

3.3.2 Research topics on Laptop Orchestras

Research on issues related to Laptop Orchestras is mainly addressed in the New Interfaces for Music Expression (NIME), International Computer Music Conference (ICMC) and Sound & Music Computing (SMC) conferences. On a survey of the NIME proceedings for the whole 2001-2015 period, research papers citing or dealing with Computer Ensembles, Laptop Orchestras or Mobile Orchestras show a consistent growing trend, peaking in the 2012-2014 period both in absolute and relative terms (see Figure 3.6). Indeed the year 2012 the First (and only, up to date) Symposium of Laptop Orchestras took place in Louisiana, where over 30 research papers on Laptop Orchestra related topics were presented.

Summarizing the main Laptop Orchestra related research topics outlined by the NIME Conference and SLEO Symposium will give us an idea of active areas of research in this field. By keyword and content analysis we conclude that the most relevant research topics in Laptop Orchestras relate to *compositional strategies* (usually illustrated by repertoire reviews), *setup, infrastructure and instrument design methodologies* and *analysis and frameworks for collaborative interfaces* (see Figure 3.7). Other relevant topics are the discussion of performance strategies and challenges, the educational perspective -indeed, most laptop

[1] `https://silpayamanant.wordpress.com/new-music-ensembles-in-the-us/laptop-orchestras-and-ensembles/`

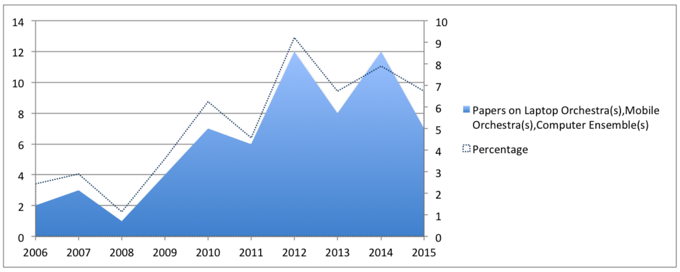


Figure 3.6: Number papers on computer ensembles in the NIME Proceedings

orchestras belong to academic institutions- and the development of frameworks for network infrastructure. Finally, there are some studies on aesthetics and sociology of Laptop Orchestras as well.

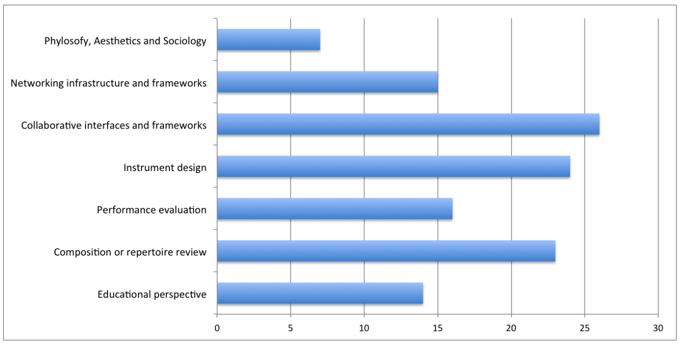


Figure 3.7: Relevant research topics on Laptop Orchestras in the NIME Proceedings

Closely related to computer ensembles is the research on design and evaluation of collaborative musical interfaces, such as collaborative tangible interfaces (such as music tabletops, or collective musical interfaces for novices),remote performances and interactive installations

Ensemble models

Winkler [Winkler, 1998] defines four performance models for traditional instrumental ensembles in the context of interactive music systems: the conductor model, the chamber music model, the improvisation model and the free improvisation model. The social roles participating and defining a computer ensemble extend beyond Winkler instrumental models because both the compositional stage and the design and implementation of the software and hardware infrastructure are usually taken care by the ensemble as an essential part of their creative work [Wang et al., 2009].

Moreover, the instrumental model is rather dynamic in a Laptop Orchestra, and widely varies according to the chosen repertoire. We might, however, tentatively assign the former performance models to some well established ensembles according to their regular practice. Under this perspective, the institutional ensembles adhering to the LOrk model tend to follow a conductor model , while smaller ensembles better adhere to the chamber music or free improvisational models. We will call this approach the HUB model, as such ensembles tend to follow many of the practices of The League and the HUB.

In a way, both models reflect in the music collective context what Jordà identifies [Jordà, 2005a] as the tension between the *luthier-as-composer role* -for the LOrK ensemble model- and the *improviser-as-composer* (an luthier!) role -for the HUB ensemble model-.

Mobile music ensembles, by the other hand, belong to the theoretical framework of Mobile Music as well, as defined by Gaye et al. [Gaye et al., 2006], at the same time deeply influenced by the broader concept of Locative Arts, as introduced by Drew Hemment in 2004[Hemment, 2006]. The authors identify novel possibilities for music collaboration which take advantage of context awareness by location and proximity sensing.

We will now discuss the two most relevant ensemble models for Laptop Orchestras, the LOrk model and the HUB model.

LOrK model As previously stated, and according to Knotts[Knotts and Collins, 2014], the LOrk style ensemble is now the dominant model in the USA, with a higher level of university affiliation, whereas European groups are more often unaffiliated and of smaller size.

The so-called LOrk model advocates for medium to big sized ensembles in which performers use laptops as their sound generating devices, extending them with occasional gestural devices and usually incorporating hemispheric multichannel speaker arrays for to provide an individually localized sound radiation pattern.

Those ensembles may consist of more than a dozen of players, and use to adopt some practices commonly seen in large orchestral ensembles: explicit use of the space, by distributing performers on stage and resorting to instrumental sections, and a performative ritual built around a programme of composed works, preventing scenarios of indeterminacy and at most resorting to highly structured improvisations.

Additionally, large sized Laptop Orchestras usually have a conductor, who may employ different conduction strategies, from expressive gestural shaping to signs to control a structured improvisation or instrumental gestures to be mirrored by the performers. The primary use of the network is to assist the conductor and players to enforce synchronization and automating structural events in the performance, though more interactive schemes are non uncommon. Even if participants are actively encouraged to take part in the creative process, Laptop Orchestras often conceive new repertoire more in a top-down approach, providing ready-to-be-played works in which roles and behaviors are much less idiosyncratic and autonomous than in the League or the HUB.

In conclusion, the LOrk model might tend to rather autocratic. The orchestra defines a homogeneous hardware setup, provides well documented repertoire with individual authorship and is formed by undergraduate students who take part in the ensemble as traditional performers. Some relevant examples are the Princeton Laptop Orchestra, the Linux Laptop Orchestra, the Stanford Laptop Orchestra or the Mobile Phone Orchestra to mention just a few.

HUB model Smaller ensembles may tend to adopt a more democratic model, closely adhering to the HUB model of performers-as-composers-as-designers and focused in the exploration of the practices and paradigms of Network Music. Rapidly shifting roles and improvisational practices are much more common, therefore we might observe behaviors more characteristic of chamber or free jazz ensembles.

Those ensembles tend to be smaller, though there are exceptions, as for example with the Cybernetic Orchestra (10 members) [Ogborn, 2014] or the Republic 111 ensemble (cited by [Knotts and Collins, 2014]). The approach is more collaborative and may resort in heterogeneous setups. A strong Do-It-Yourself approach is prevalent and performers are usually given the responsibility to prepare their own setups [Hewitt et al., 2010].

Knotts observes that such ensembles may tend to be more democratic [Knotts and Collins, 2014], despite the eventual presence of a conductor, whose role is to coach, mediate and otherwise indirectly guide the performance rather than explicitly commanding performers.

HUB-inspired ensembles resort to multi-user networking paradigms and/or live-coding practices, and play more improvised or collaboratively composed pieces. The collaborative process in the infrastructure building may extend to the performance itself, thanks to the facilities provided by the network regarding the sharing of information, resources and code [Wilson et al., 2010].

As relevant examples of computer ensembles adhering to the HUB model we could mention the BiLE (Birmingham Laptop Ensemble), PowerBooks Unplugged or Benoit and the Mandelbrots for example.

The Barcelona Laptop Orchestra closely adheres to the HUB ensemble model, despite being a conducted ensemble, and its performative paradigms will be further discussed in Chapter 4. Finally, we should emphasize here that such categories are in no way mutually exclusive. Indeed, a relevant amount of the repertoire developed by the former Lork-style orchestras explore network music scenarios more akin to the HUB-style ensembles.

Beyond the acousmatic experience

> *For an unhappy audience who is inclined to believe that a performer with the movement patterns of a clerk is probably playing back a cd while checking his mail, laptop performances are lacking appeal.*
>
> Julian Rohrhuber [Rohrhuber et al., 2007]

Even if computer ensembles are usually known by the kind of computer device that performers use to play - and so we may have laptop orchestras or mobile phone orchestras, for example- the very nature of computer devices makes it hard to associate any kind of instrumental identity to computer ensembles. Computer devices, and sets of co-located computer devices as well, are culturally associated to their primary design purpose -office use- or to some other tangentially well established contexts, such as videogame LAN parties. The audience faces the challenge of building expectations and associations for devices which are not readily linked to a sound identity nor a style of playing nor a known repertoire and performative tradition [Cascone, 2002]

Live computer music performance may be regarded as hardly believable, both by its lack of expressive gesturality, by our acquired cultural knowledge which relates computers to office work and by the lack of physical causality between actions and sounds. Expectations regarding embodied expressivity might not be so critical in larger computer ensembles, but are nonetheless present. Not to say that some ensembles advocate for this acousmatic approach as an aesthetic statement, this is the case of the Florida League for Intedeterminate Performance [Peuquet et al., 2012].

On the other side, to provide a non-acousmatic listening situation Ruviaro states that Laptop Orchestras must account for the need to provide presence, movement and cultural references to their performance[Ruviaro, 2011]. This challenge of embodiment is, paradoxically, increased by the incorporation of complex networking schemes which may enhance social interaction among performers but which, by the other hand, tampers with the connection between ensemble participants and audience members[Vallis et al., 2012]

Presence and movement can be greatly enhanced through the user of gestural controllers. Laptop Orchestras do make regular use of such controllers, in line with the current NIME design practices and approaching the laptop as a freely expandable and configurable meta-instrument [Trueman et al., 2006][Trueman, 2010].

Choreographic and idiomatic gestures are a useful resource to provide cultural anchor points for the audience to better understand the (see, for example, Forkish et al.' Cop de Cap as described in [Ruviaro, 2011]. For some ensembles physical presence, gesturality and choreography are their *raison d'etre*, this is the case of the Linux Laptop Orchestra [Tech et al., 2010].

Mobile orchestras, on the other hand, provide more opportunities for embodied interaction. Mobile devices do offer more gesture-based possibilities of interaction thanks to their embedded range of gesture sensors[Oh et al., 2010].

The incorporation of visual information is expected to reinforce the listener experience as well, by providing additional expressive cues and making the whole performance more understandable and believable. Visual cues may come from graphical elements in users' interfaces or projected to the audience, and from performer's gestures, conveyed by the use of gestural interfaces.

Wang [Wang, 2014] elaborates a set of principles of expressive visual design for computer music, both for musical visualization and for user interfaces. Regarding HCI considerations, Wang advocates for real-time visuals, concurrently designed with audio processes, offering intelligible affordances, being content-oriented, constrained, and guiding and encouraging physical interaction. Visuals must be informative and suggestive at the same time, therefore the designer should find the optimal equilibrium between functional and aesthetic considerations, which ultimately reinforce each other. To Wang visual interfaces for computer music should be essentialistic but aesthetically appealing as well. Musical processes are best displayed with dynamic, organic visuals which provide defined behaviors and identities to elements.

Infrastructure

In this section we will review the approaches and proposals related to the development of technological infrastructure for Laptop Orchestras, a key research topic in the literature on computer music ensembles as we have outlined.

While the trend towards computer based meta- an extended instruments is a rather distinctive feature in LOrk ensembles, some other technological resources have been consistently adopted by a majority of ensembles, notably the networking infrastructure for ensemble interaction.

Sound reinforcement infrastructure Since its inception, the Princeton Laptop Orchestra emerged from a body of research encompassing the design of sound radiating devices to achieve a more instrument-like presence and the embodied interaction through specifically tailored computer interfacing devices. It should not surprise us, therefore, that a strong emphasis was given to the specification of a laptop-based meta-instrument incorporating such features. A typical PLOrk setup consist of an ensemble of 15 or more stations consisting of a laptop, sound interfacing hardware and a six-channel speaker with individually addressable drivers to emulate different sound radiation patterns. Performers play onstage, sitting on floor and communicating with each other and with the conductor through a local area network [Trueman et al., 2006].

Compared to a shared PA setup, this approach offers some advantages. At an individual level, Cook and Trueman noticed that BoSSA [Trueman and Cook, 2000a], the violin-inspired device which first incorporated the speaker arrays the PLOrk would be based on, provided a presence similar to an acoustic instrument, a sense of space which allows for a sort of intimacy and physical relationship with sound, which emanates in front of the performer, surrounding him. In turn, that intimate relationship with the instrument leads to increased opportunities for expressive performance and ensemble awareness. At an ensemble level, individual sound sources contribute to a clarification of textures by means of their spatial and personal roles.

Individual sound sources sitting next to performers evoque a traditional acoustic ensemble in terms of sound localization, and also suggest performative constraints related to the speed of sound and visual connection between players. While such constraints ultimately pose a limit for coordination between players, such limits can nonetheless be violated or bypassed by means of network coordination, which may be both thrilling and disturbing, as Cook and Trueman noticed [Trueman, 2007].

Another approach in terms of sound setup comes from Alberto di Campo's ensemble Powerbooks Unplugged [Rohrhuber et al., 2007] resorts to the built-in speakers of laptops, with their restricted high frequency range. Performers sit among the audience thus blurring the audience/performer/stage conventions and at the same time maximizing the spatial qualities of such quasi-acoustic experience [2]. Ensembles like the Huddersfield Experimental Laptop Orchestra [Hewitt et al., 2010] adhere to this approach for its convenience and portability.

Finally, the very assignation of sound sources to instrumentalists may be challenged by the nature of the instrument (for example, in the case of multi-user instruments or shared environments). The IEM Computermusic Ensemble (ICE) goes far beyond the traditional *sound-source-per-player* paradigm by decoupling the physical locations of performers from the actual placement of sound sources. Their spatialization framework [Ritsch, 2014] consists of a virtual concert hall (VCH) which simulates a 3D acoustic space into which performers may freely incorporate their instrumental voices. A specific rendering engines instantiate this virtual space as required, for example as an Ambisonics decoding into a multichannel audio setup or as a decoding for binaural listening.

[2] http://pbup.net/s/

Network synchronization for computer ensembles Most Laptop Orchestra performances highly rely on, and build upon network coordination, synchronization and timing techniques. Issues like connectivity protocols, network topology, bandwidth, latency and jitter control techniques have been individually addressed by most Laptop Orchestras. Nowadays, most if not all Laptop Orchestras extensively employ Open Sound Control over UDP to communicate because of its flexibility and extensive implementations available. UDP is usually chosen for the minimum latency requirements, despite its unreliable delivery.

A centralized clock is the most employed synchronization mechanism since the first laptop orchestras: Perry Cook, Gee Wang and David Trueman addressed some strategies to enforce tight synchrony in laptop ensembles, using a conductor laptop which sends pulses the stations synchronize with. [Smallwood et al., 2008] The most common issues detected in networked music ensembles are dropped packets, non-uniform latency between stations, jitter, and the mixed effect of audio and network latency. Some toolkits have been developed to analyze those issues, such as LOrkNeT) [Cerqueira, 2010]. Provided a shared absolute time reference is available, time-stamping events allows stations to trade lower jitter for a longer latency by deferring, processing or discarding packets according to their time of arrival, that is, by *forward synchronization scheduling*[Cerqueira, 2010] [Schmeder et al., 2010]

Platforms and frameworks for collaborative instrument design and performance Most laptop orchestras resort to generic programming environments for developing real-time musical software applications, which adhere to a modular dataflow paradigm indebted to Max Mathew's Music N family. Graphical environments such as Puredata [Puckette and Others, 1996] or MaxMSP [Puckette et al., 1990] allow for visually programming sound synthesis algorithms and interactive music processes by wiring boxes which implement lower level functionality, following Puckette's Patcher programming model [Puckette, 1988], while text-based environments such as SuperCollider [McCartney, 2002] or ChucK [Wang and Cook, 2004] resort to custom programming languages aimed at facilitating the expression of compositional and signal processing ideas.

Such environments excel at prototyping single-user DMI but, as all of them provide native networking capabilities, may seamlessly extend to multi-user and collaborative performance environments. Many Laptop Orchestras develop their own OSC-based frameworks to handle intercommunication and to facilitate the implementation of network based repertoire in which there is need for resource sharing and collective user awareness. Such frameworks and toolkits are often tailored to the specific environments and requirements of that particular Laptop Orchestra, and therefore lack enough generalizability as to be quickly adopted by other ensembles.

Rebeca Fiebrink contributed with the development of two distinct frameworks for Laptop Orchestra instrument design. SMELT [Fiebrink et al., 2007] was based on the requirements of the PLOrk repertoire and is an open-source toolkit aimed at facilitating rapid prototyping of interfaces build on the native input capabilities of laptops, such as keyboard and mouse.

The Wekinator [Fiebrink et al., 2009b] is an OSC-enabled Java environment for supervised, interactive machine learning for music based on Weka [3], a popular software for data mining. It allows the performer to build controller mappings in real-time. It includes a built-in number of low level audio and video feature extractors and sensor data capture for training, as well as the wide repertoire of learning algorithms provided by the Weka machine learning library. Fiebrink[Fiebrink et al., 2009b] also proposes a play-along strategy for training, in which the user freely performs gestures in accordance with a provided "score". They show examples for realtime mapping of gesture to musical generation and physical model control. Many more similar mapping frameworks and toolkits do exist, most of them resorting to a variety of supervised and unsupervised machine learning techniques (see for example [Bevilacqua et al., 2005], [Van Nort and Wanderley, 2006], [Gillian et al., 2011], [Smith and Garnett, 2012], [Gillian and Paradiso, 2014], [Malloch et al., 2014]) though Wekinator's approach, allowing the incorporate the very process of instrument building into the ensemble performance itself, is rather unique.

While the two former environments are specifically designed for mapping input controllers, some other frameworks or toolkits are more focused towards simplifying the setup, network infrastructure and user management of a computer ensemble. Regarding the setup, GRENDL[Beck et al., 2011] is designed to handle the deployment and management of hardware and software setup, and distributes scores to performers .

As for network infrastructure, NRCI, developed by Christopher Burns and Greg Surges[Burns and Surges, 2008], is a set of abstractions for user interfaces, GUI-based controls, audio generation and processing routines and their own networking protocols to handle offline delivery of performance data and flexible negotiation of control commands, all designed to facilitate the construction of software instruments. It was developed an tailored to suit the MiLO (Milwaukee Laptop Orchestra) requirements.

Another package aimed at facilitating ensemble networking is Republic. Republic [4] is an extension to SuperCollider which allows for organization and cooperation in changing groups. It transparently takes care of synchronization and administers users' addresses for easy session management. It was developed to

[3] `http://www.cs.waikato.ac.nz/ml/weka/`

[4] `https://github.com/supercollider-quarks/Republic`

Table 3.3: A sample of frameworks used for instrument design, networking infrastructure and performance with computer ensembles

Framework	Year	Platform	Main purpose
CoAudicle	2005	ChucK	networking
SMELT	2007	C++	mapping
NRCI	2008	PureData	design and networking
Wekinator	2009	Java	mapping
Republic	2009	SuperCollider	networking
GRENDL	2011	C++	deployment
LOLC	2011	Java	performance

suit the live-coding performances of PowerBooks_UnPlugged but it is currently being used for other SuperCollider live-coding ensembles[de Campo, 2013].

LOLC [Freeman et al., 2003] is a Java environment for collaborative text-based laptop performance. It supports chat, exchange of musical materials throughout the performance and real-time music notation generation.

CoAudicle[5], is a graphical environment which extends ChucK live-coding paradigm to interactive network-enhanced performance between multiple performers [Wang et al., 2005], allowing for a client-server model, where performers may collaboratively edit others' code run on a centralized server, or a peer-to-peer model where each performer runs his own ChucK virtual machine but may exchange code and data with other performers. It therefore allows for highly interdependent processes to take place concurrently.

Finally, table 3.3 summarizes the surveyed frameworks, their year of creation, platform and main purpose.

Live-coding environments for ensemble performance Some computer ensembles strongly adhere to live coding practices. For example the network band Powerbooks_Unplugged, one of the largest unconducted laptop ensembles, utilizes their aforementioned SuperCollider-based Republic[de Campo, 2013] system for collaborative live coding, while Ge Wang developed the ChucK[Wang and Cook, 2004] language to introduce live coding in the Princeton Laptop Orchestra. Many other generic programming or computer music oriented environments have been used for live-coding, the only requirement being that they are realtime oriented [Zmölnig and Eckel, 2007]. We could cite here Impromptu [Brown and Sorensen, 2007], Fluxus [6], ChuK, LuaAV [Wakefield et al., 2010] or ixi lang [Magnusson, 2011b], to mention just a

[5]http://audicle.cs.princeton.edu/doc/faces/co.html

[6]http://www.pawfal.org/fluxus/

few.

The graphical counterpart to collective text-based live-coding practices is dynamic patching, that is, the ability to concurrently edit a patch while performing, therefore breaking the traditional separation between editing and performance views and working modes. We should mention here the dynamic patching features incorporated into PureData by Kaltenbrunner et al. [Kaltenbrunner et al., 2004] which were motivated by the ReacTable design requirements [Kaltenbranner et al., 2006] [Jordà et al., 2007]. Some other approaches for collective dynamic patching are Hans-Christoph Steiner's Xtreme Programming for PureData [7], Sarlo's GrIPD remote interface for concurrent patching [Sarlo, 2003] and the Barcelona Laptop Orchestra's HUB, a concurrent live-patching environment built on PureData's dynamic patching features (see 4.2.5).

Frameworks for mobile musical interface design We have come a long way since Essl, Wang and Rohs[Essl et al., 2008] stressed the need for generic instrument building tools for mobile devices, in pair with the rich set of features provided by their hardware.

Nowadays several frameworks for DMI development for mobile devices are available. Many aim at simplifying the mapping from device sensors to synthesis parameters, for rapid interface construction, usually taking into account the different I/O capabilities of mobile devices, with smaller displays than their desktop counterparts, but a rich set of sensor devices. They usually focus on touch-based interaction on multi-touch screens, instead of the alphanumeric keyboards and single-point mouse interaction which characterize the prevailing editing paradigms for laptop computers.

Among those, a number of frameworks focus on providing users a flexible platform to replace the traditional fader control boxes with equivalent touch based GUIs, and may be deployed either as native applications (such as Control[8] or TouchOSC[9]) or as browser-based interfaces (as is the case with NexusUI[10]). Some of those frameworks are intended to communicate with a server (usually a desktop device onstage) which is in charge of more CPU intensive processes, such as audio synthesis, and may facilitate the integration by automating the procedure, though nowadays embedded audio synthesis is entirely feasible.

In general, latency may be an issue when using such devices for musical use. Certain OS are not tailored for real-time performance, but a client-server architecture must account for the latency and jitter caused by WiFi LAN connection, much

[7] `https://gem.iem.at/pd/pd/Members/hans/xp4pd`
[8] `http://charlie-roberts.com/Control/`
[9] `http://hexler.net/software/touchosc`
[10] `http://nexusosc.com/`

Table 3.4: Frameworks for mobile musical interface design

framework	**environment**	**platforms**	**comments**
mMTCF	PureData	Android	OSC control, synthesis with libpd [11]
MobMuPlat	PureData	Android, iOS	OSC control, synthesis with libpd[12]
Csound mobile	Csound	Android, iOS	uses the CsoundAPI [13]
massMobile	Max MSP	Android, iOS	uses the MaxMSP Api [14]
Mira	Max MSP	iOS	GUI to Max bridge [15]

higher if resorting to broadband connectivity as in environments where the audience participation is expected. In the latter case, the large and essentially unpredictable latency of cellular networks might advocate for performative paradigms in which precise real-time is not an issue, such as in collective voting systems or collaborative sequencing, to mention two audience oriented projects using the massMobile framework[Weitzner et al., 2012]

The proliferation of different architectures and OS associated to mobile devices pose a problem for the developer of musical applications. A possible workaround are browser-based application development, which would theoretically maximize multi-platform interoperability. Developing for the browser allows the designer of musical interfaces to take advantage of already existing solutions. If earlier java or flash applets, the current trend aims to take advantage of native browser capabilities, employing Javascript and the Web audio API for audio synthesis and processing and CSS and HTML for simple visual layout. This is the case of the NexusUI platform, which integrates with WebAudio.

Additionally, many generic sound synthesis environments and interactive music programming environments have ports for mobile platforms at different stages of development. Some of them assist in building GUI bridges to desktop computer patches carrying the DSP, though OSC or specific APIs provided by the platform, while some others are direct ports of the original environments to the new platforms, either as standalone packages or as libraries to be incorporated in third-party applications. Table 3.4 lists some mobile frameworks suitable for musical ensemble interface design.

3.3.3 Compositional and notational approaches

This section surveys relevant repertoire for computer ensembles adopting the paradigm of interdependent multi-user instruments. Of special interest are the social hierarchies which the topology of interdependence promotes, and the role and performative options given to players. The role of the conductor, if present, is analyzed as well.

The purpose of this section is not to carry out an exhaustive review of the existing repertoire for computer ensembles, but to provide illustrative examples of compositions for computer ensembles in view of the aforementioned taxonomies and performative paradigms.

For a related survey on interfaces, ensembles and contexts for collaborative musical experiences covering the period from 1993 to 2002 see [Blaine and Fels, 2003] . More recently, Xambó carried out a similar survey on systems, musical pieces and computer ensembles as a whole, which illustrates specific cases of interdependent networks, computer and mobile music ensembles, collective tangible user interfaces as well as historical examples of network music systems [Xambó, 2015].

As this review is based on actual compositions and systems-as-compositions, it largely omits the contribution of improvisational live-coding practices, which are hardly documented. Additionally, the most documented repertoire for Laptop Orchestras is to be found at PLOrk[16] and SLOrk[17] websites, therefore this survey will exhibit an unavoidable bias towards their repertoire.

A final section will review the main compositional and performative notation techniques to document the repertoire for Laptop Orchestras.

A review of the repertoire for Laptop Orchestras

Networking and governance

As previously described, a prevalent use of the network in computer ensembles following the LOrK model, in accordance with its more centralized and hierarchical approach, is to enforce synchronization and provide tools for proper ensemble governance.

In some cases, networking synchronization techniques are incorporated whenever the need for them arises. In *Supreme Balloon* (PLOrk), for example, players manually trigger arpeggiato sequences. A server shows a tempo which must be manually followed be performers by ear. Lack of proper instrumental training, or maybe latency issues as well, made the human coordination unfeasible up to the point that the server ultimately enforced the tempo to guarantee proper synchronization [Reinecke, 2010]. A similar approach may be seen in N*on-specific Gamelan Taiko-Fusion Band* [Trueman et al., 2006] [Wang and Cook, 2006], where multiple sequencers are kept in tight synchrony by a central clock.

CliX(Ge Wang, 2006) epitomizes a piece composed for a LOrK ensemble: all performers play the same sound generator (it is a highly homogeneous setup) and the networking implementation is just designed to enforce synchronization, acting as a conductor assistant, but not to allow for a more democratic interaction

[16]http://plork.princeton.edu/index.php
[17]http://slork.stanford.edu/

between performers. Additionally, the use of the space is relevant as the distribution of performers onstage is a relevant factor in the performance. In CliX, which consists of a conducted improvisation, two conduction levels take place simultaneously. A conductor gesturally instruct what players must play, and with what density and frequency by typing keys on their laptops. At the same time such actions are quantized and synchronized to a shared network clock sent by the conductor's laptop over WiFi, which surprisingly wasn't reliable enough[Cerqueira, 2010]. Finally wired ethernet was required for a more reasonably accurate network synchronization. The ensemble is actually played by the conductor, who controls density, frequency and spatial distribution of the sound textures, while the role of performers is minimal, limited to follow the conductor commands with no space for significant individual contribution.

The hierarchical use of the network may play a more salient musical role. Quite often, it may serve to structure an improvisation by triggering certain scenes or configurations or may be used by the conductor to balance and shape the dynamics of a performance. We will illustrate both approaches with two examples.

Adam Scott proposed three elements that create a work's identity, and which may exhibit a variable degree of determinacy: form (scale and structure), fixicity (or sonic identity) of the performers and concept or method of execution (from score-based to free improvisation)[Neal, 2008]. He proposes a number of different pieces for laptop ensemble with variable degrees of indeterminacy. In *Presets* (2008), the least determined one, players perform a free improvisation with indeterminated form. Their laptops are networked and player 1 acts as a conductor: by triggering certain cues, his computer reads the state from the other players and redistributes a certain parameter to all the other players, which is automatically updated. When no cues are being triggered, the quartet works effectively as a set of independent instruments, but cues lock temporally all players to player 1, unifying a certain timbrical parameter or musical state for all.

In *Autopoetics I* (Ted Coffey, 2009), players are free to input certain musical sequences, undo or clear them all into a cue which a central server processes. The *clear all* option had so much dramatic effects worked as a kid of anonymous sabotage. Some students asked to incorporate some accountability to identify user's actions and thus enforce discipline, while the conductor just asked for proper ensemble etiquette, but without enforcing norms or rules. [Reinecke, 2010]. On further iterations voting schemes or further constraints were incorporated to enforce a suitable balance between performers' activity and musical result. Needless to say that such unlimited power granted to players is unthinkable in a traditional ensemble. Ted continued this line of research with Autopoetics II (for laptop quintet) and III (in interactive installation format (2010).

On the other side, providing collective awareness through the network may facilitate user negotiation in more non-hierarchical scenarios. For example, Paula

Matthusen's *Lathyrus* (2007), another popular piece in the repertoire of several Laptop Orchestras, is a structured, improvisational game-like piece where performers must self-organize and seek a proper consensus on the patch chosen to navigate the musical areas available for exploration.

Game inspired pieces

As a co-located, synchronous activity relying on a computer network, it should not surprise that many composers highlighted the parallelism between computer music ensembles and multi-player LAN parties, an already well established social context for computer networking. Multiplayer games may be used to design a goal-oriented performance by imposing rules to performers, they may provide the infrastructure to design a shared performative environment, or they may inspire visual-driven multi-user interaction. We will show some relevant examples of each possibility.

In Smallwood's *On the Floor* [Smallwood et al., 2008], for example, players freely run a sonified game operating autonomously, trying to globally recreate a casino soundscape. The conductor monitors all the players and can change their individual game states to enforce a few global parameters such as the overall character and the piece duration. This is an extreme example of game sonification in which players are only expected to play a game, while the sonic environment is provided by the samples triggered and a certain musical dynamics is indirectly enforced by the conductor.

By re-adapting popular multiplayer environments, it is possible to rely on an already available rich infrastructure which provides opportunities to develop what Alvaro Barbosa calls Shared Sonic Environments[Barbosa, 2003]: synchronous or quasi-synchronous environments which allow for concurrent improvisation both in co-located and remote scenarios.

In *nous sommes tous Fernando...* (Robert Hamilton,2008) [Hamilton, 2009] players navigate within q3osc, a modified, OSC-enabled Quake 3 game engine. Instead of the rules provided by the game, players center on the musical possibilities provided by the sonification and spatialization of their actions, like throwing projectiles. Even if it is a freely improvisatory work, the virtual camera operator projecting the scene to the audience acts as a conductor, shaping the balance and spatialization of the ensemble and typing messages to performers onscreen. A similar approach may be seen in *GG Music*[Cerqueira et al., 2013], which sonifies Starcraft 2, another popular real-time strategy computer game. The sound engine collects gameplay data and extensive player's metrics, and the game engine allows for scheduling and parallelization of musical events, but the overall musical discourse is mostly dependent on the game rules and dynamics. As both environments don't allow for mutual sonic influence between players (each sonic event is attributed exclusively to a single player) we might consider the contribution of performers as summative.

Less constrained by the inherent game dynamics, Wang&Smallwood's *ChucK ChucK Rocket* is a shared environment built with the ChucK-based Audicle environment, consisting of mice moving in a grid. The visual interface lets players to share mices and direct mice movements by placing arrows and sonified obstacles. The conductor shapes the global sonic character by modifying some simple game rules like mice speed and density of mice, but he can directly instruct players as well. In contrast to the previous piece, players are encouraged to pay attention to the musical outcome of their collective actions[Smallwood et al., 2008]. Therefore, even if music arises from visual activity, there is an inherent tension: if the visual rules, it is game sonification, if the music drives it, the visuals are a kind of dance [Wang and Cook, 2006].

Physicality

The expressive use of physicality, both with the use of gestural interfaces and the expressive use of space by performers has been addressed in repertoire for laptop orchestras and more predominantly in repertoire for mobile ensembles.

For example, L2Ork's performance aesthetics emphasizes physical presence through gesture and choreography. Indeed, his conductor, Ivica Ico Bukvic, aims for an amalgamation of performer and conductor roles, resorting to performative gestures inspired on Taiji choreography. In *Citadel*(2010), for example, the whole ensemble gesturally follows the conductor to build a continuum of string chords by simulating bowing movements with Nintendo Wiimotes and Nunchuks[Tech et al., 2010]. This repertory is particularly suitable for undergraduates and even K-12 students, a specific educational target for the L2Ork.

A more sophisticated approach combining physical gesture and interdependence may be seen in *in Line* (Jascha Narveson, 2011), which explores gesturality as a required resource to enforce emergent synchronization between players. Modified Gametrak controllers [18], which allow for two-hand position tracking through retracting cable reels, are utilized by all the performers. Conceived as a structured improvisation, every performer may trigger independent melodies gesturally controlled. Pitch and texture may be locked by careful matching of players' frequency and phase of their autonomous processes respectively. The overall result is a slowly evolving musical and choreographic performance.

Combining the hierarchical, conductor-based approach of L2Ork repertoire with computer-mediated human interaction, the Physical Computing Ensemble explores the expressive use of physical interactions which are metaphors of real

[18] `https://en.wikipedia.org/wiki/Gametrak`

non musical actions, such as in *Skipping Stones*[Hattwick and Umezaki, 2012], where players perform throwing motion which triggers sound events. An automated score acts as a software conductor which sends timed cues to performers. They are invisibly sent to their WIImotes through vibrotactile feedback, acting as reminders to players, much as a prompter in an opera or theater performance.

As previously stated, mobile ensembles do make extensive use of gesture and space. For example, *SoundBounce* (Luke Dahl,2010) is based on the metaphor of a bouncing ball -a metaphor addressed as well in Just Continue to Move, by the aforementioned PCE ensemble. Performers throw sounds gesturally with their mobile devices, moving around the stage. This physically based interaction aims for a cooperative or competitive play and provides space for skill-based performance. The piece is structured as an improvisation which finally evolves to a sonified game. Performers interact spatially with independently generated sound sources, but at the end they can knock others' sounds, therefore the instrument becomes interdependent.

Interdependent networks

We will now survey some illustrative examples of repertoire specifically focused on the exploration of interdependent musical networks. While this is a performative paradigm more often addressed by ensembles adhering to the HUB model, a significant amount of repertoire for LOrk's explores network interdependence as well.

We will provide relevant examples which explore interdependence through sharing different musical resources, namely higher level musical structures, low level control parameters and audio objects.

Sharing musical structures *The PLOrk Tree* [Trueman et al., 2006] [Smallwood et al., 2008] explores interdependence and collective awareness though an ensemble of step sequencers in which performers may access and copy their neighbors' parameters according to a binary-tree topology. The conductor is at the root of the tree and sends a master clock to all the performers. Performers may see the state of their immediate root node, if they copy and modify it, the state will locally propagate up through the tree, eventually the conductor may access the state of all the terminal nodes, feeding back the information back into the network. A built-in chat system is equally graph-directed, with players being able to communicate to their immediate neighbors' and the conductor being able to communicate with everyone. The musical outcome, which alternates anarchy and the emergence of localized structures, is deeply related the topology of interdependence.

Dannenberg's *FLO*(2012) [Dannenberg, 2012] was possibly the largest scale composition ever written for laptop orchestra, targeting not one but six laptop orchestras (from Stanford, Texas, Boulder, Baton Rouge, Pittsburgh and Belfast). It was heavily inspired in HUB's *Borrow and Steal* and the author's former research with the Carnegie Mellon Laptop Orchestra, and will be the only composition specifically addressed to remote performance surveyed here. FLO offered an overlay, peer-to-peer network allowing players to access a Publish-Subscribe system for clock, chat, performer and conductor based messages. Performers could publish performance data (such as pitches or volume) and those subscribed to them would have their instruments constrained by their parameters -the actual musical realizations, however, did not seem to explore this concept to its full extent. The piece resorts to an asynchronous interaction scheme which may cope better with the inherent network delay of broadband area networking.

Sharing control parameters The Hub explored a number of non hierarchically and hierarchically structured collective instruments. *StuckNote*, written by Scot Gresham-Lancaster for the HUB [Gresham-Lancaster, 1998], is the paradigm of a non-hierarchical, heterogeneous interdependent network. In StuckNote each performer freely implements his own sound generator, which may incorporate only two control parameters: a volume control and a freely designed timbrical control. Both controls may be freely accessed and operated concurrently by all the performers in the ensemble. Two sources of unexpectedness arise when performing StuckNote: the manipulation of a parameter from a performer's sound generator by other performer, and the concurrent manipulation of a single parameter by more than one performer, giving rise to unexpected glitchy gestures caused by the merging of both MIDI controllers.

Shelly Knotts wrote *XYZ*(2011) for the BILE [Knotts, 2013], a piece heavily influenced by HUB's StuckNote. In XYZ therefore, as in StuckNote, the central issue is the challenging of performer's ownership of the sound by sharing its control among the ensemble. A minimum ownership is guaranteed, though, by not sharing the volume control of each performer's sound. XYZ provided three shared parameters -it must be noted that the complexity of interdependences increases geometrically with the number of shared parameters - gesturally controlled by accelerometer-equipped devices, such as game interfaces.

The piece consisted of an structured improvisation whose sections delimited distinct interactions between performers: fight sections (taking over other performer's control, generating glitchy passages as the parameter fluctuates between both simultaneous sets of controller values) alternating with passages where an automated process probabilistically warrants full control of that shared parameter to a single "winner" performer. As previously stated, the concurrency of several gestures on a musical control generate novel gestures which can hardly be attributed to the performers. It is also hard for the audience to understand the whole process. Therefore Knotts suggested using visuals inspired on video-game aesthetics to emphasize the theatrical elements of the piece.

While not properly a work but a performative environment, Graham Booth's *Inclusive Interconnections* [Booth, 2010] is a framework designed to explore flexible reconfiguration of player interdependencies in heterogeneous networked ensembles. Much in the Hub style, Inclusive Interconnections is basically decentralized, allowing performers to implement their own sound generation processes. At the same time, the performers themselves may modify the topology of interdependence, freely taking, giving, exchanging or sharing influence with other players.

While the former examples were designed with highly specialized computer ensembles in mind, Nathan Bowen's *4Quarters* (2012) [Bowen, 2012] is an example of network composition targeted towards audience participation. It was intended to go beyond score-matching games or simple event triggering, giving to the participants the possibility to exert a greater and more creative influence on the musical outcome by setting up an *ad hoc* improvisation. The audience could control a central server from a TouchOSC based interface on their mobile devices. A central projection screen provided feedback of all the performer's contributions. The performers themselves could join one of four teams to concurrently select sound files and alter their volume, panning and equalization.

Sharing audio objects In *GroupLoop* [Ramsay and Paradiso, 2015], Ramsay explores ensemble collaboration through a browser-based feedback performance system. Performers may freely shape their own sound by controlling the gain and equalization of their own feedback loop as well as send and receive audio from the other players. The software establishes a fully connected peer-to-peer network topology allowing real-time audio streaming through WebRTC. As shown in its premiere, it extends seamlessly to remote performances.

Because of the unstable nature of audio feedback, and because there was no global awareness of the amount of audio signals fed back to every laptop, the performance was notoriously difficult to control according to the author. In terms of interdependence, GroundLoop may exhibit a wide range of topologies, from independent feedback instruments to complex feedback paths involving several laptops, therefore the contribution may be summative, multiplicative or multi-directional.

As a final example of interdependent instrument which incorporates both shared control and shared sound objects we could mention *Telephatic* (Scott Wilson,2011) [Wilson et al., 2010]. Conceived as a structured improvisation for live-coders, Telephatic provides performers a Republic-based environment which allows them to improvise sharing both code and audio objects. Additionally, a server computer provides a master clock to which performers can optionally synchronize to the desired amount, allowing for a complex rhythmic interplay. The network facilitates coordination and formal structure as well, indeed, its identity as a work lies not on the varying compositional materials but in the specific behaviors encouraged by the constrained interaction scenarios. The server sets up a minimal framework to structure the collective improvisation, for example enforcing a decrease in the quantization base for generated events and a final *accelerando*.

Finally table 3.5 summarizes the characterization criteria for the repertoire surveyed, focusing on the more relevant properties to identify the topology of the multi-user instruments: their architecture (homogeneous or heterogeneous, and centralized or decentralized), the social balance of roles (whether there is a potential unbalance between performers, beyond the role of a conductor) and the characteristics of the interdependence: whether it is summative, multiplicative or multidirectional, the amount of shared resources and the extend to which such resources are actually shared.

Table 3.5: A survey of network music repertoire for Computer Ensembles

Year	**Piece**	**Ensemble**	**Composer & ref**	**Contribution**	**Players**	**Role balance**	**Texture**	**Centralization**	**Resources shared**	**Sharing**
1995	StuckNote	HUB	Gresham-Lancaster[19]	Multiplicative	3	Balanced	Heterogeneous	Decentralized	Two out of two	Complete
2008	Presets	ad hoc	Scott Neal [Neal, 2008]	Summative. Multiplicative.	4	Unbalanced	Homogeneous	Decentralized	Any (one at a time)	Complete (enforced)
2006	The PLOrk Tree	PLOrk	Trueman [Trueman et al., 2006]	Multiplicative	15	Unbalanced	Homogeneous	Decentralized	Any (tree topology)	Complete
2006	CliX	PLOrk	Wang [Smallwood et al., 2008]	Summative	variable	Balanced	Homogeneous	Centralized	No	No
2006	ChucK ChucK Rocket	PLOrK	Wang [Wang and Cook, 2006]	Summative	variable	Unbalanced	Homogeneous	Centralized	variable (game rules)	Complete (enforced)
2006	Non-specific Gamelan Taiko Fusion	PLOrK	Trueman [Wang and Cook, 2006]	Summative	15	Balanced	Homogeneous	Decentralized	No	No
2007	Lathyrus	FLEA	Matthusen [Matthusen, 2009]	Summative	4	Balanced	Homogeneous	Decentralized	No	No
2008	Nous Sommes tous Fernando…	SLOrK	R Hamilton [Hamilton, 2009]	Summative	variable	Balanced	Homogeneous	Centralized	No	No
2008	On the Floor	PLOrK	Smallwood [Smallwood et al., 2008]	Summative	15	Unbalanced	Homogeneous	Decentralized	1	Complete (enforced)
2009	Autopoetics I	PLOrK	Coffey[20]	Multiplicative	15	Unbalanced	Homogeneous	Centralized	All	Complete
2009	Supreme Balloon	PLOrK	Smallwood [Fiebrink et al., 2007]	Summative	variable	Unbalanced	Homogeneous	Decentralized	No	No
2009	SoundBounce	MoPho	Dahl [Oh et al., 2010]	Summative, multiplicative	variable	Unbalanced	Homogeneous	Decentralized	1	
2010	Citadel	L2OrK	Bucvik [Tech et al., 2010]	Sumative	variable	Balanced	Homogeneous	Decentralized	No	No
2010	Skipping Stones	PCE	Hattwick [Hattwick and Umezaki, 2012]	Sumative	6	Balanced	Homogeneous	Centralized	One	Complete
2010	Inclusive Interconnections	HELO	Booth [Booth, 2010]	Variable	3 to 6	Variable	Heterogeneous	Decentralized	One	Variable
2011	XYZ	BILE	Knotts [Knotts, 2013]	Multiplicative	6	Balanced	Heterogeneous	Decentralized	Three out of four	Complete
2011	in Line	Sideband	Narveson [Narveson and Trueman, 2013]	Multi-directional	6	Balanced	Homogeneous	Decentralized	Two	Partial (negotiated)
2011	Telephatic	BEER	Wilson [Wilson et al., 2010]	Summative, multiplicative	3 to 5	Balanced	Heterogeneous	Decentralized	2	Variable (user selectable)
2012	FLO	6 LOrKs	Dannenberg [Dannenberg, 2012]	Multi-directional	+60	Variable	Homogeneous	Decentralized	Any (variable)	undocumented
2012	4Quarters	audience	Bowen [Bowen, 2012]	Summative. Multiplicative.	variable	Variable	Homogeneous	Centralized	Any	Complete
2013	GG Music	SLOrK	Cerqueira [Cerqueira et al., 2013]	Summative	3	Balanced	Homogeneous	Centralized	none	none
2014	GroundLoop	ad hoc	Ramsay [Ramsay and Paradiso, 2015]	Summative, Multiplicative, Multi-directional	5	Variable	Homogeneous	Decentralized	All(audio stream)	Variable (user selectable)

Notational approaches for Laptop Orchestras

A desirable goal of notated scores for Laptop Orchestras would be the dissemination of their repertoire. Two approaches are possible here: either seeking unified setups to guarantee hardware and software compatibility, much in the line of the meta-instrument methodology of the Princeton and Stanford Laptop Orchestras, or possibly resorting to scores which are implementation independent, relying on the performer for the actual implementation of the required instruments, as advocated by HELO laptop orchestra [Hewitt and Tremblay, 2012] and generally considered as a better recipe for perdurability [Baguyos, 2014] The former use to provide packages consisting of software environments for generic computer music performance software, setup notes, performative instructions and textual indications which serve as an illustrative score. The later seeks to provide scores with details on the required instrument capabilities and performance commands. Of course both approaches are not mutually exclusive.

In his PhD Thesis, Scott Hewitt [Hewitt, 2014] evaluates both graph and text based approaches to notation of electroacoustic music. An interesting remark is the fundamental link between graph scores, text scores and computer code. Ultimately, any precise and unambiguous specification of gesture, timbre and instrument characteristics can be implemented as computer code. At such level of specification, graph and text scores are interchangeable, being the choice a matter of readability and ease of use. However, only if there is ambiguity (in the events or in the instrument specification), repeated performances make any sense.

Common Western Music Notation is rarely used for Laptop Orchestra scores. It may be efficient provided a shared understanding of this notation and can be easily adapted to different pitch-based instruments following a note-based paradigm, for example analogue synthesizers or acoustic instruments [Hewitt and Tremblay, 2012]. Incidentally, Western Notation may be a useful resource when the score includes musical citations or patterns which may be easily transcribed, such as in Dan Trueman's PLahara [Smallwood et al., 2008].

More open approaches are possible with graphic notation, a resource widely used in the repertoire of laptop orchestras. Graphic notation may be employed as well to indicate literal mappings directing physical gesture, becoming a script for choreography [Rebelo, 2010]).

Some authors consider that performance patches for Laptop Orchestras are the main specification of both the composition and score. In a way, they equate instrument design to score, because instrument affordances and constraints define timbrical boundaries and indirectly shape the performance. This compositional approach in which the description of the piece is mostly the specification of the instruments involved, is characteristic of some approaches to Laptop Orchestra composition. The founders of the Princeton Laptop Orchestra fully discuss this approach in [Smallwood et al., 2008].

Real-time notation for Laptop Orchestras Several interactive composition systems are capable of real-time notation generation. This opens a new range of creative possibilities, such as the concurrent generation of scores by algorithms and musicians, the automatic transcription of improvisation or the active involvement of the audience in the generation of musical scores or in the control of algorithms for score generation, which sets up a feedback loop between performers, scores and audience [Freeman and Street, 2008].

The Quintet.net, by Georg Hajdu is possibly the first networked music environment utilizing real-time notation, and was being actively developed since 1999, in close relationship with the artistic endeavors of the European Bridges Ensemble. The environment does allow for a wide variety of notational approaches, from static to dynamically generated scores, utilizing any combination of textual instructions, open graphical notation and standard music notation [Hajdu et al., 2011].

Another environment supporting real-time music notation for networked performance is Jason Freeman's LOLC [Freeman et al., 2010]. It allows for mixed performances in which laptop musicians generate notation on-the-fly for instrumental musicians, expected to sight-read in performance.

Score for instrument design: the case of live coding Live coding is the primary instrumental practice for a number of laptop orchestras. A particular paradigm of musical notation is the use of code both for documentation, performance and interface, in the specific case of live coding practices. As Hewitt notices [Hewitt and Tremblay, 2012], there is a wide range of possible approaches, from challenging the role of the performer to letting him focus on higher level processes and deferring actual note triggering and micro gestural execution to generative algorithms, for example, or any strategy in-between. Magnusson, indeed, sees live coding as the natural continuation of the XXth-century tradition of experimentation with the musical score [Magnusson, 2011a].

If computer code is to be considered as score and instrument, we are actually blurring the distinction between performers, composers and *luthiers*. See also [Kaltenbrunner et al., 2004], who observes that, in a dynamic patching performance, building an instrument is conceptually equivalent to playing it and rebuilding that instrument from scratch could be compared to the reproduction of a musical score.

Live coding, usually resorts to video projection of text code to allow the audience to get a sense the process of development of the performance, and at the same time adhering to an open aesthetics which avoids the obscurantism of processes hidden from the audience [Zmölnig and Eckel, 2007].

However, as only a fraction of the audience is expected to actually understand the lines code as the piece develops, there is a strong aesthetic statement in this act. Collins note that the arcane text coding systems might look appealing, but prevents against a snobbish, deliberate obfuscation of musical processes [Collins et al., 2003]. To improve code understanding, there is the possibility to provide a more verbose or better illustrated representation of the code to the audience [Troyer et al., 2012].

3.4 Conclusion

The aim of this chapter was to review the relevant literature on computer ensembles, framed within the traditions of electronic, computer and network music, and focusing on characterizing ensemble models and surveying instrumental taxonomies and evaluation criteria.

We first historically contextualized the origins of the first Network Music ensembles and their main contributions. Then, we defined, characterized and reviewed design and evaluation criteria for multi-user instruments, the most archetypal instrument paradigm

Then, we introduced the broad topic of Laptop Orchestras, a nowadays well established instrumental ensemble. Two well delimited, but often overlapping ensemble models were identified and characterized, the tentatively called HUB and LOrK models respectively.

We reviewed the most relevant literature on Laptop Orchestras and related computer ensembles, particularly on sound, network and programming infrastructure and the diversity of compositional and notational approaches. The review proved useful for understanding the role of scientific and technical research in the genesis of a new ensemble identity.

In the next chapter, we will present an overview of the Barcelona Laptop Orchestra as our particular approach to network music performance. A detailed review of the repertoire developed will serve to illustrate many of the instrumental taxonomies reviewed above, and will show how design criteria is inextricably linked to the very creative process, the performative experience and the aesthetic goals pursued in a computer music ensemble.

Chapter 4

THE BARCELONA LAPTOP ORCHESTRA

4.1 Introduction

This chapters contextualizes the taxonomies of multi-user instruments discussed in the former chapter in the particular creative praxis of the Barcelona Laptop Orchestra (henceforth, BLO), which strongly adheres to the HUB ensemble model, though incorporating technical resources and compositional approaches proposed by Laptop Orchestra related research over the last decade.

The purpose of this chapter is twofold. First, to illustrate the principles of network music throughout the performative practice our Laptop Orchestra. And second, to provide preliminary reflections on the challenges and opportunities of network music in terms of performative engagement and creative outcome.

This chapter is structured in three broad categories of repertoire, according to the main multi-user instrument paradigms which oriented its design. Though such paradigms are by no means mutually exclusive, each one brings out a number of relevant issues concerning creativity in performance.

- *Score-based pieces* illustrates the new approaches in Music Notation for performance, from graphical open scores to score as computer code and mixed approaches.

- *Collective Sequencing* explores the concurrent edition of musical patterns to evaluate

- *Gestural Interfaces* highlights the expressive potential of gestural expressivity in a collective performance

- Finally, *Interdependent Networks* explores the implications of shared resources and interdependence of musical processes.

Each category is illustrated with a number of designs we have developed and explored throughout these years.

4.2 The Barcelona Laptop Orchestra. Genesis and motivations

The Barcelona Laptop Orchestra (formerly Esmuc Laptop Orchestra) was established in 2008 and celebrated its premiere performance in June of 2008. It was conceived as the instrumental ensemble of the Sonology Department from the Escola Superior de Musica de Catalunya (ESMUC)[1], to give prominence to the instrumental and creative practices carried out within the department. The membership comprises teachers and students from this institution and researchers from the Music Technology Group (Universitat Pompeu Fabra)[2], as well as freelance artists and programmers.

The Barcelona Laptop Orchestra, as otherwise typically seen in similar ensembles, soon became an active meeting point for sound researchers, *digital luthiers*, performers and composers. It focused on the development of new musical interfaces and new performative paradigms, with the laptops playing a central, but not exclusive role. At least, these were the original goals. Eventually, interface development turned out to be less prominent, favoring the research on network music paradigms instead.

Since its inception, the BLO tried to build bridges between two seemingly disparate traditions and contexts for computer music performance: the experimental avantgarde around the network music tradition, and the electronic music practices featured in popular electronic music genres and festivals. Therefore, the ensemble both played in electroacoustic venues, computer music related conferences and electronic music festivals. Just to mention a few, the BLO performed at the Sound and Music Computing Conference (Porto, Portugal, 2009, Barcelona, 2010), the Network Performing Arts Production Workshops (Teatre del Liceu, Barcelona, 2011), at the Sonar Festival (MACBA, Barcelona, 2011 and Fira de Barcelona, Barcelona, 2013), at the Network Music Festival (Birmingham, UK, 2012), at the Festival Mixtur (Fabra i Coats, Barcelona, 2013), at the Esmuc (L' Auditori, Barcelona, 2013) and at the Mobile World Center (Barcelona, 2013), offering occasional workshops on laptop ensemble performance as well. Since 2013 it is the

[1] `http://www.esmuc.cat/eng/The-School/Departments/Sonology`
[2] `http://www.mtg.upf.edu/`

Resident Orchestra at the Phonos Project[3], where it performs regularly once or twice a year.

The BLO adheres to the HUB model characterized by smaller, chamber like computer ensembles with a strong collaborative approach in repertoire and instrument design. The current repertoire features free and conducted improvisational pieces exploring distinct multi-user instruments and network topologies. Additionally, the ensemble has premiered over a dozen commissioned compositions and re-arrangements of classics of the XXth century musical avantgardes, written or adapted by members of the group as well as other prominent composers.

It is currently co-directed by Alex Barrachina and Josep M Comajuncosas. Indeed, a distinctive feature of the BLO compared to other ensembles adhering to the HUB model is the regular presence of a conductor, both for rehearsals and performance. In the next section we will discuss the motivations behind the decision to rely on a conductor for our performances.

4.2.1 BLO Repertoire analysis

In this section we will perform a comparative and evolutionary analysis of a representative sample of our repertoire.

The repertoire selected is listed in Table 4.1, which summarizes the main characteristics of the surveyed pieces in the perspective of the previously reviewed taxonomies for collaborative musical interfaces and multi-user instruments.

The CSCW inspired scheme proposed by Barbosa and extended by Xambó lets classify our repertoire as *co-located* and *synchronous*. Turn-taking scenarios don't challenge the essentially synchronous nature of the musical activity on the long term. Xambó proposed a mixed category of simultaneously synchronous and asynchronous interaction in her thesis[Xambó, 2015]. Only in one specific case, our arrangement of H.C.Steiner's Solitude for Laptop ensemble (4.2.2), we might consider that interaction takes place both synchronously (performers compose their parts concurrently) and asynchronously (each performance loops through the same excerpt, which is incrementally reworked).

In view of Weinberg taxonomy of interconnected musical networks [Weinberg, 2003], our repertoire shares some similitudes to *collaborative instruments for novices*, because it emphasizes performers' experience over product, offering a low floor learning and suitable for short interactions. This it the case of Nuvolet, which might resemble a sound installation, and Six Pianos or Clix ReduX, as they may be played with low interactions between performers and guarantee a successful outcome with little effort because of their built-in constraints. But even such instruments have notable potential for experts.

[3] http://phonos.upf.edu/

However, as we will see, most of our repertoire explores more complex interdependences, do have a higher ceiling learning and allow for longer performances, with a stronger *focus on performance* and only occasionally tailored for large audience accessibility. Ultimately, when we achieved enough mastery with certain collective instruments, our approach was more *product oriented*, but with a clear focus on the performance dynamics rather than on actual music products.This was the case with La Roda and The Hub, out most explored instruments (see 4.2.5 and 4.2.5 respectively).

In terms of size, all our repertoire is scaled for *small scale* (4-8) ensembles, according to the average number of stable members of the Barcelona Laptop Orchestra. Interdependent works are highly *process-centered*, with a simultaneous focus on social and creative experience. Finally, the topologies are diverse, both *centralized* and *decentralized*, with different levels of hierarchy according to design goals.

The multifaceted design criteria which Blaine and Fells provide to categorize collaborative musical interfaces are useful to identify and outline the diverse performative contexts of our repertoire. From this perspective, the analyzed repertoire mainly consist of *local setups* (with occasional use of remote performances), with players sharing *uniform responsibilities* in *free or directed interaction*. While some pieces employ gestural interfaces, we mainly resort to *native input devices* for user interaction. As for media employed in performances, our orchestra usually combines sound with visuals and light aimed to inform the audience on the performer's activity, but occasionally we address the performance as a mostly acousmatic experience as well. Finally, the target musical genres are diverse, from electroacoustic improvisations to League-style open explorations and more dance oriented, IDM realizations for large audiences.

Jordà highlights a number of relevant considerations for multi-user instruments. Beyond what has already been discussed in the former taxonomies, a deeper analysis on the nature of mutual interaction in our repertoire is possible through Jordà considerations. Some of our repertoire are actually shared performance environments without mutual interdependence, that is, a parallel, *summative* kind of multi-user instruments. For example some of our collective sequencing instruments like CliX Redux and our multiplayer gestural interface Nuvolet belong to this category. Some others do allow for more intense interplay though mutual interdependence. La Roda exemplifies a *multiplicative*, *single-threaded*, circular, unidirectional interdependent network of sequential nature and *static* but *essential roles*, Streams is a *multiplicative*, *multi-threaded*, bidirectional and sequential interdependent network with *role flexibility* but *non essential roles* and the Hub is a *multiplicative*, *multi-threaded* , bidirectional and concurrent interdependent network with *role flexibility* and *non essential roles*.

Table 4.1: Taxonomy summary of the surveyed Barcelona Laptop Orchestra's repertoire

Year	Piece	Contribution	Target	Performers	Role balance	Texture	Centralization	Real-time	Synchrony	Resources shared	Sharing
2010	Rimandi	Sumative	Experts	4	Balanced	Homogeneous	Centralized	Real-time	Synchronous	none	none
2012	Solitude	Sumative	Experts	5	Balanced	Homogeneous	Centralized	quasi real-time	Synchronous	none	none
2013	Variations II	Summative	Novices	5	Balanced	Homogeneous	Centralized	scheduled	Synchronous	none	none
2013	BLO_2013	Summative	Experts	4	Balanced	Heterogeneous	Decentralized	Real-time	Synchronous	none	none
2013	CliX ReduX	Sumative	Experts, Novices	variable	Balanced	Homogeneous	Centralized	quasi real-time	Synchronous	none	none
2013	Streams	Multidirectional, Multiplicative	Experts	3 or 4	Unbalanced	Homogeneous	Centralized	quasi real-time	Synchronous	variable	none
2014	Six Pianos	Sumative	Novices	6	Balanced	Homogeneous	Centralized	quasi real-time	Synchronous	none	none
2011	Nuvolet	Summative	Novices	3	Unbalanced	Homogeneous	Centralized	Real-time	Synchronous	1(corpus)	none
2014	Nuvolet2	Summative	Experts	3	Balanced	Homogeneous	Centralized	Real-time	Synchronous	1(corpus)	none
2010	La Roda	Multiplicative	Experts	4 to 8	Unbalanced	Heterogeneous	Decentralized, later Centralized	Real-time	Sequential	1(audio)	none
2014	Conducted Roda	Multiplicative	Experts	4	Balanced	Heterogeneous	Centralized	Real-time	Sequential	1(audio)	none
2015	Conducted Roda	Multiplicative	Experts	3	Balanced	Homogeneous	Centralized	Real-time	Sequential	1(audio)	none
2012	The Hub	Multidirectional, Multiplicative	Experts	variable	Unbalanced	Homogeneous	Centralized	Real-time	Synchronous	variable	partial
2014	The Hub proxy	Multidirectional, Multiplicative	Experts	4	Unbalanced	Heterogeneous	Centralized & Decentralized	Real-time	Synchronous	2	variable

4.2.2 Score-based pieces

We will begin this survey with Score-based pieces because, even if a piece realized into a musical score might be conventionally considered as a work with a rather stable entity (as we discussed in Chapter 1), we have nonetheless incorporated many of the notational approaches discussed in Chapter 3, including open notation, notation as computer commands and live notation. They are therefore illustrative of the diversity of approaches and the kind of *fluid identity* that musical works for computer music ensembles may offer.

Rimandi (2010)

Rimandi (Ivano Morone), commissioned by the BLO, is a work for laptop quartet in which performers manually operate on piezoelectric elements attached to their laptops and, at the same time, control a series of ring-modulators in parallel that process the sound coming from the piezos. The score (see Figure 4.1) consist of a permutation of gestural blocks of precise duration. It incorporates two notational strategies as described by Hewitt in[Hewitt, 2014], namely, commands for physical gestures and accurate, platform-independent instructions for sound synthesis processes.

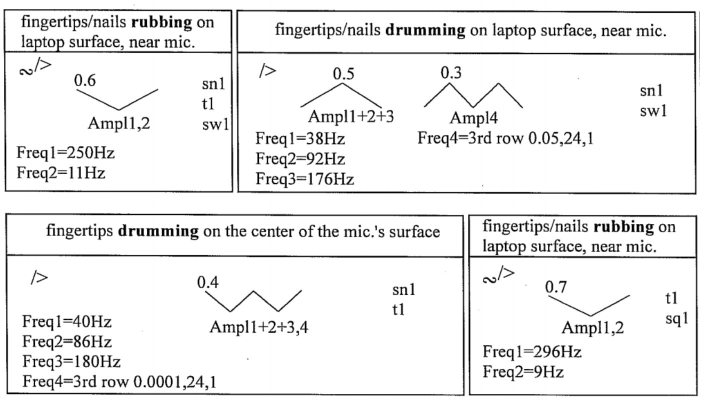


Figure 4.1: Rimandi, by Ivano Morone. Excerpt of the score.

The composer provided a standalone software for performance but we eventually implemented a centralized version which managed the score, scheduling the appropriate instructions to players, for a more precise alignment to the score and at the same time automating some rather mechanical tasks, thus liberating performers of the need to operate with their graphical interfaces for scene changes or slider movements.

This is an example of piece in which the level of specification of certain aspects in the score is unambiguous enough as to be automated, therefore getting rid of the human performer and addressing commands directly to the software. Had it been completely specified this way, the performance would have been absolutely superfluous of course -unless the theatrical aspect of seeing performers extremely busy justified it!

The role of the remaining, not automated notated commands is to provide loose gestural orientations which give some space for improvisation. For example a command *with oscillating gestures* is ambiguous enough with respect to the evolution of the oscillatory speed and depth in that passage. Similar commands are found throughout the score, such as *tapping the keys* or *scratching the piezo surface*. When players were close enough or their manual activity was projected on a single screen onstage, a parallel discourse emerged, for players learned to imitate from, or suggest to other players gestural variations. We thing that the performance under this context was more varied and cohesive at the same time.

Solitude (2012)

Even if Rimandi was successful at keeping performers busy by following the score commands and allowing them to freely implement their gestures, playing it soon became a routine job. The was no enough *space for improvement* and the score was not complex enough to keep sustained interest. Eventually we resorted to different notational approaches which could challenge performers and promote a more active involvement.

We tried a different strategy with the admittedly free version of Solitude, a graphical score for computer performance by Hans-Christoph Steiner[4]: collective real-time scoring. While Rimandi provided limited, *ornamental* improvisation opportunities to the performer, Solitude relies on the performers for the complete specification of the piece in real-time. We will now briefly present the original piece, our adaptation for real-time ensemble performance and finally a comparative discussion of its performative challenges.

Steiner's Solitude combines minimalist textures with rapid sonic transformations. Timbrically it explores the contrast between time-stretched samples, one with very tonal characteristics and the other more clearly timbrical. It was implemented using PureData graphical structures. Thus the score generates, and completely defines the piece (except for the characteristics of the audio samples utilized). Colors represent samples, and each sample comprises two timelines, the bigger, upper one functions as an stochastic mask for the time pointer for the sample to be played, while the lower, thinner one controls panning and amplitude.

[4] `http://at.or.at/hans/solitude/`

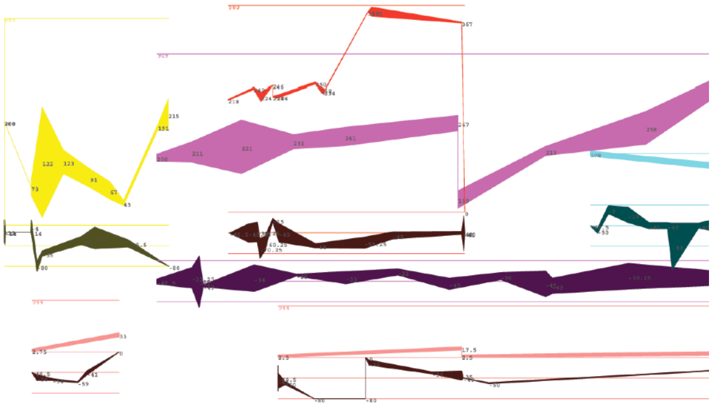


Figure 4.2: Solitude, by H.C.Steiner. Excerpt of the score.

Inspired by the graphical notational approach of the piece (see Figure 4.2), we implemented a set of networked graphical interfaces and an proxy in the original PureData patch to allow an ensemble to concurrently edit new scores in realtime. Each player was in charge of the temporal evolution of one of the sample granulators, and all the players where looking at the global score as it was being drawn on a projection screen which the audience could follow as well. Due to the resolution limits of the projection screen, the score was rendered as a 1-minute loop, which was being retriggered several times.

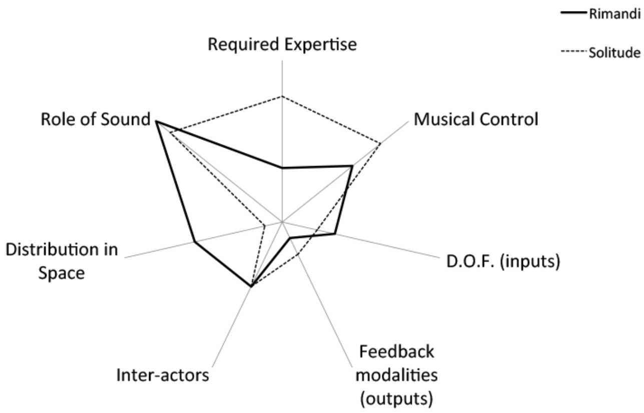


Figure 4.3: Dimension Space plots for Rimandi and Solitude instrumental setups

In terms of multi-instrument design, we actually shifted from a rather decentralized setup in Rimandi, in which performers play independently on their laptops and the only role of the server is to facilitate performers' synchronization and partially automate the performance, to a completely shared workspace in which the server centralizes the sound synthesis. But resources are still exclusive and there is no mutual interdependence: each performer is in charge of an independent score for his assigned sound sample.

As an improvisational environment, Solitude offered an interesting challenge to performers: to be able to incrementally edit their part within the 1-minute span required to go through the score. Performers could adjust the sample pointers to search for novel timbrical features, or the whole shape to modify the playing speed. They were encouraged to listen to the overall mix to balance their part and to look at the whole score to fill the gaps or give space to others, to detect interesting gestures or patterns in others' parts or in the whole score and think of suitable complementary or contrasting ideas.

Admittedly, the near realtime constraints of the loop-based performance pose a severe challenge on such compositional procedures. This kind of realtime score-based collaborative sequencing places additional cognitive demands to performers, who must quickly decode others' contributions from a score and map their ideas to precise graphical representations. The brief lapse of time available to react to, or anticipate to others' events only partially overcomes those limitations.

Playing Solitude clearly showed a tension between the desired level of immediacy for quick feedback and comparison between rounds and the need for enough time for proper planning. Indeed, the approach shows the difficulty to balance the tension between two opposing forces: a realtime context that impels performers to react to others' score gestures in a quasi improvisational style together with a collaborative working environment which is optimally suited for a careful process of crafting together the most accurate timelines.

Compared to Rimandi, we concluded that the performative freedom offered by Solitude didn't compensate the lack of instrumental performability. This was a consequence of the design. As we see in Figure 4.3, we traded physicality, immediacy, expressive musical control and easiness of use just for a slight improvement in performative freedom.

Variations II

Still keeping the collective real-time scoring idea, but with an approach closer to an audiovisual interactive installation, our next notation-based piece was a free realization of Variations II by John Cage (1967), a work also addressed by The HUB[5]. Variations II was originally implemented for a single computer by William

[5] http://crossfade.walkerart.org/brownbischoff/hub_texts/var2_f.html

Brent. The author kindly provided a network version of the piece which allowed for concurrent elaboration of scores which where played concurrently.

Variations II is an archetypal example of Cage's abstract graphical works, both combining elements of chance music and absolute predetermination. The performer must first elaborate the score by freely arranging points and lines on transparency sheets, then he must measure some distances which are mapped to a number of musical features which serve to realize the score. Brent's implementation took care of the score computation and final rendering - with a set of percussion samples- and performers only had to draw the sketches. The server scheduled the sketches, and the task of the performers was to feed the server fast enough as to keep the score renderer busy. All the scores being played were shown on a projection screen, with the score elements displaying in synchrony with the associated sound events(see Figure 4.4).

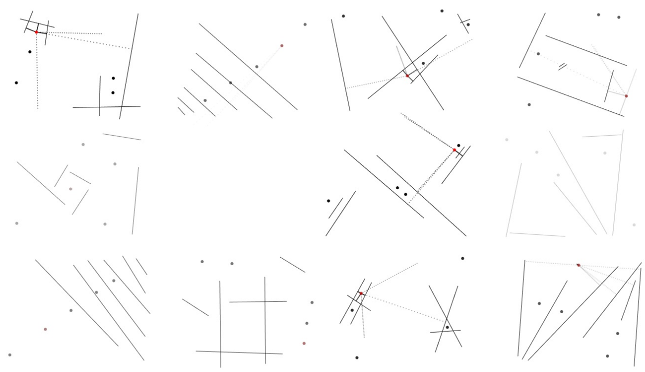


Figure 4.4: William Brent realization of Cage's Variations II. Projected score

In Solitude, as a looped sequenced score, a minimal amount of interaction was possible between players as they could modify the score being played in account for others' contributions, both before the software player reached a certain passage or planning for the next iteration. In Variations, there were less opportunities for interaction: once the score is written a complete specification for the performance is set up, sent to the server and scheduled for performance.

From the audience viewpoint, the use of scores to engage them in the performance was not entirely successful. Showing actual scores to the performance is an invitation for a more active listening provided the mapping is properly understood, which was not the case. The causality is not always clear either. Some attendants reported they were not sure whether the graphics projected were generated before or simultaneously, neither if they were driving the musical outcome or they were concurrently generated by audiovisual algorithms. Ironically, both performances turned out to be more believable because the performers looked incredibly busy while drawing new sketches.

BLO_2013

The former approaches to engage the performer in a score-based work was, indeed, turning the performer into a composer, and deferring the performance to the computer. Drawing from previous experiences, we resorted to a notational approach which would return to the performers their original role but at the same time challenging them with a more complex and open score.

The next season the German composer Orm Finnendahl collaborated with the orchestra in the production of a collective improvisation with his networked software environment. BLO_2013 is both a score and a canvas for improvisation [Finnendahl, 2012]. The scores are computer generated by recursive generative algorithms written by the composer. A central computer triggers the scores to be performed but otherwise each laptop acts as an independent sound-processing station. The score (see Figure 4.5) indicates to every performer when the computer starts and stops recording him, and displays when it will play back time-stretched and/or frequency-shifted replicas of the recorded passages. To minimize the notational complexity, the author resorts to color schemes to distinguish the sets of replicas associated with distinct recordings.

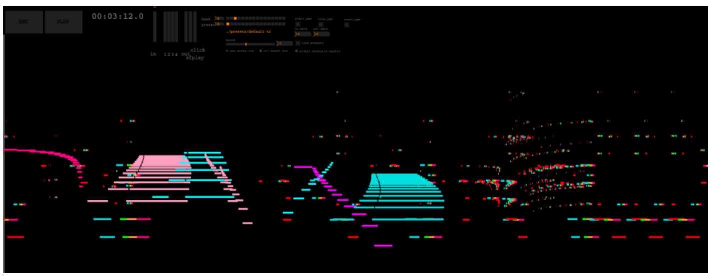


Figure 4.5: BLO_2013 by Orm Finnendhal. Score for section II

While composition is mainly addressed to expert players with solid experience in contemporary improvisation techniques, in was successfully addressed by the BLO with intensive tutoring by the composer. Performers freely decided what sound generators would utilize: from toy instruments, to electronic guitars, synthesizers or non musical sonorous objects.

Compared to Rimandi, the score itself is not a set of commands for performance, but a graphical representation of the automation scheduled for performance. It actually works as a framework for structured improvisation, but with elements of interactive performance, and even of chance music. The performer is expected to rehearse the score to be able to accurately follow the timed processes and develop successful strategies to cope with them. Any instrumental gesture will have long term consequences and the performer may plan in advance all sort of musical dialogues, counterpoints and polyphonic overlaps between what is being currently played and the eventual replicas of former recorded passages. Eventually, when the amount of different and concurrent replicas is sufficiently high, the computer takes over the performance as no human performer could actually plan and accurately follow its complexity.

4.2.3 Collective sequencing

Some of the aforementioned approaches to collective scoring, particularly Solitude, may be considered as well as examples of collective sequencing, the concurrent modification of musical patterns. The Barcelona Laptop Orchestras has been investigating this particular performative scenario with a number of less notation-centric pieces which explore distinct representations of the shared environment and its musical outcome.

Half way between collective composition and performance, concurrent edition of sequences and cyclic patterns do provide opportunities for interaction and improvisation in a quasi-realtime performative environment. Collective sequencing and pattern editing incorporates elements of traditional composition, such as the ability to prepare structures in advance or rework on the ones already developed, and elements of live improvisation, such as the short time available to take musical decisions and the opportunity to quickly react to others' contributions throughout the performance.

Concurrent edition of musical patterns serves as well as a useful platform for research in experience-based design for collaborative musical interfaces. We should mention here the studies of Nick Bryan-Kinns [Bryan-Kinns and Hamilton, 2012] [Bryan-Kinns, 2013] around musical collaboration and mutual engagement in free improvisational scenarios with Daisyphone, his semi-synchronous, web-based collaborative environment which allows concurrent edition of cyclic musical patterns (see Figure 4.6).

CliX ReduX (2013)

CliX ReduX, the first piece/instrument based on collective sequencing which we will discuss, is loosely inspired in a classic in the repertoire of Laptop Orchestras, the aforementioned CliX by Ge Wang[Smallwood et al., 2008]. Because of the simplicity of the setup and open structure, CliX has become one of the few composed pieces for Laptop Orchestras adopted by a number of different ensembles in their repertoire.

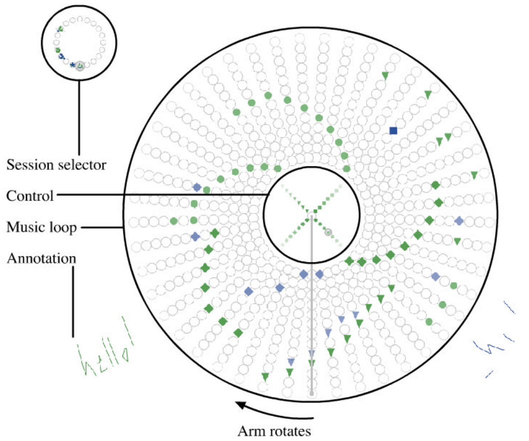
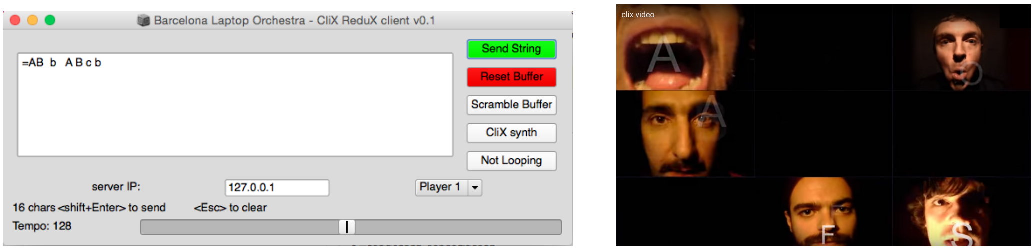


Figure 4.6: Bryan-Kinns' Daisyphone, a collaborative editor of musical patterns.

The original piece, implemented ChucK, consisted on a conducted improvisation in which performers type their laptop keyboards to trigger click sounds. The conductor instructed players when to play (to achieve dramatic spatialization effects) and the typing speed and *frequency* (which was mapped to keys according to their ASCII code). A centralized clock enforced synchronicity though, as the SLORK fellows and ourselves noticed, this was greatly interfered by jittering caused by the wireless networking infrastructure.

After some initial rehearsals players reported that, apart from the aforementioned technical issues which would require either a wiring network or a centralized DSP, the instrument lacked performative engagement, at least for a small sized ensemble. Arguably it is used to great effect by Ge Wang with much larger computer ensembles, compensating the limited performability with a consistent use of global textures and spatialization of performers. But in smaller groups such a restricted interaction scheme and audible result could hardly be compensated by the cumulative effect of many players following the conductor's gestures at once. Another disturbing factor is the simultaneous use of a realtime typing action with a buffered and re-quantized triggering of the keys, which ultimately challenges the agency of the performance.

Our piece greatly departed from this original concept, eventually becoming a structured improvisation performed with an audiovisual instrument, and just keeping from the original piece the interaction model. At implementation level, it consist of centralized video and audio applications, respectively written in OpenFrameworks and SuperCollider, which play back the video snippets triggered by the performers. Each performer has a player interface in which he pre-buffers text strings which may be easily copied, looped and scrambled to generate a wide peri-



(a) Client interface (b) Projected matrix

Figure 4.7: CliX ReduX. A client interface and the projected video matrix for all the players.

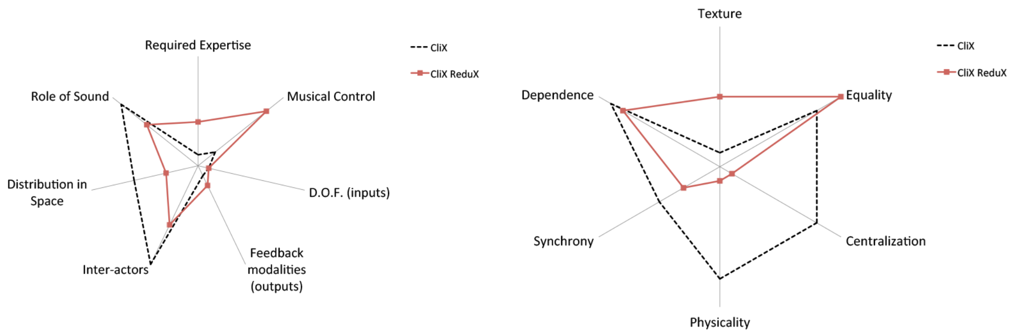
odic patterns of finite length (see Figure 4.7a). When sent to the server, the buffer of characters trigger short video sequences and their respective audio samples on a matrix, in which each performer has a specific cell assigned (see Figure 4.7b).

As a sequence-based environment, CliX ReduX allows for an off-line compositional stage, in this case by editing the string sequence. Once thrown no further intervention is possible, but a new sequence may be prepared to continue and further evolve the previous one or, as players usually did, in response to other players' contributions. The pre-buffering phases functions like a private space, however we did not provide a pre-listening functionality as in [Fencott and Bryan-Kinns, 2010], therefore its main advantage is to facilitate a sudden activation of new complex patterns instead of incrementally building them in realtime.

Performing with CliX ReduX involved a combination of predefined structures and reactive behavior to unexpected emergent patterns. As an example of the first process, players collectively agreed on keeping some keys with salient timbral or visual identities for later use and/or using them coordinately when appropriate. As for the second process, when a certain combination of patterns between two or more players was found to be interesting, the performance stabilized around it and other players joined by elaborating on them.

CliX ReduX is an example of an homogeneous, conducted instrument turned into a more heterogeneous, largely improvisational scenario. The multidimensional plots in 4.8 allows us to evaluate their differences: CliX is more scalable and greatly benefits from spatial arrangement of players, while CliX ReduX offers greater creative possibilities and feedback modalities, at the expense of a slightly increased difficulty (see Fig 4.8a). Players' identities are enforced by placing their visual contributions into specific locations in the projection screen and by triggering some video snippets with their own face as well, but otherwise the shared projection screen and access to the remaining snippets allow for a common ground for collective improvisation.

On the other side, CliX ReduX allows for better differentiated sonic roles if



(a) Dimension Space plots

(b) Collaborative Dimension Space plots

Figure 4.8: Multidimensional plots for Ge Wang's CliX and BLO's CliX ReduX

wished, whereas CliX is much more homogeneous timbrically. Finally, it must be noted that CliX ReduX is highly centralized and less scalable than CliX, in a way each piece was designed to suit the needs and take advantage of the characteristics of the ensemble it was addressed for.

Emergent Streams (2013)

While Clix ReduX, despite its centralized architecture, kept the traditional individual assignation of musical resources to performers, preventing from concurrent access to others' contributions, a shared environment for collective sequencing may be readily extended to support co-editing and thus sharing the authorship of the whole musical outcome[Bryan-Kinns, 2004].

This was precisely our approach with Emergent Streams, an interdependent, collective pattern editor conceived as a perceptual experiment on emergent multilinearity. We will discuss Streams in this section as it is a relevant example of environment for collective sequencing, but we will refer to it in the section on interdependent instruments as well.

Emergent Streams offers a single, shared loop consisting of a brief sound sample being repeated indefinitely at a rather high speed, on which sound effects are gradually applied. All performers access to the same loop, but each one of them may choose to have exclusive access to selected sections of the loop, and have exclusive access to the sound process to be applied to them as well. This process is schematized in Figure 4.9.

Compared to CliX ReduX, *Streams* trades multiple musical threads with exclusive access for a single thread with shared access. In this respect it is simpler

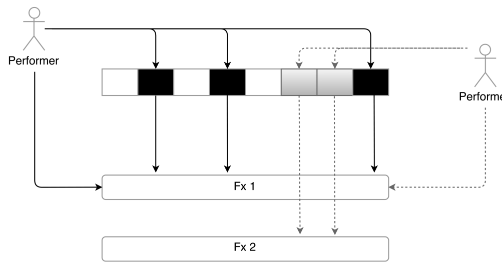


Figure 4.9: Sound and process ownership in *Streams*. Performer 1 owns the three black sections of the loop and Fx 1 while Performer 2 owns the two grey sections and Fx2.

than the aforementioned Daisyphone design, which allows to deal with multiple, shared concurrent threads, though as we will see it does provide time-interleaved multi-threading, the main focus of the instrument (see Figure 4.10 for a schematic of the player-pattern access policies in the three instruments).

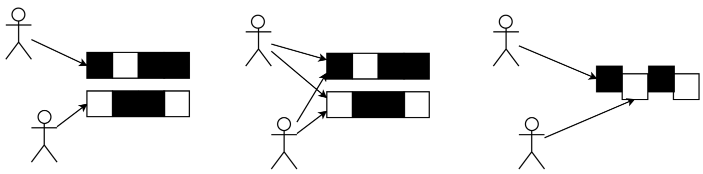


Figure 4.10: Schematics of player-pattern access policies in CliX Redux, Daisy-Phone and Streams

A simple PureData interface (see Figure 4.11) provides access to a sequence which is continuously looped, though in this case it is a single one shared by all players. A performer may choose up to a certain number of available time slots (the red ones are already taken by other performers), choosing an available sound effect and increasing the dry/wet mix each player may suggest a new stream, a group of events which will be perceived as a unity. It is, in this respect, both a single threaded instrument (a single process taking place) intended to be perceived as a multi-threaded one.

Sound effects ranged from very subtle timbrical modifications like partial pannings, small volume shifts and equalization, to more obvious ones like pitch changes, distortion and reverb.The conductor gradually removes constraints, allowing players to own more samples or to apply new, more drastic sound effects to them.

Figure 4.11: Streams. Graphical interface for performers.

After some rounds keeping a stream stable, a performer can free his events and look for others. Once free, the sound transformations applied are kept, which was a disturbing feature for players. This was the main point of the piece, however, the cognitive challenge associated to flexible appropriateness and continuous perceptual restructuring.

The performance is expected to proceed as an incremental accumulation of perceptual streams, which continuously allow for new reinterpretations of an originally static sequence of identical sound samples, by means of timbrical modifications applied to them. Such modifications should be as much subtle as possible at the beginning, hardly above the just noticeable difference for each timbrical parameter, so as to allow a new perceptual stream to gradually emerge. As the piece evolves, however, faster and more abrupt changes should be applied in order to stand out of the formerly applied effects and others' streams. Eventually, the overall complexity of the accumulated sound processes, the collision between competing streams and the quick succession of new patterns giving rise to a sort of cocktail party effect.

This interface exposes performers to a *competitive* scenario because of the politics of resource ownership. First, the acquisition of time slots may frustrate other players looking for the very same ones, and secondly, streams compete perceptually and players feel forced to increase the number of time slots owned or the amount of effect applied. What we learned from Emergent Streams was that competition, being more intrusive than collaboration, tends to focus user experience too much in the social side of the performance, at the expense of a more music centered experience. Performing could become enjoyable just as a social game, but for the same reason it will be frustrating as soon as social interactions interfered with our musical goals.

Although not entirely a failure, Emergent Streams warned us against exacerbating certain social interactions in network music. Being social is at the heart of the performative experience with multi-user instruments. When designing collective instruments in which performers are aware of others' actions and may negotiate strategies on their shared resources, either in a collaborative or in a competitive style, the mediated social experience may take over the musical experience. At the extreme, the activity could resemble a sonified network game instead of an environment for collective music making.

A corollary of these observations could be

Design strategies to enhance social interaction may positively contribute to creative group flow ...but the social experience should not come at the expense of music itself.

Six Pianos arrangement (2014)

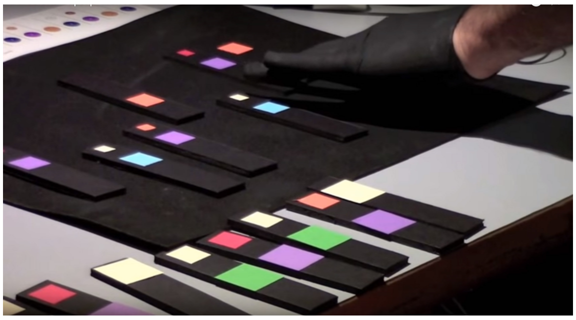
The two previous instruments were targeted mainly to experienced users. CliX adds an additional mapping layer from letters to video clips which requires prior knowledge of the corpus of video clips available for the performance, whereas Streams, despite its easiness of use, asks for a very subtle operation and interaction between performers. The last environment for collective sequencing we will discuss, Six Pianos, adopts a different approach: the design of a tangible interface for unobtrusive, robust collective pattern edition addressed equally to expert players and novices.

With this interface we wanted to address the lack of fluidity, immediacy and agency perceived in the former approaches, which suggested us to experiment with shorter loop-based schemes, which would provide almost real-time feedback to performers.

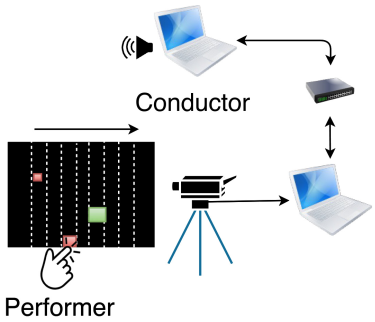
Six Pianos (1973), the piece which inspired our instrument, is a process-based, minimalist piece from Steve Reich. While similar in scope to other phase music pieces from the author, it focuses here on the accumulative buildup of patterns through addition of notes and volume fades. The basic process consists of some players keeping a steady pattern while some others gradually introduce the same pattern out of phase. The repetitive, phasing pattern relationships is just one 4/4 measure at 192bpm, so each loop lasts only 1,25sec. We though designing an interface to recreate Six Pianos would provide a suitable framework for further improvisation on collective note-based pattern sequencing which could overcome some of the drawbacks of Solitude.

The Six Pianos instrument is not really a multi-user instrument, but a set of six independently threaded musical interfaces which only at the implementation level share the sound synthesis engine. It consists of six sets of laptop and a computer vision system which tracks a tangible sequencer, made of a grid painted on a sheet of paper where players operate by putting and dragging colored tokens (see Figures 4.12a and 4.12b). Mapping pitches to colors was motivated by the evidence in the pedagogical literature that color-coded notation aids in the acquisition of performative skills, thanks to audiovisual synesthesia [Rogers, 1991].

The attack times are quantized to eighth notes as in the original piece, while



(a) Live performance



(b) Interface schematics

Figure 4.12: The Six Pianos. Playing the interface and schematics for a single player.

frequencies are restricted once again to the D major pitches required to play the original Steve Reich's Six Pianos. Finally, the player could choose three different octaves represented as different sizes, and the volume could be changed on a note basis in a sound console fashion, by dragging the squares upwards (volume up) or downwards (volume down).

Each station sets a network link with a server which acts as a conductor (setting a master clock and controlling the isochronous triggering of events) and plays all notes received with a sampler engine. By routing audio channels on the server we provided audio feedback to each player with individual monitors. We implemented the stations as an standalone OpenFrameworks application while the server runs a SuperCollider patch, and both exchange standard OSC messages.

We provided individual visual feedback to performers in the adjacent laptop screen, which overlays performance information to the captured image, such as a moving timeline which illuminates the notes being triggered. We provide a global collective feedback in a projection screen, both for the performers and the audience.

The fact that all performers shared the same timbre (a sampled piano) was both a key feature to recreate the subtle emergent patterns of the original piece and a challenge for proper collective awareness. It provided a strong sense of communal, shared experience, but performers had to learn not to overplay but instead appreciate the dynamic balance between masked and salient contributions. A tangible interface was an advantage here, as it increases the sense of appropriateness and embodiment, resulting in a more intimate experience and a feeling of being in control.

Because the most challenging factors involving the original Reich performance (rhythmic accuracy, control of dynamics and precise ensemble synchronization) are mostly dealt by means of software, our interface guarantees a rhythmically and modally good-sounding performance with little effort and might therefore resemble to be exclusively targeted to novices. It actually prevents performative errors, which to many authors are a key indicator of the potential for virtuosity of an interface.

As a consequence of an insufficient collective feedback, both visually and aurally, performers hardly improvised together at the beginning, resembling their activity more a building game. Once the patterns are well established, though, they quickly started improvising by building new patterns and gradually adding or removing markers on the time slots, much similar to the original Reich concept . Some players devised new unexpected strategies as well, like using their gloves to force sudden mutes, or moving the whole sheet and the markers above to cause a phase shift in the sequence or a global volume change.

Because of such simplicity, the interface allows performers to concentrate on the ensemble, complementing incipient gestures or filling gaps, gradually fading in and out of the mix while listening to the emergence of novel patterns.

The interface have some additional advantages. Playing with a tangible interface allows for a closer relationship with the audience, which we invited to come around us and participate actively in our performances. We also noticed that its affordances (putting and drawing colored squares) and mappings (xy directions, color and size) were easily grasped, and successfully used it in musical workshops for children, even with multiple users interacting on a single interface.

4.2.4 Gestural interfaces

He discussed in Chapter 3 the recurrent use of gestural interfaces in computer ensembles. As novel input devices, they may be used to extend the laptop as a meta-instrument in a similar way that the incorporation of additional sound production devices, or more often they may become the principal input interface to offer a more embodied experience, both to the performers and to the audience.

The Barcelona Laptop Orchestra has made limited use of gestural interfaces. Possibly the most explored ones have been gesture capture interfaces for open air interaction, such as the Kinect[6] and the Leap Motion controllers[7], which we used for collaborative concatenative sound synthesis and audio mosaicing. We will discuss the genesis and strategies adopted to perform with both instruments.

[6] `https://dev.windows.com/en-us/kinect/hardware`
[7] `https://www.leapmotion.com/`

Nuvolet (2011-2014)

The Nuvolet [Comajuncosas et al., 2011], originally developed for a musical work written by the catalan composer Ariadna Alsina for singer-reciters and laptop ensemble, was designed to let a number of performers (one to four) of the Barcelona Laptop Orchestra to explore a multidimensional representation of audio snippets by moving their hands in the space (see figure 4.13). Later, we designed Nuvolet 2.0 as a setup more appropriate for Laptop Ensemble performance, with a similar approach but with a more distributed setup.

Figure 4.13: Playing with Nuvolet

Concatenative Sound Synthesis is a process whereby audio is created by the concatenation of many small segments of audio, called units, from a source unit database, called a corpus. In this process, unlike in traditional granular synthesis methods, the grain selection is not arbitrary but rather determined by the characteristics of the audio itself. This data driven process [Schwarz and Others, 2004] may take a given audio input as a *target* from which a list of audio features called descriptors are derived. Source units from the corpus are then selected based on how well they match selected descriptors of the target. Typically, the multi-dimensional descriptor space is searched using a path search algorithm (e.g. Viterbi [Schwarz and Others, 2004]) or an adaptive local search algorithm (e.g. Zils [Zils and Pachet, 2001]). This process is called unit selection. The target specification is often derived from a piece of audio or from user navigation through the corpus of source units.

Diemo Schwartz has been exploring real time improvisation with CataRT by analyzing and segmenting live audio captured onstage from a musician[8]. Several authors investigate as well how to navigate the multidimensional descriptor space, for example plumage [Jacquemin et al., 2007], which uses a custom 3D interface to control CataRT. Compared to it, Nuvolet relies on direct mapping from the spatial dimensions to a three-dimensional sound space, thus achieving a touchless but direct manipulation of the virtual timbrical space

The use of concatenative synthesis, and particularly graphical navigation through sound corpuses, for computer ensembles was also addressd by Catork, a CataRT skin optimized for laptop orchestras developed by P Ruviaro. His work *Intellectual Improperty 0.6* (1999) for the SLORK[Ruviaro, 2011], explores concurrent navigation of individually constructed sound corpuses, with a conductor providing textual commands through the network. Compared to his environment, Nuvolet 2.0 provides performers a shared view of a single sound corpus and supports collective awareness. Conductor's commands may have a more unambiguous and gestural quality too. However, it is less scalable for larger ensembles.

Nuvolet was more targeted to an interactive installation or a theatrical performance, incorporating a wide range Kinect sensor, suitable for full skeleton tracking of several performers over a relatively large are. The physicality of the instrument is rather high, performers must negotiate the spaces to explore concurrently and may resort to verbal and non-verbal interaction cues.

In Nuvolet 2.0 we switched to Leap Motion controllers placed in front of player's laptops, which allow for precise hand and finger tracking on a much smaller area. In Nuvolet 2.0, therefore, we replaced a singe interface capturing all performers by individual interfaces, and instead of a large shared projection screen, each player used his own laptop screen for navigation. Physicality is more reduced both in spatial extent and in possibilities of interaction between users, which are now mostly network mediated. All the performers share the same visualization of the corpus to guarantee consistency, and corpuses are controlled by the conductor's laptop which acts as a server. Performers only need to navigate with their hands, and their laptops only send the gestural data to the server, which provides individual audio feedback through a multichannel soundcard, connected to individual speakers placed next to each performer.

Both setups are highly centralized: all the sound processing takes place on a server computer, which runs a PureData, CataRT based concatenative synthesizer using William Brent's TimbreID external[9]. However, in Nuvolet 2.0 each performer operates on his own laptop to interact with the sound corpus, whereas in Nuvolet the interaction takes place on a shared space onstage. We may compare Nuvolet 2.0 to Catork as well, because they share similitudes: from the perspective

[8] `http://www.youtube.com/theconcatenator`

[9] `http://williambrent.conflations.com/pages/research.html`

of interdependence, both are homogeneous and hierarchical systems, being Nuvolet 2.0 centralized and more similar to a shared sound environment and Catork decentralized and closer to a set of loosely coordinated independent instruments.

From now on we will discuss our work with Nuvolet 2.0, as it is more specifically targeted toward computer ensemble performance. Several strategies were investigated regarding the mapping of sound corpuses, their assignation to players and strategies for performance.

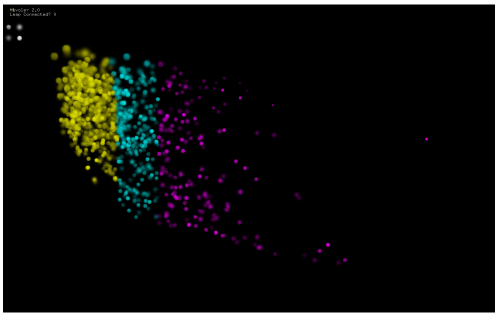
Descriptor space navigation We first addressed several strategies for collective navigation through a sound corpus. The corpus explored consisted of 1-minute passages of Berio's Sequenzas for three distinct instruments: female voice, cello and trombone respectively, therefore it provided distinct instrumental identities, with strong timbical variety and occasional overlap. Their individuality is expected to be useful for source-based clustering, while their overlap (ie they are not disparate timbrical sources, but originate from musical instruments with some shared range of pitches and timbres) may help to make more coherent timbre-based clustering.

Player assignation may be carried out by clustering the whole corpus and assigning each cluster to a different player. We may employ supervised clustering, by assigning manually sound snippets to player according to a certain attribute, for example the source instrument or its temporal location in the original audio source. Alternatively, we may employ unsupervised clustering techniques to automatically assign snippets to players according to a weighted set of timbrical descriptors. The four figures 4.14 provide four distinct strategies which will will now discuss.

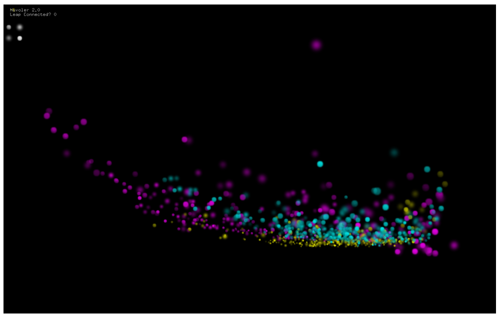
Different clustering criteria had profound implications in the group performance. If clustering criteria matches the visualization axes (see figure 4.14a) the clusters are graphically isolated and each player has a well defined, non overlapping space for performance. Performers found that this mapping didn't promote interaction between performers and was perceived as an enforced isolation. Besides, the assigned space tends to be too reduced and much of the display area (and its corresponding spatial volume available for performance) becomes unusable.

It is preferable to provide mappings which spread the available snippets among the whole display area, allowing for overlapping between players' assignations. By mapping snippets to axis and performers according to different criteria we will therefore achieve more satisfactory results.

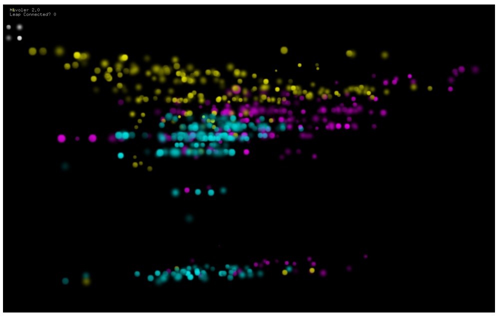
Figures 4.14b and 4.14c show two more different distributions, generated by supervised clustering. In both cases the snippets where sufficiently large as to be perceived as musical notes. The first one was generated with a supervised cluster in which each performer was assigned a different instrument, but snippets were



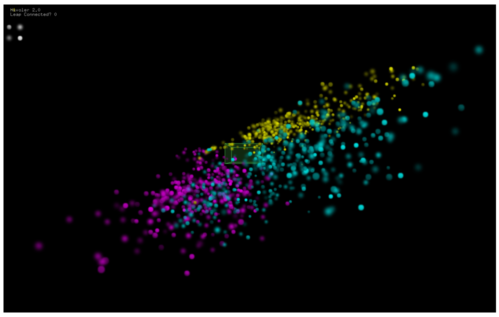
(a) Timbre-based mapping and user assignation



(b) Timbre-based mapping and instrument-based user assignation



(c) Pitch-based mapping and instrument-based user assignation



(d) Conducted navigation. The green cube sets the available space for performance

Figure 4.14: Four different mappings for collective exploration of sound corpuses. Each color corresponds to sound snippets assigned to a different user

mapped according to their spectral centroid and mean rms amplitude. Therefore ay player seeking a given combination of volume and timbre will explore a similar area, and expressive gestures may be readily understood and imitated between performers. The second, more obvious mapping, distributed once again sound snippets according to source instruments. This strategy encouraged performers to experiment with harmonies and simple imitative counterpoints.

In the first rehearsals we noticed that players continuously performed wide gestures covering most of the available space. The overall experience was that of a monotonous, dull, limited and frustrating environment. To address it, we explored different possibilities: incorporating more timbrical variety in the sound corpus, experimenting with different descriptor mappings, or adding constraints. The constraint consisted of a rectangular region that the conductor could dynamically shape to frame the available space for performance (see Figure 4.14d). Players were allowed to freely play only within the region.

To our surprise the last strategy proved to be the most successful. The performance turned out to be more varied and structured, even displaying a wider set of distinct gestures. Players consistently reported a higher level of engagement with the new constraint, as long as it was kept steady enough as to allow them to explore the possibilities of the region. Once the conductor started to guide the improvisation by moving and resizing the playing region more ostensibly, however, performative flow was tampered, enforcing players to adopt a passive tracking behavior.

As a corollary, we might summarize the former observations stating that *predictable constrains challenge and stimulate creativity*.

Interactive mosaicing A different strategy, already suggested but not fully explored in my previous work with the Nuvolet [Comajuncosas et al., 2011], consisted on a mapping suitable for concatenative data-driven synthesis (see [Schwarz, 2003]). This synthesis technique consists of an automated selection of sound snippets according to their similarity to a target sound source, and may be considered a sort of content-based, descriptor-driven granular synthesis technique.

The selection of the sound corpus available for exploration greatly differs from the former approach. We establish a path (in our case, a straight line along the X axis) which will correspond to the maximum similitude to the target source. Then, the corpus is built in slices on the ZY plane, perpendicular to the path, by assigning a number of sound snippets which will be mapped according to their distance to the target (see figure 4.15). The whole corpus displayed allows to gesturally control the process of resynthesis to the desired degree of accuracy, and similar clustering strategies may be employed to distribute the sound snippets among the performers.

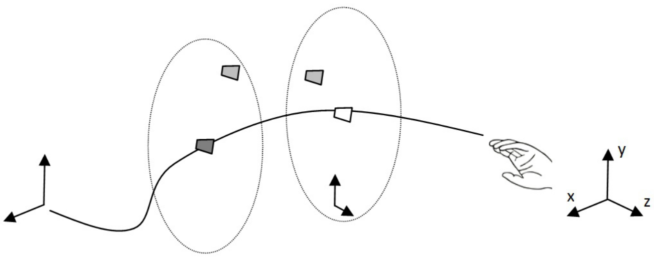
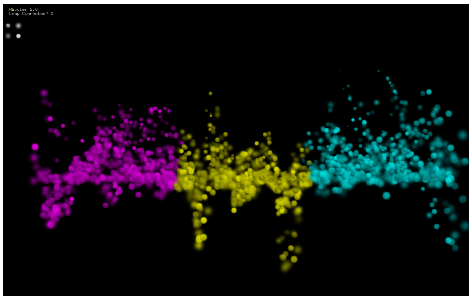


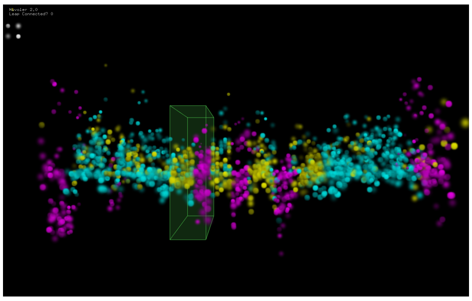
Figure 4.15: Mapping of the sound corpus for gestural target-based mosaicing

We experimented, for example, with the sample of a spoken voice because it provided a clearly identifiable target. Specifically the recording consisted of a palindromic sentence by Jesús Lladó[10]. Analysis and synthesis employed a large

[10] `http://palscat.blogspot.com.es/`



(a) Target-based mosaicing with time-based snippet assignment.



(b) Conducted, target-based mosaicing with timbre-based snippet assignment.

Figure 4.16: Two different mappings for target-based mosaicing.

window size of 8102 samples to allow for easy recognition of whole phonemes. The corpus consisted of the original soundfile, additional recordings of the same sentence by the same speaker, with different intonations, and to add more vocal variety and surprising utterances, the first 2 minutes of the first movement of the Ur Sonate rendition by Blonk-Jaap[11].

To build the final sound corpus the original sound file was segmented in windows of 8102 points and scrambled and then for each target window, the best 10 candidates for each 10% of the corpus file plus an additional randomly selected corpus window were selected. The purpose of the additional random snippets, wich are expected to be displayed isolated, was to incorporate some unexpected variety in the corpus. The axes represent normalized time(X),normalized centroid increment (Y) and normalized amplitude increment (X) (see figure 4.16a and 4.16b).

Once again, the mapping and constrains significantly influenced the performative experience. The semantics of the original sentence transformed the improvisation into a goal driven, conducted performance. As before, each performer had only some areas assigned, for example, according to the time offset of the target sound file, as in figure 4.16a or, by extracting the MFCC coefficients of each slice, to the automatically clustered target phonemes as in figure 4.16b). Collective reconstruction of the sentence was best achieved by negotiating a player acting as a leader with all the performers tracking him. The result was a collective reconstruction of the original sentence with different sections emanating from different performer's speakers.

If the path followed by the leader was predictable enough or if, as before, a conductor set up a visual constrain (see figure 4.16b), performers could combine a goal-driven performance with the additional freedom to depart or converge to

[11] `http://www.jaapblonk.com/Pages/ursonate.html`

the target. This lead to an additional social experience: an emerging collective dynamics which explored the variable intelligibility of the reconstruction process as an expressive resource. This final scenario turned out to be the most satisfactory for the performers both in terms of user engagement and musical outcome.

We could conclude stating that *a loose conducting scheme combined with additional degrees of freedom let performers to establish parallel, complementary creative dynamics*.

4.2.5 Interdependent networks

La Roda (2010)

La Roda is a collective instrument which the Barcelona Laptop Orchestra has been performing with for the last four years. It was conceived as a musical analogy to the popular Chinese Whispers children's game. La Roda consists of a turn-based iterative sound processor, in which performers gradually mutate a sound snipped with their custom made effects. This configuration proved to be rather flexible, being played from electroacoustic venues to more mainstream electronic music events.

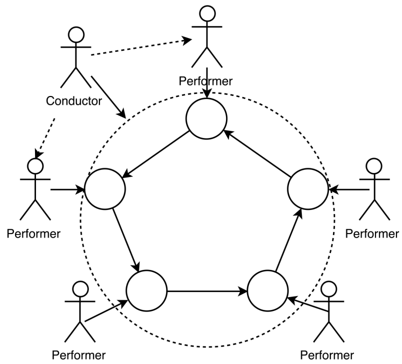
The basic schematic of La Roda is illustrated in Figure 4.17a. Each performer has a time slot allocated to manipulate a sample, which is then sent to the next performer in a circular fashion. A software conductor controls the turn mechanisms while a human conductor guides performers.

The original analog architecture of La Roda consisted of a set of laptops with their audio input and output daisy chained. All of them kept a buffer which was filled while the previous player was playing and then played back with the sound processes of choice by the performer. A central server governed the recording and playing turns.

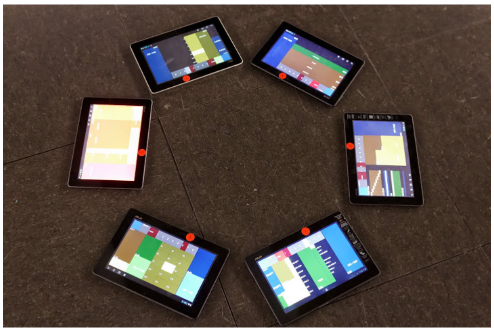
This approach proved to be problematic because of the progressive degradation of the buffer caused by the accumulative effect of successive AD/DA conversions and the severe desynchronization caused by mismatched audio latencies.

In a further iteration, La Roda was implemented as a centralized DSP which was in charge of the whole audio process and which incorporated the sound processing part of each performer, while players laptops or mobile devices kept just a graphical control interfaces to directly access their own effect in the server (see Figure 4.17b). The server was implemented in PureData and each processor was a PureData abstraction which every performer was in care of.

In either configuration, strong individuality is kept with very idiosyncratic processes tied to single performers, much in the style of The League. Players were encouraged to develop their own ideas and, after testing then within the network, usually adjusted and balanced the effects, often incorporating new functionality



(a) *La Roda*. Schematic diagram.



(b) *La Roda*. A setup with six table-based clients ready for performance

Figure 4.17: *La Roda*. Schematics and an illustrative setup

which could match others' sound processors. Another player was in charge of the centralized DSP and turn controller, effectively acting as a conductor or supervisor of the whole system.

The high interdependence between them comes from the fact that instruments are linked in a loop configuration, with a single shared sound stream flowing from one player to the next. Indeed, the accumulative behavior and heterogeneity of processors involved in the instrument lead irreversibly to unexpected timbrical evolution.

By allocating specific processes in unshared turns, a high sense of collective agency is achieved, though. La Roda performances are mainly gestural, each performer exerts a certain, rather unambiguous, timbrical gesture to the sound stream, which is permanently imprinted on it. As such, it is decoupled from the performer just at the very next turn, though we observed that a certain amount of turns were required to actually incorporate each gesture to the morphology of the audio sample. This perceptual effect prevented performers from over-processing: la Roda works best when players leave time for gestures to be assimilated and try to operate on different temporal sections of the sample and/or on different timbrical dimensions to build on former ideas and avoid overlapping at the same time.

La Roda challenges the expected rituals of a public performance, because of its inherent unpredictability, mainly dues to its collective feedback loop structure. Novel patterns often emerged as an unexpected combination of the timbrical features being continuously incorporated by every sound processor.

It took painful rehearsals for the players to learn to cope with it, by knowing intimately not only their sound processors but the others' as well. In a way, we felt that the instrument itself was continuously impelling us to seek new perspectives in the sample, new anchor points and new timbrical gestures which served as a basis for further developments. In a live performance players must learn to quickly react to others' ideations, either by compensating for an unbalanced processing, or by giving a new interpretation to a former timbrical gesture. It thus becomes a mixture of free sound exploration and collective improvisation.

The most idiomatic context for a performance with La Roda is a free, or loosely guided improvisation. However, no matter the level of proficiency of the performers and the subtleties of indirect conduction employed, its irreversibility and accumulative character may easily lead to an unplanned timbrical evolution, requiring great coordination to achieve a unified discourse. Otherwise, the emergent sonic gestures resulting from the collectively assemblage of timbrical gestures may be more reminiscent of a surrealistic *Cadavre exquis* game.

La Roda may eventually reach states in which performers feel frustrated and uninspired, because they loose the possibility to exert any kind of sensible modification to the audio sample. The typical scenarios of extremely soft or disgregated sounds, or wide-band noisy passages, place a creative barrier to the performers.

Moreover, La Roda may eventually converges to certain states, mainly determined by the nature and ordering of the effects employed. Though a huge timbrical variety and sound gestures can be achieved, ultimately most samples employed as a seed for the performance will evolve to paths that performers recognize. While this usually happens only when setting effect parameters to extreme values and there is no sensible coordination of everyones' contributions, the suspicion of irrelevance leads to a vicious cycle which strongly inhibits creative engagement.

From such observations we could derive the following corollary

Serendipity in shared environments encourages creative practice ...but decreased collective agency invariably leads to frustration and uncreative behavior

This is an example of performative context where a certain level of supervision may benefit the global creative outcome. Eventually La Roda incorporated a number of resources available for the conductor. Those resources go beyond the automatic turn control and provide mechanisms to guarantee a more sustained

Table 4.2: Summary of conductor's resources for La Roda.

task	**responsibility**	**functionality**
declicking envelope	automated	suppresses clicks when changing turns
post-gain balance		compensates for sudden, excessive gain changes
turn time	automated, supervised	time between turns
turn ordering		changes global behavior, allows for solos, duos...
global mute	manually operated	cleans the buffer may add rhythmic patterns based on silences
global audio gain		imparts additional gain, or fade out at the end
injection		adds new audio elements
typed commands	condutor's cues	global or user-specific commands
gestural commands		global commands mostly for feedback, appraisal, dynamics and character
section reminders		prearranged behaviors, verbally or gesturally

and gradual evolution or a more structured performance. A sample of some of the available resources is provided in Table 4.2.

The Hub (2012)

We surveyed in Chapter 3 different approaches in the repertoire for Laptop Orchestras based on interdependent networks, and classified them according to the musical resource shared: parametric control, musical processes or audio objects. In La Roda it was an audio object which was shared among performers. In our next instrument, The Hub, we addressed the other two resources at the same time: musical control and processes.

The Hub is a multi-user instrument conceived as a shared workspace for collective algorithmics. Players collectively build generative processes by patching algorithmic modules onscreen. It shared the principles of immediate feedback from other interactive programming practices, such as live-coding, and the paradigms of modular, dataflow interactive programming of environments for computer music programming such as Max or PureData.

At implementation level, The Hub consists of a PureData based central server which handles the processing of nodes and the routing of control signals, eventually sending them to software based sound processors or to external MIDI synthesizers. The nodes are themselves simple PureData abstractions which may be dynamically instantiated. When more than one link reaches the same node input, all control signals are just injected as they arrive without any mixing process. For nodes forming cycles, a specific feedback detector prevents data overflow.

The players run a custom OpenFramworks GUI on mobile devices (usually tablets) which is used to add and remove new nodes and links between them by sending OSC commands to the server. To focus on the principles of network interactivity, the nodes don't exhibit any parameters for real-time control, they can only be modified through control data coming to their inlets (see Figure 4.18).

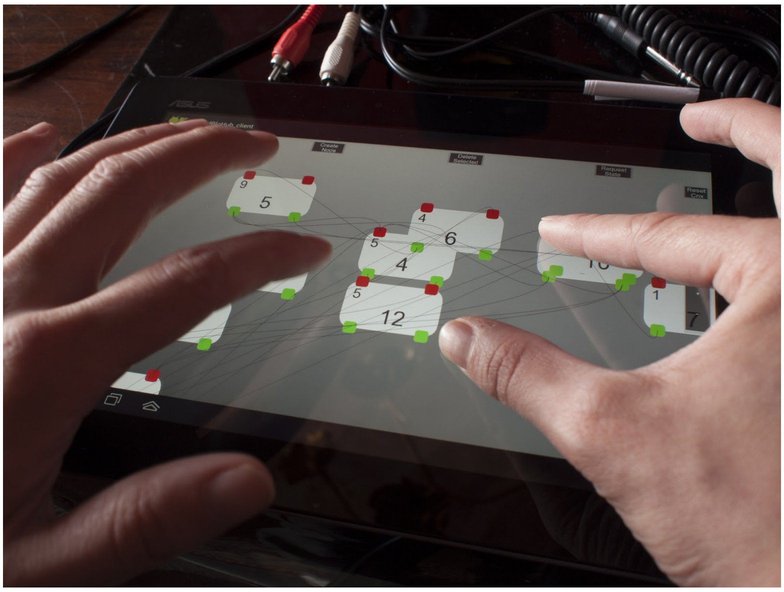


Figure 4.18: Screenshot of the clients' GUI for the HUB

An additional, embellished version of the clients' graphical interface was projected onstage for feedback for the ensemble and the audience. Most notably, it showed the activity of nodes and the signals being carried by the links.

The Hub was designed to allow for arbitrary sets of nodes to be dynamically incorporated, though the ensemble preferred to restrict themselves to a smaller and stable set in order to learn to combine them more purposefully.

A performance with The Hub was originally conceived as a structured improvisation. Typically, all performers were wandering around the stage holding their tablets but keeping visual connection with the other players and the conductor. Each player was in charge of a certain role: adding generators, shaping the activity by adding or removing certain connections, choosing what signals to send to what synthesizer controls or adding some previously proven self-autonomous subgraphs. The conductor shouted commands aloud throughout the performance, at the beginning they might consist of very precise instructions as to guarantee a self-sustained generative patch, but eventually the commands were more generic, giving freedom to performers to socially interact through their preestablished roles.

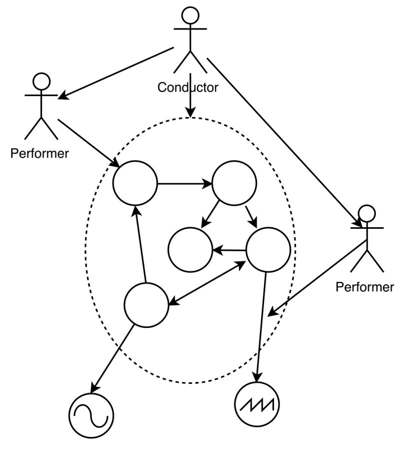
The main musical interest in The Hub lies in the unexpected behavior of complex networks of nodes incorporating feedback and unusual links between them. The most inspiring networks exhibited a self-sustained behavior which was at the same time varied and non repetitive but exhibiting a definite character. Those networks were saved as presets during rehearsal and automatically regenerated in a performance, thus performers had a complex functional network to work with.

The original topology of the Hub was essentially centralized: all the performers were concurrently operating on a single shared workspace (see Figure 4.19a). On a subsequent iteration (see Figure 4.19b), we provided a specific proxy object (P nodes in the figure) which allowed performers to communicate with the central patch from their computers through OSC, running their own processing algorithms (I nodes). Sound generators could be linked to the central patch or, more frequently, to performer's own processing algorithms. Performers can therefore program more sophisticated instruments and plan for a number of typical scenarios in advance, and may devote the performance to a high-level control of the algorithmic processes involved.

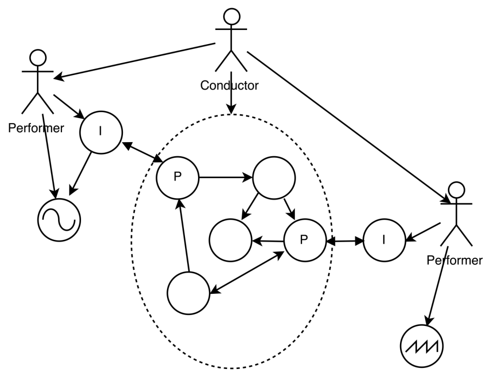
Figure 4.20 displays a capture of the conductor's control screen for a rather complex patch operating according to this second topology. The conductor screen shares the same features as the performers' control screens with some added capabilities for global management of the network topology. This control screen is actually a graphical front-end to the PureData patch which manages the HUB network and the signal processing taking place in the nodes.

On the left we may see the available objects, on its right the active network of chosen nodes and their interconnections is displayed. Each node provides two inlets and two outlets, and works as sort of oscilloscope as well, to easily monitor the flow of data through them. The wider object labeled OUTPUT in the lower part of the main display is used to route any control signal coming out from the network to the centralized synthesizers' parameters (if any).

Finally, the conductor can trigger automatic generative processes to quickly mutate the network, such as adding and removing nodes according to some hard-



(a) Centralized topology



(b) Semi-decentralized topology

Figure 4.19: *The Hub*. Schematic diagrams for a centralized and a semi-decentralized topology

coded probabilistic rules (Auto Node button), and Wiring and unwiring them (Auto Wire button) or from them to the OUTPUT synthesizer's inputs (A/W Synth button). The conductor can also load a previously stored network (with the Choose Preset button).

The nodes with a picture of the performer inside are the proxy objects which link the HUB with the local processing algorithms that performers implemented..

Compared to the former topology this semi-decentralized scenario combines two performative paradigms, a unified shared programming environment and a network of interconnected, heterogeneous processes.

The role of the performers consisted of incorporating their generative processes and gestural inputs to The Hub, while receiving and freely interpreting incoming data from it. The role of the conductor consisted of managing the inter-

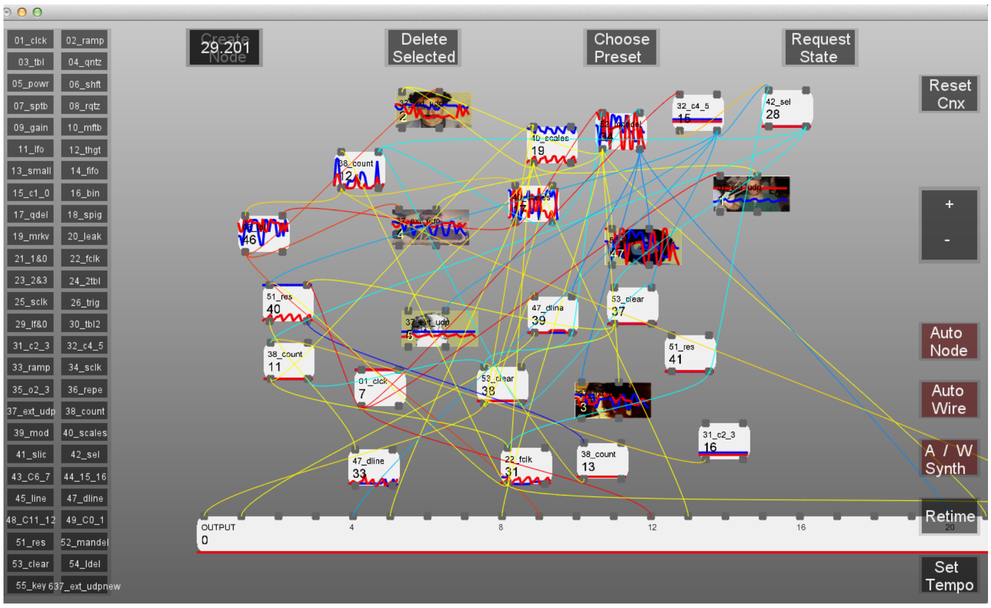


Figure 4.20: Screenshot of the conductor's GUI for the HUB

connection between performers (up to this point, the setup would be equivalent to one of the performances from the original The Hub band, though with a dynamically reconfigurable network) and eventually incorporating additional generators and processors which would affect, and be affected by the player's.

Despite being more decentralized, this is a more hierarchical scenario, because the conductor has a higher control of the flow of data between all the processing nodes. For example, the conductor may provide a shared stream of data to all performers, for a highly coherent performance, he may interconnect them all to promote emergent behaviors, finally performers' nodes may be part of a more complex global patch and thus a higher degree of interdependence and complexity will be achieved.

One of the most interesting features that this scenario offered was the ability for the performers to inject gestural data to the shared patch, whose signature was easily distinguishable from the more quantized streams generated by the HUB network. It gave a sense of organicity and embodied involvement to the performance, and allowed performers to easily highlight their own contributions with distinctive gestures.

The Conducted Roda (2014)

Going back to La Roda, this instrument was further re-worked in the context of a study on conduction of collective instruments carried out by this author ([Comajuncosas and Guaus, 2014]).

We already discussed that both the traditional human conduction commands given in a performance and the software driven procedures to control the turns did not provide the level of coherence and control which we desired. We address in this section the development of a software module which was conceived to assist both the conductor and the performers in order to be able to perform in a more coordinated way.

A software assistant for La Roda consist of a set of modules which follow the structure of a closed-loop controller (see Figure 4.21):

- The conductor sets a **target**, consisting of a desired set of audio characteristics which the sound sample should achieve.
- At each turn one performer modifies the sound sample with his effect $F_x(n)$
- The **Analysis** module analyzes relevant audio characteristics from the processed sound sample.
- A **Model** module lets estimate the optimal ordering and effect configurations for the next processors to guarantee a proper convergence to the desired timbrical target.
- A **Planner** module schedules the estimated effects and their configuration and sends them to the players.

Our first approach consisted of building a predictive model for each performer's process, effectively turning them into centrally controlled adaptive effects. The processes were heterogeneous and provided two parameters each, and the procedure comprised a training phase consisting of an analysis of the effects' behavior for a large combination of varied sound sources and parameter combinations. Once this offline stage was finished, La Roda would infer in realtime which combination of effects and parameters better matched the desired target descriptors (for further details on the processes involved, please see [Comajuncosas and Guaus, 2014]). By injecting the parameters into the system, the whole instrument could be run automatically or provide suggestions to performers.

While the results were promising, the heterogeneity and non-linearity of the processes selected often prevented the system from achieving the desired convergence. Our next approach consisted of a simpler, homogeneous system in which

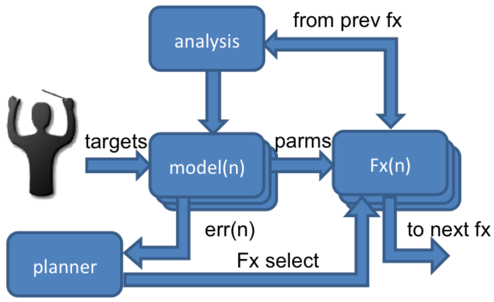


Figure 4.21: The *Conducted Roda*. Functional block diagram

performers explore iterative equalization through a series of band-pass filters distributed among them.

This conducted Roda is a sort of distributed, asynchronous vocoder which applies the processing in turns. At the beginning of every turn the software conductor analyzes the buffer which acts as the carrier signal in a vocoder. The analysis extracts 12 bands in 200-mel spacing, covering from DC to 6.5kHz approximately, and representing the mean RMS amplitude for each band and for the whole turn.

The modulator is the target spectral envelope provided by the conductor, who freely draws it in a custom interface. Such envelope represents the desired RMS amplitude for each band. A planner module computes the two largest differences between the carrier and the modulator and suggest the appropriate gains to be applied.

Performers control four bands each, consisting of peaking/notch filters with a configurable gain of $\pm 24dB$ (to allow for a very dramatic equalization if desired). As this procedure is repeated at every turn, eventually all players will get suggestions for some of their bands, and most of them will be received with enough turns in advance as to allow them to accurately plan the requested slider movements.

If the target envelope is kept steady and performers accurately follow the suggestions, the buffer will eventually converge to the desired mean equalization, and the suggested movements will approach zero.

The whole process for one turn is depicted in Figure 4.22: performer 1 operates on one of his bands (band 10 in this case), the Model compares the spectral envelope of the processed sound sample at the end of the turn against the desired equalization curve set as the target by the conductor. Then, the Model suggest Performer 2 to operate on his band 11 to further approach the desired target.

The graphical control interface for this second Conducted Roda is displayed in Figure 4.23). On the top, the performer can monitor the turns (the green an

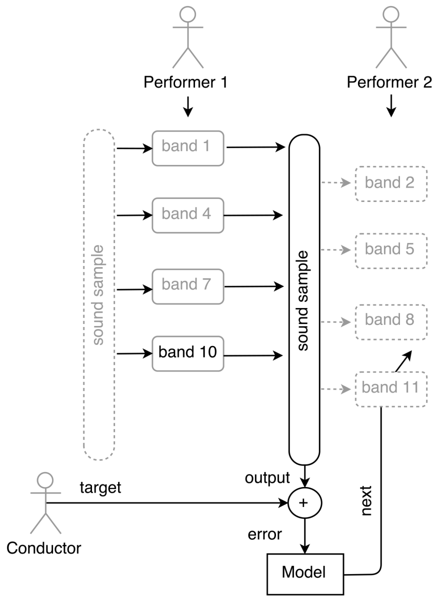


Figure 4.22: Steps performed in one turn of the Conducted Roda

red colors indicate that the player 2 is currently active and the player 3 should be prepared to play in the next turn) and the available turn time (the slider on the right). The central graph displays the current mean Eq curve (in solid green) and the desired target equalization (as a superimposed yellowish line). The slider on the right shows the correlation between the two, that is, the agreement between the actual and the desired equalization. Finally, the lower area shows the sliders available to the performers. Each performer has four slider bands to play with (the green one) but can see what others do in their own bands (the gray ones). Additionally, the optimal slider value to achieve the fastest convergence is indicated as a small yellow ball on the slider.

There are notable differences between the design of La Roda and this new Conducted Roda. It is a homogeneous, top-down designed instrument in which all performers operate on the same musical dimension. A bare minimum individuality is kept by assigning frequency bands to performers, but they are interleaved in the frequency domain to encourage all players to participate most of the time and to potentially affect each other. Because of this interleaving and the relatively low Q of the filters employed, any target equalization will require the collaboration of several if not all performers.

In our scenario, commands are unambiguous, addressing precise indications to everyone, but performer's might fail to follow them exactly, or might just decide to do *something else*. We might consider those eventual imprecise actions or severe

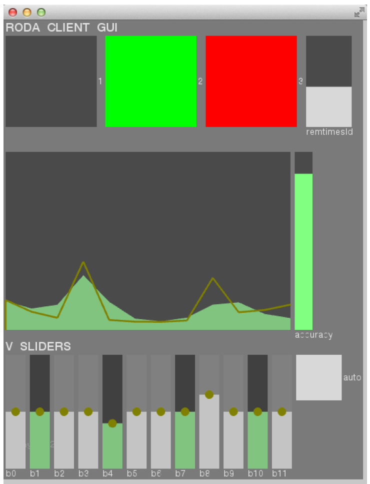


Figure 4.23: The client GUI for the Conducted Roda

departures from the planned activity as *ornaments* (from the perspective of the performers) or as *errors* (from the perspective of the software conductor).

The system will not halt the performance but, instead, it will plan a new strategy to achieve the same stated goals. Such actions will generally slow down, but not completely defeat the convergence to the desired state.

The challenges posed by performers, therefore, ask for a creative reevaluation of the strategy (by the software-assisted conductor) and possibly of the goal itself (in case the conductor just decides to assume the new path and suggest another goal in response to the ensemble dynamics). Finally, this partial departure of an otherwise rather predictable process might actually add valuable interest and novelty to the performance.

That the system can successfully react to unexpected behaviors does not mean that they should be looked for, of course, but it proved to be one of the main engaging features of the new Roda.

To illustrate the impact of a software-based conductor we provide four performance logs with the new Conducted Roda. Figure 4.24 displays the evolution of the correlation between the conductor's targets and the ensemble outcomes in a sequence of turns for four different scenarios (the conductor proceeded with the same sequence of targets, which were provided as soon as the previous was reasonably achieved).

- **Scenario 1**. Non guided goal oriented performance. The interface showed the desired target Eq's but performers had to collaboratively achieve it.

- **Scenario 2**. Guided, goal oriented performance. Same as Scenario 1, but now the software-assisted conductor instructed each player what to do at each turn.

- **Scenario 3**. Ornamented guided goal oriented performance. Similar to Scenario 2, but performers were encouraged to disrupt the expected evolution by adding some unplanned modifications to the current equalization.

- **Scenario 4**. Automatic performance. In this scenario the software-assisted conductor took complete command of the performance.

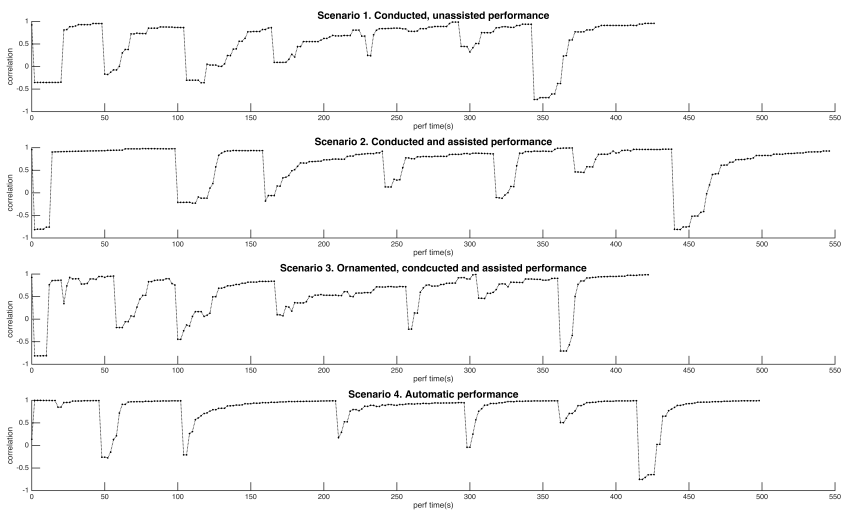


Figure 4.24: Four conducted performance logs with the Conducted Roda

By comparing the first two performances, we see that computer assistance effectively allows performers to start converging faster and can keep achieving better and better scores compared to an unassisted performance (incidentally, the first score is shorter because new targets were set as soon as no further convergence pas observed). The third performance allowed performers to include *ornaments* (or otherwise, performative errors) which then they should cope with, with software assistance. While some ornaments eventually defeated an optimal convergence, the system nonetheless kept on guiding performers successfully. To the ensemble this was the best scenario, both in musical terms - their ornaments provided plenty of opportunities for novel and surprising directions to be taken- and socially, as everybody was more involved by actively re-planning their participation and a successful dialogue was established between the conductor and the ensemble. The best *theoretical* scenario would be the one depicted in the last plot, which corresponds to an automatic performance without performers: compared to the second plot, the convergence is much faster and smoother, and an optimal correlation is achieved for all conductors' targets. It, however, lacked the excitement and sense of complex interaction conveyed by the third performance.

We see how, as performers may only contribute by modifying a single dimension of the sound, and that dimension is the only one the conductor may resort to during the performance, a certain equilibrium between performative freedom and adherence to stated goals must be established to avoid an undesired lack of cohesiveness. In rehearsals, a meta-structure of the conducted improvisation was usually planned and discussed beforehand, deciding the balance between ensemble precision and ornamental departures on different passages.

Soundpainters, for example, don't take such unexpected disagreements between conductor's commands and performer's actions as *errors* as we might call them. Either a command may be ambiguous or a performer may misinterpret or plainly ignore it, but this unpredictability is usually considered an opportunity to creatively reevaluate the flow of the performance. Under such perspective, imprecision and unexpectednesses are not to be seen as a *necessary evil* that the conductor must learn to cope with, but rather a result of a bidirectional creative flow between the conductor and the ensemble. A conductor in this context is not an absolute reference, just a guide which may be challenged by the ensemble as well, and this is the precise role of the conductor in the new Roda.

4.3 Summary and discussion

To summarize this introductory overview of our repertoire, we will show two multidimensional plots which show the whole range of possibilities explored.

The first plot (see Figure 4.25) analyzes our repertoire according to the axes proposed in Birnbaun's Dimension Space for Musical Devices [Birnbaum et al., 2005]. It shows a shift towards the top/right zone, as corresponds to designs addressed to (expert) musical performance: they are targeted to medium sized ensembles, with an strong emphasis on sound and musical control which require generally a sensible degree of expertise, and the use of the space as a resource in performances, and the amount of additional feedback modalities being significant but comparatively more reduced than in designs focusing on non-expert participation, as would be the case with interactive installations or setups involving the audience.

Finally, Hattwick and Wanderley summarize most (but not all) of the former classification criteria for multi-user instruments in their Dimension Space for evaluation of Collaborative Musical interfaces [Hattwick and Wanderley, 2012]. While it is a convenient, compact way of characterizing a multi-user musical interface, we should emphasize that each axis actually may refer to several closely related criteria, so further discussion will be needed for a precise description of

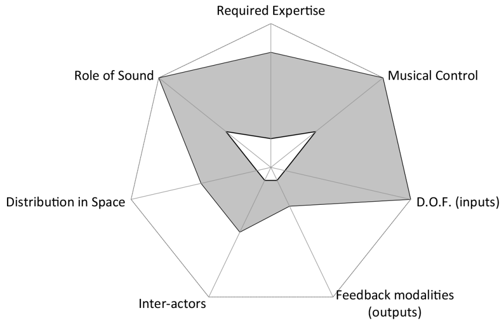


Figure 4.25: Dimension Space plot for the musical devices analyzed in the BLO repertoire survey. Total ranges.

each instrument.

The plot shows a rather exhaustive exploration of most of the dimensions, with one precise combination yet to be explored: highly decentralized and heterogeneous systems (see Figure 4.26). We believe that this fact is mostly motivated by the incorporation of a conductor and the recurrent use of networking governance techniques which motivated the adoption of uniform protocols for communication. We should mention as well the adoption of similar, if not the same programming frameworks for all members and the provision of templates for instrument development even in our more decentralized repertoire.

A similar trend could be observed by comparing the first, extremely heterogeneous and decentralized setups of The League and their late developments in the Hub, with their MIDI based setup acting as a centralized routing system and unified communications protocol.

Because of the particular relevance of interdependent multi-user instruments in this research, we will provide a more focused evolutive analysis of the relevant designs just reviewed, namely La Roda and The Hub, with their multiple iterations, and the collective sequencer Streams.

The analysis is summarized in four figures (see 4.27), which together summarize the most relevant taxonomies for multi-user instrument surveyed in the former chapter.

In terms of roles, both the Conducted Roda (because the conductor agent enforces the balance) and the Hub provide balanced environments, with equal responsibilities and opportunities. La Roda and Streams, on the other hand, may promote undesired unbalanced behavior because performers may freely compete for a shared resource. In terms of role flexibility, La Roda in his two versions was a rather inflexible instrument (roles were rigidly determined by the chosen

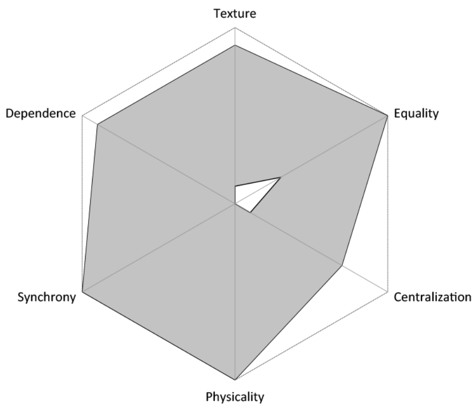


Figure 4.26: Collaborative Dimension Space plot for the musical devices analyzed in the BLO repertoire survey. Total ranges.

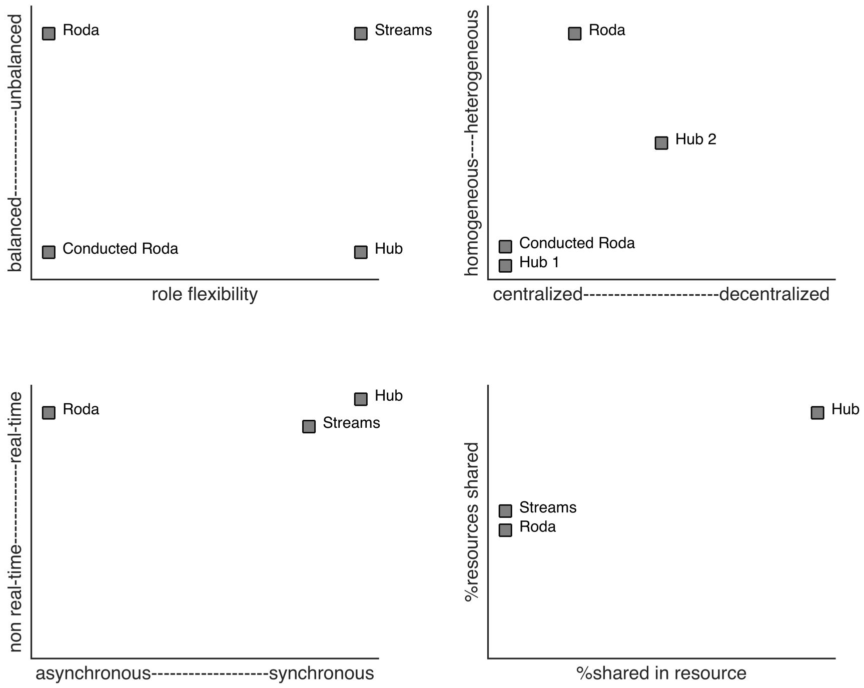


Figure 4.27: Characterization of the interdependent multi-user instruments analyzed in the BLO repertoire survey

instruments) while in the Hub performers may freely construct their algorithms or decide how to interact with them, while in Streams they may freely choose the sound elements and the processes to apply to them. We might say that the dimension accounting for role balance and flexibility has been fully explored in our repertoire.

In terms of temporal interaction, our environments are all conceived for real-time interaction, but La Roda is an asynchronous, turn-based instrument, Streams is a quasi-realtime (and quasi-synchronous) collective sequencer while The Hub is perfectly synchronous. Those are the two contexts expected for environments tailored to real-time performance, while non-realtime environments would be better suited for collective composition, either in concurrent (ie synchronous) or in iterative (asynchronous) scenarios of collaboration.

Regarding the design of the instruments in terms of degree of centralization and homogeneity, our approaches were multiple. Our most purposefully decentralized design was La Roda, even if it was eventually turned into a centralized design. The Hub was more centralized from the beginning but the incorporation of the proxy objects open the possibility for a more decentralized approach, and finally our tow Conducted Roda versions where highly centralized by the requirements of the conductor agent. Heterogeneity is more characteristic in designs where performers may implement their own processes, as in La Roda and in the proxy Hub.

As we noticed before, the dimension left to explore would be that of a multi-instrument highly decentralized and highly heterogeneous, although this is how La Roda was originally conceived.

Highly centralized and homogeneous setups, by the other hand, are a new smaller scale performative environment which depart from the archetypal repertoire from the LOrk and HUB ensemble models (respectively, highly expansible and decentralized homogeneous systems and small-scale and decentralized heterogeneous systems). They may challenge the assumption of specific roles and distinctive musical voices in an ensemble performance but otherwise they do offer the opportunity to explore highly unified and coherent performative environments which may multi-user instrument.

A last dimension explored refers to the topology of interdependence, and evaluates both the amount of resources shared and the extent of such sharing. As we may appreciate, we did not yet explore environments in which a resource was only partially shared. Shared resources are, at a given time, either not owned or completely owned by a single performer, except for The Hub which, as a completely shared environment, do not provide support for resource ownership. Will address this remaining scenario in the Experiments conducted on Chapter 8.

4.4 Conclusions

This chapter reviewed the network music repertoire developed and explored by the Barcelona Laptop Orchestra during the last five years of artistic practice.

We surveyed a number of score based compositions, collaborative sequencing environments and, finally, multi-user instruments which explored different topologies of interdependence. The repertoire was analyzed in the context of the taxonomies and classification criteria reviewed in Chapter 3.

This comparative analysis served to place the BLO under the *Hub performative model*, making it clear our shared approaches but highlighting as well our novel proposals. We particularly developed new strategies to balance performer's *ownership* and *shareability* in shared environments and showed the usefulness of machine listening techniques to assist the conductor and performers with the coordination processes required to guide a performance with a multi-user instrument.

In this respect, our preliminary conclusions were mainly drawn from our own involvement in the development and artistic practice with the aforementioned instruments.

Finally, it is worth pointing out that our mostly centralized, network-based multi-user instruments greatly facilitate the capture and analysis of performers' actions. Some tentative performance plots were already provided for the Roda and the Conducted Roda instruments, showing their usefulness to monitor the performance, evaluate the contributions of performers and even try to correlate the relationship between their actions and the musical outcome, indeed a first step towards a more quantitative performance evaluation.

Those preliminary analysis motivated our further research into measurement based techniques for music performance evaluation. In the following Chapter we will carry out an extensive review of the research in music performance evaluation, with a specific focus in analysis based on performance measurements and their application for computer based collaborative musical interfaces and computer music ensembles.

Chapter 5

MUSIC PERFORMANCE EVALUATION

5.1 Introduction

This chapter will review the literature on music performance evaluation.

Closely linked to the massive recent growth of music psychology [Hargreaves et al., 2012], evaluation of music performance asks for an interdisciplinary approach encompassing art, sciences and social sciences. This chapter will survey some major research topics and methodologies related to music performance. It will summarize the most relevant research carried out in the areas more closely related to the subject of this Thesis: performance measurement and evaluation and improvisation analysis. A closely related subject, models of performance and improvisation, is left for Chaper 5, where we discuss theoretical models of musical creativity.

While the vast majority of the research in music performance is concerned with the interpretation of composed (i.e. score-based) music rather than in improvisation, with a specific focus in classical Western Music, such studies nonetheless illustrate the common methodologies and challenges than a researcher must face when studying the complexity of this multifaceted phenomenon, and are therefore of significant importance to address our research.

Additionally, we will survey the smaller but growing body of research involving group performance and improvisational performance (or both simultaneously, as in research studies on Jazz), particularly relevant for this research.

The last sections will discuss the emergent research on performance evaluation in the context of computer-supported Collaborative Musical Environments and Laptop Orchestras

5.2 Music Performance analysis

Alf Gabrielson carried in an exhaustive review of the field of Music Performance research in 2003 [Gabrielsson, 2003]. If dated, it may serve to delimit the main areas of empirical research on music performance. According to the author, the majority of research was on measurements of performance, but with growing contributions related to performance planning, performance practice and formal models of music performance.

Performance research historically developed in parallel with psychology, and later music psychology studies. Since its beginning, notably around the 1920s and 1930s, the focus was on performance measurements, specifically on timing issues. The first empirical studies, carried out by Carl Emile Seashore in Iowa, already stated the basic directions of music performance analysis: the study of the relationship between expressivity and performance deviations from the score.

The artistic expression of feeling in music consists in aesthetic deviation from the regular – from pure tone, true pitch, even dynamics, metronomic time, rigid rhythms, etc. [Seashore, 1938]

After the War, there was a revival in such studies, aimed at confirming the hypothesis of systematic variations in timing. We could mention here the work of Ingmar Bengtsson at Uppsala University throughout th 60s and 70s, and the research on expressive timing in Bach harpsichord performances by Povel in 1977. Empirical performance research proliferated since the mid 70s, still keeping the focus on measurements of performance but gradually incorporating novel research areas such as motor processes, performance models and psychological and social factors. As the focus was traditionally on high-art Western Classical music, performance research on improvisational genres, either in Western or non-Western cultures, was comparatively much scarcer.

With the advent of MIDI technologies in the 80s, and particularly the release of the Disklavier MIDI piano in 1987, a solid platform for experimental research was finally available and, indeed, most of the subsequent research was focused on piano performance.[Clarke, 2004]

5.2.1 Measurement of performance

As a prevalent research topic in music performance, the literature on measurement of performance of is rather extensive. We will now summarize some key researches on the field of empirical performance analysis, both focused on single performers or measuring music performance in music ensembles.

One of the main problems in music performance measurement deals with the data acquisition problem. While the study of timing and dynamics in keyboard performance may benefit of readily available performance encoding solutions for capture, recording and analysis -mainly through MIDI interfaces and software-, addressing other musical parameters, instruments and musical practices must be addressed with specific technological resources which typically deliver massive amounts of multi-modal data. Alternatively, performance data may be extracted or inferred from the acoustic audio signal itself.

Virtually all the parameters which may be significantly altered through the performance of a music score have been addressed in specific studies. Apart from timing, understood as deviation from strict adherence to the norm or score, the traditionally primary subject for most papers, some researchers devoted to the study of variations in dynamics, intonation and vibrato as well - though they are more instrument specific.

It is rather well acknowledged that there are two main, closely related causes for such performative deviations: as a resource to frame, or highlight specific structural traits [Palmer and Deutsch, 2013] or as a resource to express mood or emotion [Juslin, 2003].

Additionally, while performance measurements usually address deviation and variation from the score for the purpose of musical expression, some papers account for the analysis of plain performance mistakes as well. Such studies may be relevant in the context of pedagogy research and may provide insight into the cognitive factors involved in the acquisition of performance kills (see for example [Repp, 1996b] and [Palmer and Drake, 1997]) or in studies of interaction between auditory feedback and motor performance,as in [Maidhof, 2013].

Timing

One of the classical and most cited series of research papers on timing is due to Bruno H. Repp. His research focus on the analysis of timing in piano performances of classical music. His studies were based either in data capture from actual performances of transcribed from historical recordings.

On comparisons between performers playing Schumann, and later Debussy works, Repp highlighted the differences between students and seasoned performers: a higher homogeneity was observed in students, adhering more closely to performative canons and displaying more similarity, repeatability and consistency of timing patterns compared to a noticeably higher originality in famous pianists [Repp, 1992] [Repp, 1995] [Repp, 1997a]. On another study on a Chopin piano édude, Repp employed Principal component analysis to study the expressive timing deviations at the end of melodic gestures, noticing a complex and idiosyncratic use of a few shared, basic timing strategies [Repp, 1999b].

Still focused on piano performance, Repp and Palmer both studied the timing of sustain pedal with respect to note onsets. As actual piano practice confirms, its ordinary use in performance showed to be consistent and maximizes the legato while avoiding undesired, dissonant overlap [Repp, 1996a] [Palmer, 1996a] [Palmer, 1996b].

Another well know performative deviation in piano performances is the subtle asynchrony between the melody and the accompaniment to enhance the perception of cantabile. Sometimes the melody is slightly anticipated and sometimes slightly lagged. Palmer observed that such performative resource was rather more pronounced in expert performers [Palmer, 1996b]. Melody lead is much more pronounced in jazz performance, up to the point that notes may anticipate or delay several beats. Ashley noticed a consistent pattern consisting of lagged notes followed by accelerando and alignment with the accompaniment at cadencial locations [Ashley, 2002].

Still another timing-related research is the investigation of systematic variations in timing. Humair conducted and study addressing a number of popular danceable musical genres, from baroque dances such as the Bourrée and Menuet to the Vienese waltz and XXth century genres such as swing or tango, and confirmed the existence of systematic rubato patterns [Humair and Genève., 1999]. Johnson also found typical arch-like rubato phrasing patterns in the study of timing deviations in a Bach 's Suite for Cello [Johnson, 1999], while Davis found evidence that violinist resorted to rubato to bring out the implied polyphony in Bach's solo violin music, but was largely avoided if the structural complexity was already providing enough expressive cues by itself [Davis, 2009].

Other performance parameters

Although articulation is an important and particularly effective cue in music performance, greatly contributing to the emotional character of a piece [Gabrielsson and Juslin, 1996], considerably less work has addressed it. Articulation studies can be easily carried out with MIDI keyboards as well, because both note onsets and releases are precisely recorded. However, it is more complex to analyze than timing for other musical instruments. Besides, timing instructions in a score can be much more unambiguous than articulation marks.

Bresin carried out measurements on articulation, including amount of overlap time and inter-onset intervals, for a number of performances of an excerpt of a Mozart piano sonata. The author asked participants to perform an optimal performance and then some expressively biased renditions. Similar strategies were observed among all the participants regarding the relationship between IOI ratios and overlap or note duration both in legato and staccato passages. The aim of his research was to build rule grammars for automatic music performance models [Bresin and Umberto Battel, 2000].

Deviations of equal tempered tuning, either static or dynamic (pitch inflections), may be an expressive resource for players, specifically singers and string instrumentalists. In the case of string ensembles, such deviations may be adaptive as well, coping for the harmonic requirements of an ensemble performance.

Another performance variable is the vibrato. Its frequency extent and intensity do not only obey to stylistic constraints but to the musical context, the performance setting and the idiosyncrasy of performers. For example Prame [Prame, 1997] studied the use of vibrato in singers and violinists performing Schubert's Ave Maria. The vibrato extent was larger for shorter notes and larger for singers compared to violinists.

A more holistic approach aims at the study of the relationship between several performance dimensions. As a first step towards this goal, several authors analyze the coupling of two variables. The relationship between timing, or tempo and dynamics has been addressed by several researchers, questioning an hypothetical simple correlation between tempo and dynamics [Clarke and Windsor, 2000] [Repp, 1999a]

Relationship between tempo and timing is addressed, for example, is studies of swing in jazz performance. Friberg and Sundström, for example, observed that the swing ratio decreases when tempo increases, though the swing ratio of the soloist was lower [Friberg and Sundström, 2002].

The study of swing deviations has attracted interest by a number of researchers. Prögler [Prögler, 1995] measured the timing of drummers in relation to a stable beat and observed a systematic, though heterogeneous and dynamic use of timing asynchronies of the ride.

Finally, comparative studies between performers over multiple expressive variables may be facilitated through techniques for summarized performance visualization. In this respect, maybe the most is the Performance Worm, by Dixon et al., a visualization of the tempo and loudness evolution which may provide insight into the expressive patterns applied by performers [Dixon et al., 2002]. As a technique for data reduction, it may be used to computationally evaluate and compare performances and reveal both diversities and commonalities among performers [Goebl et al., 2004].

5.2.2 Subjective evaluation and Judgment

While music performance is commonly evaluated by all the actors involved - the performer itself, the audience, the critics and so on - it is unclear under what criteria and how the judgments are done [Gabrielsson, 2003]. A common approach is to resort to performance assessment schemes drawn from educational contexts. However, as Thomson observes, assessing the quality of a musical performance is a fairly controversial topic, as the typical inconsistency on ratings by expert evaluators demonstrate [Thompson and Williamon, 2003] [Bergee, 2007]. Both technique and musical expression factors have proved to have a significant, direct effect on the assessment of overall performance quality of a performance [Russell, 2015], but the whole picture is rather more complex as we will see.

Some papers incorporate listener ratings to confirm the relevance of certain observed performance patterns, sometimes with initially paradoxical conclusions. For example Repp compared listener judgments between individual performances and an artificial, "averaged" performance of the same piece, which surprisingly received the highest ratings in terms of quality, but not of individuality [Repp, 1997b]. Similarly, Johnson confirmed a correlation between perceived musicianship and increased use of rubato in a Mozart Concerto [Johnson, 1996]. A plausible conclusion would be that listeners may trade originality and idiosyncrasy for accuracy or adherence to aesthetic canons, at least to some extend.

Johnsson studied the effect of rubato magnitude in the perception of musical musicianship on a Bach's Cello suite, by providing to listeners performances with different degrees of alteration of a reference mode derived from performance measurements. His findings indicated that both the performances with less rubato or exaggerated rubato were perceived as less musical [Johnson, 2003].

Saunders and Holahan carried out a large scale, multiple-item evaluation involving high school music students. The judges evaluated their performance by applying criteria-specific rating scales to a number of technically related dimensions, such as technique, intonation, melodic accuracy, rhythmic accuracy and so on. Through step-wise multiple regression analysis the authors were able to predict the total score, with good accuracy ($R = 0.96$) as a combination of just five dimensions [Saunders and Holahan, 1997].

In a similar study involving teachers, peers and the performers themselves, Bergee observed good correlation between faculty and peer evaluations and poor correlation with self-evaluations [Bergee, 1997]. There was a tendency to over-criticism, as Daniel observed as well on a study targeted to self-assessment through the evaluation of video recordings of students' own performances [Daniel, 2001]. In another large-scale study involving 373 music students, a similar pattern of inaccurate underestimation of self-assessment was observed as well [MA, 2005].

While evaluating students does usually focus on technical issues, with expert performers the focus shifts towards a higher number of dimensions. Therefore, instead of analyzing the ratings, some studies resort to more qualitative analysis techniques to determine what are the parameters that judges utilize to evaluate a performance. For example, Davidson and Coimbra used a qualitative analysis to assess the main parameters which judges utilize to evaluate professionals singers onstage. Questionnaires and interviews were utilized to elicit the criteria for assessment. Factors such as technical control, but also appeal, bodily communication or facial expression where taken into account, while surprisingly the timbre of the voice was not mentioned [Davidson and Coimbra, 2001].

The effect of such so-called non-musical factors in the evaluation of a performance was clearly demonstrated in Howard's study on solo vocal performances [Howard, 2012]. Adjudicators assigned higher ratings to performances presented in audio format as compared to audiovisual format, showing in the later case that quality ratings were significantly affected by factors such as soloists' performance attire and stage deportment.

Thomson used a more formal approach, resorting to *Repertory Grid* methodology (see [Fransella et al., 2004]) to elicit the constructs utilized in the evaluation process by professionals adjudicators, and overall assessments were found reliable, with moderate degree of agreement between judges. Pedalling was the most common construct and showed strong correlation for overall preference, but even stronger correlation was found with right hand expression, phrasing and expression at the end of the piece [Thompson et al., 1998].

The authors suggest that their methodology may contribute to be aware, refine and develop the skills of adjudicators and ultimately increase their reliability by standardizing the evaluation criteria. It may be useful to competitive performers as well, to gain greater understanding on the process of adjudication.

5.2.3 Performance evaluation in music ensembles

Interpersonal coordination

Ensemble performance analysis must account not only for the compliance to (and deviations from) the score, but to the joint action between performers, a complex interplay based on continuous proposals and re-adaptation to a continuously shifting context. Therefore, in an ensemble performance we may simultaneously study individual deviations and overall ensemble deviations, and the relationship between both.

Once again such deviations, now not only or not particularly from the score but between performers as well, set a driving force for musical interplay and expression. In 1987 Charles Keil introduced the term Participatory Discrepancies [Keil, 1987] to refer to such '*slight human inconsistency in the way that a musician executes rhythm, pitch and timbre*'. In a later article, Keil [Keil, 1995] defends that such discrepancies are actually the basis for musical creation, the true source of meaning and groove in music performance.

How such inconsistencies are negotiated throughout the performance has been a topic addressed by cognitive psychologists. And similarly to individual performance, ensemble performance measurements have been also focused on timing issues, particularly on relative asynchrony and mutual synchronization between players. The early quantitative research on ensemble coordination resorted to laboratory studies of sensory-motor synchronization (SMS). Repp carried out an extensive review of the tapping literature on SMS [Repp, 2005], including extensive references on studies of SMS in musical contexts. Keller and Pecencka continued this line of research in inter-personal synchronization through laboratory studies[Pecenka and Keller, 2009][Pecenka and Keller, 2011].

Summarizing his research on the subject on a theoretical framework, Keller [Keller, 2007] [Keller and Repp, 2008] [Keller and Mills, 2008] [Keller, 2013] pointed out three key cognitive processes present in ensemble musicians: auditory-imagery, where the musicians have their own anticipation of their own sound as well as the overall sound of the ensemble, prioritized integrative attention, where musicians divide their attention between their own actions and the actions of others, and adaptive timing, where each musician adjusts the performance to maintain temporal synchrony. This perspective makes it possible to identify the following important points: first, that each musician incorporates the ensemble score as well as the performance of the rest of the ensemble into an anticipation of the produced result. Second, the musician defines the saliency of each performed note with respect to the ensemble as a whole, shaping their performance so that it integrates both with the ensemble's actions as well as personal expressive choices. Lastly, the above choices must be made while maintaining ensemble synchrony at the same time.

More recent research addresses multiple player synchronization by taking into account multimodal non-verbal communication cues and social hierarchies within the ensemble. For example, Goebl and Palmer [Goebl and Palmer, 2009] studied the ensemble synchronization in restricted auditory and visual feedback scenarios. They found that decreased auditory feedback increased temporal asynchronies and motivated higher reliance on visual cues. Additionally, they observed that bidirectional adjustments where present even in a leader/follower scenario.

Research on interactional coordination in ensemble performance is not only limited to the analysis of synchronization of note onsets, any may be addressed through multidimensional analysis of the ensemble performance, eventually incorporating multimodal data to cope with gestural and non-verbal communication cues taking place during the performance. We should mention here the research carried out within the SIEMPRE project (2011-2013)[1], aimed at the study of the mechanisms underlying social cognition and co-creativity in ensemble musical performance and audience experience.

For example, Glowinski [Glowinski et al., 2012] identified patterns of dominance and inter-musical communication in a string quartet through the analysis of the musicians' heads movements, and Papiotis [Papiotis et al., 2012] [Papiotis et al., 2011] carried out measurements of ensemble synchrony through intonation, timing, dynamics and articulation adjustments through audio feature extraction. In a related research, Wing [Wing et al., 2014] studied the synchronization in two string quar-

[1] `http://www.infomus.org/siempre/`

tets and revealed that tempo adjustments reflected the ensemble hierarchy, with the first violin exhibiting less adjustments in the first, more autocratic ensemble as compared to the second, more democratic one.

In his PhD thesis, Marchini [Marchini, 2014] addresses the analysis of expressive ensemble performance in string quartets, by both statistical and machine learning techniques. By comparing the timing of a solo versus an ensemble performance, the author observes the restrictive effect of synchronization demands on performative licenses, and a tendency to exaggerate timing accents when playing in ensemble. More significant was the machine-learning approach, which confirmed a preference for horizontal features in solo playing and a preference for vertical, or inter-voice features in ensemble playing in his predictive model.

5.2.4 Group Improvisation

As a particular research topic in ensemble performance evaluation, improvisation studies are scarce and mostly centered on jazz. Beyond the aforementioned studies on timing in jazz performance, this genre provides a suitable framework to evaluate improvisational strategies and the development of improvisational resources by a performer. Performance studies on jazz either focus on the syntactical structure of jazz improvisation, emphasizing the formal analysis of transcriptions of improvisations, or may otherwise insist upon the centrality of interaction in a jazz performance, therefore placing greater emphasis in coordination issues through the performance.

As a relevant example of an analytic approach, we should mention Järvinen [Jarvinen, 1995], who studied bebop improvisations based on a popular chord progression and found statistical similarities to European art music in terms of tonal hierarchies. In this article and specially in a later one [Järvinen and Toiviainen, 2000], Järvinen confirmed that jazz improvisers match key notes -with respect to tonal centers- to strong beats.

As for coordination and interaction studies concerns, Reinholdsson [Reinholdsson, 1998] studied the processes of negotiation and musical and social experiences within small-group jazz performances. Through investigation of group performances, Reinholdsson observed contrasting episodes of cooperation and flow versus problematic passages of resistance and distrust. The author advocated for an interdisciplinary approach for such an study, encompassing social psychology, ethno- and socio-musicology.

Combining both approaches, Hodson [Hodson, 2007] evaluated improvisational decisions in the context of interactive possibilities within the ensemble. By resorting to transcriptions of progressive and free jazz improvisations, Hodson studied how jazz musicians interact in a performance, with specific examples on re-harmonization practices and techniques to clarify or obscure structural and phrase boundaries in improvised jazz solos.

Weisberg [Weisberg et al., 2004] carried out a statistical pattern analysis of transcribed solos of bebop musicians. It allowed him to determine how much "formulaic" improvisers were, that is, what were the most utilized melodic patterns and to what extend the whole solo was constituted of a concatenation of such patterns. Additionally, they compared patterns among several performers. The projection was statistically significant in one case, suggesting stylistic similarity or an influential relationship between two performers. A limitation of the analysis was that the it only accounted for exact pattern repetitions, therefore it disregarded any eventual processes of pattern transformation, the authors suggest a hierarchical analysis of patterns would provide further insight into the creative process of constructing a solo improvisation.

5.2.5 Performance evaluation for CSCME and Laptop Orchestras

Studies of mediated human-human interaction in computer based collaborative musical interfaces have been growing only in recent years. Through analysis of musical performance with collaborative interfaces, researchers study group flow and interaction, which in turn can inform interface designers about usability issues from a social perspective. Related studies analyze user interaction when performing or improvising collectively with tangible musical interfaces. Most studies resort to study users behavior in undirected, free improvisational scenarios. Studies on conductor-ensemble interaction in collective digital ensembles are, to this author's knowledge, limited to ethnographic studies.

Concerning computer ensembles, and particularly Laptop Orchestras, the literature on performance analysis is still scarce, and mainly resorts to a few ethnographic analysis to investigate the social interaction within a performance.

We will review in this subsection the main literature concerning the analysis of musical performances in both performative contexts.

Music performance evaluation in CSCME

The analysis of musical performance on Computer Supported Collaborative Music Environments, compared to performance analysis on Computer Ensembles, is both quantitatively and qualitatively more relevant, and is aimed both towards a better understanding of the dynamics of the collaborative and creative processes taking place within the computer mediated environment, and as a tool to inform and provide valuable design criteria.

Weinberg's research in the 1998-2003 period [Weinberg, 2003] was focused on the design of tangible musical interfaces with high-level controls and built-in interdependence to allow children and novices to participate in meaningful, social musical experiences. He designed a number of collaborative interfaces halfway between actual collective instruments and interactive sound installations, both performatively simple but offering complex interaction to the players.

The evaluation was largely informal, through retrospective discussion with participants. During the design phase, composition or evaluation of his interfaces and repertoire written for them [Weinberg, 2002] [Weinberg and Gan, 2000] he noticed novices preferred high level musical controls, rated as more expressive, whereas musicians were more interested in direct control of lower level parameters. Players also stated they had to decide whether to concentrate on the effort required to conduct social interaction through non verbal cues or the nonetheless substantial effort to perform expressively with their own interfaces.

When performers resort to computer interfaces, playing remotely or even co-located, non-verbal cues are severely limited and new issues of user awareness and graphical interface design for music interaction come into play.

Drawing on models of CSCW research,several authors [Murray-Browne and Mainstone, 2011] [Stowell et al., 2009] [Bryan-Kinns and Hamilton, 2012] study how improved shared awareness increases mutual engagement in shared representations (in short, collaborative interfaces), providing design guidelines. The author designed custom distributed collaborative interfaces (Daisyphone and Daisyfield) for concurrent music-making specifically tailored to his research. Additional observations refer to the usefulness of additional communication channels between users (like shared annotation mechanisms) but taking care of the cognitive load, for example applying a decay to the visualization of performers contributions [Bryan-Kinns and Healey, 2006].

Fencott [Fencott and Bryan-Kinns, 2010] conducted an experimental study targeted to musicians, to determine the effect of awareness and privacy in musical interaction, collaboration, contribution and emergence of roles in a collaborative music software environment. The software, which offers a number of sound-generating modules to users, could provide personal and shared workspaces with configurable awareness mechanisms. The evaluation was carried out by an introductory questionnaire, video recording, group discussion and interaction log analysis. The authors observed more contributions but less co-editing with private spaces, cognitive load issues caused by excessive visual and aural simultaneous awareness of all contributions and spontaneous assumption of specific instrumental roles.

In another related study [Fencott and Bryan-Kinns, 2012] the authors show that different audio delivery configurations (headphones, own speaker, public speakers) and interface layout affect the way participants interact through a collaborative performative environment. Through quantitative analysis of interaction logs, video recordings and user questionnaires the authors provide valuable insight into user and mutual awareness, territoriality and privacy issues, which have implications in the design of collaborative interfaces. A key finding is the evidence that private spaces increase personal contribution but greater interaction can be achieved by providing only shared spaces to users.

A relevant number of studies with collaborative interfaces resort to environments tailored to concurrent loop sequencing, with an emphasis on the drum circles metaphor (for additional studies see [Blaine and Forlines, 2002], [Beaton and Tech, 2010] and [Derbinsky and Essl, 2012]).

Klugel [Klügel et al., 2014] asked participants to freely explore two collaborative rhythmic sequencers. Their activity was evaluated both qualitatively through user questionnaires, video recording and quantitatively through data logging. Individual and group flow and engagement were assessed through facial and corporal cues. The authors noticed a rather consistent activity pattern, starting with an exploratory approach which gradually transformed into a goal-oriented interaction through negotiation, labor division and spontaneous assumption of leadership and coordination roles. The performance analysis measured the complexity of the generated rhythmical patterns, but it was inconclusive on whether it was caused purposefully or due to the prototype affordances.

Ben Swift [Swift, 2013] carried out an experimental research on free improvisational interaction with *Viscotheque*, a collaborative interface for mobile devices specifically designed for this study. To conduct the research, video recording, data log and final interviews with the participants served to evaluate their joint experience. The degree of engagement was qualitatively assessed and the evaluation was mostly through user interviews. The author states the challenge of resorting to traditional data-driven HCI measurement to evaluate such elusive and slippery concepts as group flow and mutual engagement, suggesting the potential uses of expert judgments, unsupervised learning and biometric data collection to measure open-ended, experience-centered interactions in HCI. Indeed the author resorts to machine learning to assist ethnographic analysis, showing its usefulness in the automatic recognition of players and groups of players from the logged performance data.

Concerning computer ensembles, and particularly Laptop Orchestras, the literature on performance analysis is still scarce, and mainly resorts to a few ethnographic analysis to investigate the social interaction within a performance.

Hansen conducted a number of studies centered to shared music instruments targeted to non-musicians, with game-like interfaces and strategies which at the same time are designed to encourage simultaneous improvisation. In [Hansen et al., 2011] the author set up a shared music improvisation interface encouraging a solo-accompaniment interaction. Both quantitative analysis of interaction logs and qualitative analysis of video captures were employed. She observed spontaneous assumption of roles, negotiation of turn-taking schemes and other conversational related resources use of coordinated pauses, like gesture based imitation. She finally provides design guidelines to strength music focused collaboration.

The same author [Hansen et al., 2012] carried out another experiment targeted to novices required rapid turn-taking negotiation to create interleaved melodic patterns. The task included some gamified elements such as sonic *rewards* working as a background rhythm when both players coordinated to achieve certain (unknown) patterns. Sessions were video recorded. The findings indicate that users tend to perceive more a causal relationship between sonic rewards and their individual actions instead of the collaboratively built patterns.

Xambó [Xambó et al., 2011] proposes the use of design patterns in interactive musical systems, drawing from design problems identified in the literature. Such design patterns, oriented to multitouch interactions, encourage the use of private and shared spaces, for developing ideas and for sharing and modifying them respectively, aim for a well designed learning paths suitable for novice exploration and later for skilled performers. Another design pattern calls for the need for the collective interface to handle unequal participation, leadership and role division.

The same author, in a more recent research [Xambó et al., 2013] conducted a long-term study on user interaction in Reactable improvisations. It was similarly addressed to expert musicians. It revealed that the rich interactions facilitated by a tabletop interface supported practice-based learning strategies (both learning by doing and peer learning). Additional findings related to collective collaboration are reported as well, such as the negotiation of a shared "storage space" and regular change of ownership of objects.

Music performance evaluation in computer ensembles

As previously stated, the research on performance evaluation and performance analysis in computer ensembles is still in its infancy. We will survey two qualitative ethnographic studies and two quantitatively informed analysis of performances with computer ensembles.

Two relevant ethnographic studies were carried out by David Reinecke -with the Princeton Laptop Orchestra- and Graham Booth -with the Birmingham Laptop Ensemble- respectively.

Reinecke's study [Reinecke, 2010], shows through rehearsal analysis how the assignation of conduction related tasks to humans or to computers can lead to coordination and governance conflicts. He observes how informal conventions on coordination and governance, which are often taken for granted and interactively negotiated and challenged in Jazz ensembles, or largely formalized, standardized and enforced with a well established hierarchy in classical music, can be disrupted with the incorporation of code for synchronization or shared control and ensemble coordination. He provides three illustrative examples drawn from the PLOrK repertoire

Booth [Booth and Gurevich, 2012] conducted an extended study with the Birmingham Laptop Ensemble, documenting the whole practices involved in rehearsals and performance. System logs of chat and audio and transcription of video recordings of the sessions were utilized to study the musical and interpersonal interaction of the group. Boo

Freeman and Troyer carried out an evaluation of LOLC, their environment for collaborative text-based laptop performance, with eight members of the Princeton Laptop Orchestra [Troyer et al., 2012][Lee et al., 2012]. Their study was aimed at evaluating how the environment promoted collaborative improvisation, as well as the learnability and understandability of the process. It consisted on qualitative and quantitative analysis of code logs, chat logs and musical output from a series of rehearsals and a final public performance, plus interviews with the performers and the audience.

Performers could, through simple textual commands, create, schedule, share and transform rhythmic patterns, in a manner loosely inspired in the aforementioned studies on improvisational interaction [Hodson, 2007]. Additionally, a visual interface graphically displayed the whole collaborative process. As the whole process of improvisational interaction was mediated through the chat server embedded in the LOLC environment, it provided valuable insight into the whole collaborative process, both for the analysis and for the audience. The collaboration metric evaluated was the amount of patterns shared and its precedence and scheduling times, showing different strategies within the ensemble: some performers did extensive use of borrowed patterns while the rest resorted to creating and modifying their own.

Through the analysis, some design recommendations arose in order to encourage more collaboration. The authors suggested the use of a better preview/prelistening functionality and promoting a more conversational interaction instead of the multilayered textures achieved through an excessive use of automated looping. was suggested to encourage more collaboration. Compared to a jazz performance, ensemble negotiation was found to be much slower mainly because of the inherent latency of interaction in a textual environment and the disruptive offline process of pattern scheduling. [Lee et al., 2012]

Finally, Charles Martin et al. [Martin et al., 2015] developed an interface for tracking ensemble performances in realtime. It consists of an agent that measures performers' interactions on touch-screens, classifies the sequence of gestures and estimates how quickly performers change them. They finally describe some possible applications such as rewarding original performers or disrupting those who stay on certain gestures too long. Under such approach, there is a mutual feedback between performance analysis and the performance itself.

5.3 Conclusion

In this chapter we surveyed the literature on musical performance analysis, with specific focus on the most recent research on performance evaluation in computer supported collaborative music environments and particularly in computer music ensembles.

A musical performance may be analyzed by measuring a performance (either actual performance data, or audio recordings of the performance) or by carrying out subjective evaluations of it. We first surveyed the most relevant findings in classical musical performance analysis. Such studies were mainly addressed at the study of musical expressivity by measuring timing deviations from the score or determining correlations between performance variables and structural traits in the score. A notable revival of measurement-based performance analysis was due to the adoption of the MIDI protocol in keyboard instruments in the late eighties.

By the other side, the analysis of acoustic instruments my means of specifically tailored motion capture devices, as well as the analysis of ensemble performance, is more recent and comparatively scarcer. In this respect, analysis on improvisational styles such as jazz tend to be more qualitative, in this case researchers studied higher level facets such as motive development and stylistic adherence to particular genres, though timing and coordination between musicians have been addressed as well.

The evaluation of a performance is a cultural construct. Only unambiguous such as performance errors may be quantitatively assessed. Beyond that, studies confirmed that both the performer, the audience and the experts in that field are hardly reliable when assessing the artistic value of a musical performance.

Regarding the analysis of music performance in computer mediated environments (either with collaborative tangible interfaces or with multi-user instruments in computer ensembles), we conclude that the majority of research conducted mainly consist of qualitative, ethnographic studies, combined with quantitative measurements, greatly eased thanks to the data-logging facilities of computer-based musical environments. Synergies between both approaches are possible, for example in Ben Swift's studies, which resort to machine learning techniques to assist his ethnographic analysis.

It is worth noting that all of the studies conducted in the former environments were aimed at the study of user interaction on improvisational scenarios. The actual musical outcome, nevertheless, was less relevant compared to the study of user interaction. Their purpose was twofold: further our knowledge of ensemble behavior in terms of negotiation strategies, exchange of materials and emergence of roles to assist in design recommendations and provide design patterns which will maximize mutual awareness and mutual engagement in collaborative interfaces.

Our research will follow a similar approach, by drawing conclusions from analysis based on actual measurements of improvisational performances, and combining them with user-based evaluation with ethnographic methods as discussed in Chapter 3. Additionally, we will propose to go a step beyond in the analysis of performer's contributions, by developing a performance-based metric which will allow us to quantitatively measure how creative is a performer, both bringing back musical creativity as a criteria for the evaluation of musical interfaces and without relying on the inconsistency of subjective performance evaluation.

The next two chapters will be therefore focused in the development of such creativity metric. Chapter 6 will introduce the subject of Creativity, its definition, characterization and assessment, while we leave for Chapter 7 the development of our proposed metric to compute musical creativity from performance analysis.

Chapter 6

CREATIVITY

6.1 Introduction

This chapter will survey the literature on Creativity in general, and Musical Creativity in particular, with a focus on computational methodologies for creativity assessment.

We will first provide a definition for Creativity, by itself, as we will see, a rather elusive and controversial term, and the we will proceed with a review of the main theoretical frameworks on the subject. Two specific perspectives are relevant in the context of this Dissertation: the cognitive, which attempts to explain both the nature and the dynamics involved in the creative process, and the social, which puts greater emphasis on the situated nature of creativity and the emergence of collective creativity. We will review the research from both perspectives.

Next, we will introduce the term Computational Creativity, and will survey the most relevant proposals to computationally evaluate creative *products*. Specific metrics for the most relevant creativity attributes will be provided, with a focus in Maher's creativity metrics, on which our own creativity metrics will be based.

In the following sections we will discuss the main research topics in Musical Creativity, in which the social context pays an essential role. Finally we will survey the cognitive models for two facets of creativity in musical performance: musical expressivity in score-based performances, and musical improvisation, ending this chapter with a review of the incipient computational models for creativity assessment proposed in the literature.

6.2 Creativity

6.2.1 Definition

We may tentatively define Creativity as *the process of generation of products or ideas that are both novel and appropriate*, a deceptively simple definition provided by Hennessey [Hennessey and Amabile, 2010]. However,there is no a universally accepted definition of creativity.

Taylor [Taylor, 1988] reviews the literature for a wide number of contrasting definitions of creativity. Definitions mainly differ from the approach used by investigators in their research, from those that emphasize the creative process as a mental phenomenon to those that emphasize the social grounding of creative acts, from those who focus in the creative product itself to those who look at the subject who produced it.

Sarkar [Sarkar and Chakrabarti, 2011] carried a comprehensive survey of the definitions of creativity, counting up to 160 different definitions. From these definitions, though majority and relationship analysis, the author proposed a common definition as follows: *Creativity occurs through a process by which an agent uses its ability to generate ideas, solutions or products that are novel and valuable.*

This is indeed a definition with closely resembles the aforementioned definition of creativity proposed by Hennessey.

6.2.2 Theoretical frameworks

In this section we will review the most relevant theoretical frameworks on Creativity. Modern perspectives on Creativity research take a multidisciplinary approach, incorporating the latest developments in social psychology, linguistics, philosophy and cognitive science to mention a few. In the later, we should mention the more recent contributions in the fields of Artificial Intelligence and Neuroscience, which is beyond the scope of this Thesis.

We will first cite some relevant historical contributions to creativity from the psychological perspective, followed by a survey on the cognitive, social and artificial intelligence approaches to understand the creative process.

Early research on creativity

Research on nature and dynamics of the creative process has been qualitative and largely empirical, and was initially focused in the description of cognitive processes and creative leaps by exceptional creators. We must mention here the pioneering work by Graham Wallas [Wallas, 1926]. His classic -stage model is considered one of the first models of the creative process. The author identified four distinct cognitive phases in the creative act: preparation, incubation, illumination and verification. Wallas work paved the way for further models of the creativity process in the XX century.

Arthur Koestler in his book *The Act of Creation* [Koestler, 1964] describes the process of creativity in a comparative study between such apparently disparate fields as humor, science and art. He observes the common pattern of *Bisociation*, the term the author proposes to indicate the fusion (in science) or juxtaposition (in art) of two simultaneous perspectives. They set up a creative tension which is explored in art, or resolved in a new paradigm in science, ultimately leading to what is know as a creative leap.

Cognitive approaches to creativity

From a cognitive and philosophical perspective, Boden [Boden, 2004a] refers to creative processes as operations on *conceptual spaces* which may be described as structured styles of though. Individuals explore and modify such conceptual spaces though distinct behaviors to produce creative products or artifacts. Modifications on conceptual spaces consist of *combinational*, *exploratory* and *transformational* processes.

Further, the author proposes a two-axes taxonomy of creative systems. A first dimension highlights the dichotomy between historical and psychological (or absolute or personal) creativity, whereas the second dimension distinguishes between merely exploratory and truly transformational creativity (in which creative processes redefine the rules and constrains of the concept space to produce paradigm shifts). According to the first dimension, Boden establishes two basic kinds of creativity: *H-creativity*, or historical creativity, whose artifacts fall outside the range of artifacts previously produced (in any society and in any context) and *P-creativity*, or personal creativity, whose artifacts fall outside the range of artifacts previously produced by that creator.

The above mentioned notion of combinational processes which give raise to creative artifacts are the base of the *Conceptual Blending framework*, formalized by the cognitive linguists Gilles Fauconnier and Mark Turner [Fauconnier and Turner, 1998, Turner and Fauconnier, 2002] and later on by Line and Per Aage Brandt [Brandt and Brandt, 2005]. In the original blending theory two or more mental conceptual spaces are blended by mapping rules and further elaborated and completed to create new concepts.

Boden's framework has been incorporated to creativity research in other disciplines, notably in the context of design science. Gero [Gero, 2002] identifies computational processes for creative design as combination, transformation, analogy, emergence, and first principles. Analogy is related to the above mentioned Conceptual Blending theory as the transfer of concepts from different conceptual spaces, while Emergence originates from a re-representation process of preexisting concepts though a reinterpretation under a new structural framework. Finally, First Principles would be the mechanism which allows direct product generation without deriving them from any preexisting concept, but just from the production rules.

Wiggins [Wiggins, 2006] refines formally the original Boden's model and proposes three rule-sets which operate on a conceptual space, namely the rules for concept membership, to constrain the concept space and determine membership, and the rules for concept construction or detection, which allow to traverse and search such space. The last remaining rule-set is the set of rules used to evaluate the "quality" of a concept, according to the criteria we consider appropriate.

The author further classifies Transformational creativity into two distinct procedures: transformation of creation rules (which may lead to truly novel concepts) and the personal technique of the creator, which will allow him to reach pre-available concepts by following a specific path.

Finally, the author elaborates a formal model of creativity which helps in identifying properties of different creative behaviors. Under his model, *uninspiration* may be seen as the lack of concepts or valued concepts in the conceptual universe or the lack to attain them within the space constrained by the technique applied. The author terms *aberration* a concept generated employing the available creation rules which does not fill in the existing conceptual space, it may be incorporated in our space or not according to their value, and in the former case it paves the way for truly transformational creativity.

Although Boden admits that creativity implies positive evaluation and that acknowledge is implicit in H-creativity, the author focuses mainly on novelty as an attribute of creativity in both cases, which do not tell us the whole picture : ideas novel both to the individual and to its social context may be agreed to be worthless. Wiggins' model, on the other hand, may account some cases of creative but not valuable products, though the author admits that the modeling of relative social value needs more elaboration.

Dynamics of creativity

While the aforementioned cognitive models of creativity attempt to determine the cognitive procedures involved in the creative act, they give us little clue as to what precise sequence of mental processes leads to the genesis of a creative idea or product. The following authors theorized, instead, about the dynamic process of creativity.

Gabora's *Honing Theory* [Gabora and Aerts, 2009] describes the creative process as a resolution of potential *wordviews* through self-organization. A first stage of interaction of the ideas within newly generated contexts self-modify the worldview making new associations, reaching a percolation threshold which triggers cognitive insight.

A throughout literature review of Creativity Process Models is provided by Sawyer [Sawyer, 2011], who proposes himself an eight-stage Model that takes into account former frameworks. Sawyer incorporates, and describes experimental evidence [Csikszentmihalyi, 2014] of the Incubation stage proposed by Wallas, thus bringing back the importance of the unconscious processes in creativity, as some other authors recently highlighted [Ritter et al., 2012].

Only recently some studies address quantitatively the dynamics of creativity. In [Noy et al., 2012] the authors identify two distinct user behavior in the exploration of a solution space : *scavenging* similar products and sudden jumps to a new region of the product space. Such unexpected dynamics seems to be in accordance with the above mentioned notion of creative leaps, although in a highly simplified laboratory setting.

Sociological approaches to creativity

The first wave of creativity research, in the mid 20th century, focused on the personality traits and cognitive processes of individual creators. Even if relevant insights resulted from this studies, by the 80s researchers realized that creativity could be better explained by widening the focus and began to explore its social and cultural dimensions[Amabile, 1983]

Through the 1990s and specially the new millennium, creativity research focused in the paradigm of distributed creativity, which investigates how creativity is embedded in social groups. This new perspective, drawing from research by sociology and social psychology, challenges the traditional assumption of creativity as an essentially individual outcome.

Pioneering this second wave of creativity research, Csikszentmihalyi proposed in 1988 a systems view of creativity which incorporated this multifaceted perspective of creative processes [Csikszentmihalyi, 2014]. The author identified three components of a creative system: the individual, who generates the idea, the domain, either cultural or symbolic, and the field, either social or interactive. Creativity is then not an individual act, but the whole process of interactions between the components of such system.

According to Csikszentmihalyi [Csikszentmihalyi, 1999], there is no way to separate the reaction of society from the person's contribution, and this is a bidirectional relationship: individuals take concepts from the cultural domain, transforming them and, if deemed valuable in the relevant field by the society, finally incorporated in that domain. Csikszentmihalyi system components are closely related to the so-called *Four Ps of creativity* introduced by Rhodes [Rhodes, 1961] : the Person (the individual), the Product (or creative outcomes), the "Press" (the social context) and the Process by which such products are developed.

Cohen[Cohen, 2012] states that Creativity is an emergent property of socio-cultural systems. The author proposes a decentered model for artistic creativity, in which there are multiple simultaneous contributing factors: creativity coming from the individual while, simultaneously, artistic forms and conventions shape the creative practice of individuals and groups. To Cohen the whole creative process is inherently relational: in a socio-cultural context it emerges through recursive patterns of imitation and appropriation/reworking.

Computational Creativity

Computational creativity arises from the convergence of the disciplines of Artificial Intelligence and creativity research. Wiggins provides a suitable definition of computational creativity as *The study and support, through computational means and methods, of behavior exhibited by natural and artificial systems, which would be deemed creative if exhibited by humans* [Wiggins, 2006].

Accordingly, Computational creativity investigates the use of computational models and creative agents to assess, encourage or interact with human creativity.

While studies on human creativity have been traditionally focused on the cognitive behavior of creative individuals, the study of computational creativity favors philosophical as well as AI research [Maher, 2010].

Much research in the field of Computational Creativity has been focused in modeling creative behavior with software agents.The resulting computational models simulate the reasoning process with software agents which can produce creative designs [Gómez de Silva Garza and Gero, 2010]

Some authors even explore the social influence in creativity though software agents, effectively incorporating the aforementioned paradigm of distributed/situated creativity into computational models. The theoretical framework upon which such models are built was incorporated by Sosa and Gero [Sosa and Gero, 2003] into the domain of multi-agent modeling of design creativity. For example, Gero and Saunders [Gero, 2002] propose computational models of collective creative design in which social awareness and *situatedness* strongly determine the behavior of a multi-agent system.

6.3 Creativity assessment

We have so far surveyed the literature on Creativity centered on the research of the processes involved, either social or cognitive.

In order to *evaluate* creativity, though, two possible perspectives are possible: process based and product based approaches. As we have seen the internal or social dynamics of a creative process are far from being understood, therefore quantitative approaches in creativity assessment are mostly based on the study of creative products.

It must be noted that, if the very definition of creativity may be problematic and lacks enough consensus, evaluation methodologies to assess creative products are consequently far from being standardized.

For the last forty years, identification and assessment of creative products has been relied on the consensual assessment of experts. This is precisely the basis of the well established *Consensual Assessment Technique* (henceforth, CAS) [Simpson, 1982] [Hennessey and Amabile, 1999]: the idea that the best measure of creativity is the combined assessment of experts in that field, without resorting to any particular theory of creativity [Baer and McKool, 2009]

Creativity ratings obtained by CAS are always comparative, they refer to relative differences within the group of creative artifacts evaluated. But the simplicity and consistence of the method, it is still widely employed in creativity assessment. Ultimately, public recognition is a requisite for creative products [Csikszentmihalyi, 1999] and the ratings of experts are just a realization of the social component of creativity.

Many other evaluation techniques have been proposed, some of them domain-specific. For an extensive review we refer to Oman [Oman et al., 2013].

Current research in computational creativity, however, advocates for more uniform evaluation criteria, which ideally should not depend on experts, but should instead be independent of the generative process, of the domain and of the creative entity itself, be it a person, a computer agent or a combination of both [Maher, 2010],[Ritchie, 2007].

6.3.1 Characterization of creative artifacts

Several authors have proposed a number of *properties* that may help us in characterizing a creative artifact. Such contributions come from from psychology, engineering, education or design, to cite the most historically relevant fields. Ultimately such properties can be incorporated into evaluation models to assist us in creativity assessment, or in commutable metrics which be used to identify and evaluate human or computer creative products.

The most cited properties in creative products are **Novelty** ([Saunders and Gero, 2001], [Boden, 2004b], [Wiggins, 2006], [Oman et al., 2013], [Goldenberg and Mazursky, 2002], [Cropley, 2005], [Maher, 2010] [Maher and Fisher, 2012], [Sarkar and Chakrabarti, 2011]) and product **Value** ([Boden, 2004b], [Wiggins, 2006], [Goldenberg and Mazursky, 2002], [Besemer and O'Quin, 1987], [Maher, 2010] [Maher and Fisher, 2012]) or similarly *quality* [Oman et al., 2013], or *usefulness* [Sarkar and Chakrabarti, 2011] *importance* ([Horn and Salvendy, 2006]), *originality* and *uniqueness* [Goldenberg and Mazursky, 2002].

Some other properties are less agreed upon, either because they rely on the receiver reaction or on aesthetic considerations. We must mention here **Surprise** ([Horn and Salvendy, 2006], [Maher, 2010], [Maher and Fisher, 2012], which could be considered a property of the receiver and not of the product itself [Wiggins, 2006] or incorporated in a broader definition of the novelty term ([Besemer and O'Quin, 1987] as *tipicality* [Ritchie, 2007]), *affect* is also a receiver response to the product [Horn and Salvendy, 2006]. As a commonly cited aesthetic property is *elegance* [Horn and Salvendy, 2006] [Cropley, 2005]. Similarly, *interest* has been proposed as a measurable characteristic of novel products [Saunders and Gero, 2001].

6.3.2 Creativity metrics

As a relevant example of creativity evaluation though the measure of different aspects of the ideation, we should mention the pioneering work by Jami J. Shah [Shah et al., 2000]. In the context of engineering design theory, the author provides metrics for measuring novelty, variety, quality and quantity in creative designs. The functional decomposition of design ideations and the relative weight and hierarchization of their functional attributes are assessed by experts in order to compute overall scores for each property. The numerical ratings obtained showed high correlation to participant's assessments on the relative creativity of the ideations evaluated.

Shah doesn't advocate for a summarizing creativity metric, though later Oman proposes a modified set of metrics in which Shah's novelty and quality scores are combined in a weighted sum to assess a global creativity score [Oman et al., 2013].

$$Creativity(C) = W_N * Novelty(N) * W_Q * Quality(Q) \tag{6.1}$$

where W_N and W_Q are, respectively, weights for novelty and quality assessed by experts in the field, and Novelty and Quality are computed according to Shah's metrics.

Some other authors opt for product-based formulations, for example Sarkar [Sarkar and Chakrabarti, 2011] suggests the metric

$$Creativity(C) = Novelty(N) * Usefulness(U) \tag{6.2}$$

while Maher [Maher, 2010] incorporates the surprise attribute into her metric

$$Creativity(C) = Novelty(N) * Surprise(S) * Value(V) \tag{6.3}$$

As we see, the tree methods agree on the importance on novelty to evaluate a creative artifact. Additionally, both quality, usefulness and value are socially constructed properties related to the adequacy of that artifact to previously established requirements. Maher incorporates and extensively develops the role of surprise in the assessment of creativity.

As Maher's model will be the reference model for our proposed creativity metric, we will finish this introduction on computational creativity assessment with a more detailed overview of former methods for novelty and surprise computation respectively, to put in context Maher's contributions.

Finally, the last section will provide a more detailed explanation of the methodologies developed to measure such attributes in the context of the aforementioned Maher's model.

Novelty metrics

Novelty is a quality of being new, specially unusually new. Something is considered novel when it is significantly different from already existing products of the same class. Novel products, ideas or states are identified as *interestingly different*, because of our innate novelty-seeking behavior. Novelty is, for cognitive psychologists, one of the primary measures of the ability to be creative, as has been recognized to have a higher impact in creativity assessment compared to other attributes [Sarkar and Chakrabarti, 2011].

To illustrate how we can quantitatively evaluate novelty for creative ideations, the aforementioned metric provided by Shah [Shah et al., 2000] will be detailed next. The author assumes ideas have a hierarchical structure and novelty should be assessed at different stages. Then, and a pool of ideas are evaluated together to obtain relative ratings. For single stage ideations, an overall novelty score is computed as

$$M_i = \sum_{j=1}^{n} f_j S_j \tag{6.4}$$

with f_j being the weight and S_j being the novelty for a particular key attribute or function j, which is assigned an *a priori* value based on experts, or calculated as

$$S_j = \frac{T_j - C_j}{T_j} \tag{6.5}$$

where T_j and C_j are the total number of attribute classes in all the products evaluated (the author provides examples for categorical attributes) and the total number of occurrences of that particular category in all products.

Note this is a *frequentist* approach, because novelty ratings are computed according to the relative count of attribute occurrences.

Novelty assessment has been studied independently from the creative context as well, because of the intrinsic properties of novelty in data. Ultimately, some methodologies employed in novelty assessment in databases or time-series have been incorporated in the context of computational creativity, so it is worth briefly reviewing them.

When we seek not to classify data but to detect abnormalities, novelty detection may offer a plausible solution based on thresholding distance measures to normal data. This is, indeed, a one-class classification problem. Two synonym terms often employed when referring to this problem are *anomaly detection* and *outlier detection*.

Assessing novelty from data has been addressed by a number of researchers form varied domains such as signal processing,computer vision,pattern recognition or data mining, with applications in security, medical diagnosis, industrial monitoring or robotics to mention a few.

Pimentel [Pimentel et al., 2014] provides an up to date, exhaustive survey of recent studies and approaches to novelty detection, establishing five general categories: probabilistic, distance-based, reconstruction-based, domain-based and information-theoretic techniques.

In the context of music information retrieval, Novelty curves have been employed for structural analysis and segmentation of audio. The procedure, originally proposed by Foote [Foote, 2000], usually implies kernel correlation along the diagonal of self-similarity matrices. As this method strongly depends on the size of the kernel, some authors Kaiser [Kaiser and Peeters, 2013] and Lartillot [Lartillot and Cereghetti, 2013] propose multiscale analysis to in order to detect novelty at different levels (or conversely, employ novelty for hierarchical segmentation).

Surprise metrics

Surprise is the feeling caused by something unexpected, or unusual, or both. There is an overlap in the scientific literature concerning the terms novelty and surprise, because surprise is sometimes taken just as a synonymous of statistical rarity.

If *unusualness* can therefore be linked to our former definition of novelty, *unexpectancy* is more properly related to the notion of surprise. Surprise may be elicited by different types of user expectations, either in transformative or exploratory design [Brown, 2012].

Grace and Maher [Grace et al., 2015] develop a typology of expectation sources, defining four basic categories: *categorical*, *trend*, *relational* and *comprehension* structural based expectations. Categorical expectations are violated by dissimilarity to former products, thus they are triggered by novelty. Closely related, comprehension expectations are violated when the the new ideations challenge the structure of our conceptual space. By the other hand, relational expectations are built by inferring correlations between attributes or trends in attribute values over time.

Assessing surprise usually involves statistical criteria to measure expectancy. By assuming that new data changes the observer's distribution of beliefs (as in categorical and comprehension based expectations as mentioned before), surprising events may are considered to be those with a relatively low probability to happen.

The first statistical measures of surprise seem to have been proposed by Weaver in 1948. The Weaver's surprise index[Weaver, 1948] compares the probability of observing the event to the sum of the probabilities of observing all other possible results.

$$S.I._{obs} = \frac{\sum_i f_i^2}{f_{obs}} \tag{6.6}$$

Despite its simplicity, Weaver's Surprise Index allows to distinguish mere statistically rare events from truly unexpected, *surprising* events. Surprising events are those which are rare compared with the average probability of the other events (small denominator and big numerator), while rare events would not be considered as surprising (small denominator *and* small numerator).

More recently, Itti and Baldi [Itti and Baldi, 2009] propose a theoretical framework to assess surprise with a Bayesian approach. Their model defines surprise as the distance between prior and posterior probability distributions according to the Bayes theorem

$$P(M|D) = \frac{P(D|M)}{P(D)} P(M) \tag{6.7}$$

with $P(M)$ being the prior probability distribution for the model M and $P(M|D)$ the posterior probability of M conditioned to new data D. The authors suggest computing the surprise by using a distance measure such as the Kullback-Leibler (KL) divergence, which for continuous distributions is defined as

$$KL(P(M|D), P(M)) = \int_M P(M|D) \log \frac{P(M|D)}{P(M)} dM \tag{6.8}$$

As a practical implementation which exemplifies the use of Bayesian analysis for surprise assessment, Horvitz et al [Horvitz et al., 2005] propose a model of surprise in the context of traffic forecast. The authors develop a temporal, probabilistic model of traffic according to several variables. After training the model, when the likelihood of a new event taking place is considered low enough it is assumed as surprising and an alert message is displayed.

6.3.3 Maher model for Creativity Metrics

In the context of computational creativity, Maher [Maher, 2010],[Maher, 2012] proposes an entity-agnostic, AI based novelty measure for creative designs in which *novelty*, *value* and *surprise* are taken into account as a sort of *Turing test* for creative agents or cognitive assistants. As Maher model is the basis of the Creativity model proposed in this Thesis, we will discuss it with further detail.

Her methodology was originally based on unsupervised clustering techniques. The author defines a metric for creative products as

$$E(a_i) = f(N(a_i) * V(a_i) * S(a_i)) \tag{6.9}$$

where a_i is a new potentially creative artifact, and is assumed to be creative if $E(ai) > 0$. N,V and S respectively measure Novelty, Value and Surprise of a_i.

We will now discuss how Maher interprets those terms and how are them operationalized in her computational model. We must note, however, that Maher's methodology was further developed by the author and collaborators, notably Kazjon Grace[Maher et al., 2013],[Grace et al., 2015][Grace et al., 2015], consequently we must see her model as an open methodology which has been continuously reevaluated and refined over the last years.

Novelty is a measure of how different a new artifact is from already known artifacts in its domain. Such difference manifests as the presence of new attributes, or as an unused value for an existing attribute or combination of attributes. In Maher's model, Novelty is related to a distance metric measured from the novel product to existing ones.

Maher explores different clustering schemes, from iterative descent clustering methods like k-Means or SOM to hierarchical approaches [Grace and Maher, 2014]. The author reports some advantages in hierarchical clustering : namely, the fact that ideations often exhibit a multilevel structure which may be better elicited with this method, and the lack of assumptions concerning the number and scale of clusters.

In agglomerative clustering schemes, novelty may be evaluated as the distance from the novel product to the closest centroid in the conceptual map.

In a hierarchical clustering scheme, novelty is evaluated as the weighted sum of the distances between the new product and it cluster hierarchy, weighted proportionally to the depth of the tree to impart higher novelty ratings for greater distances in a higher hierarchic level.

Value tries to comparatively measure several desirable properties of the creative product, generally related to its performance and stated requirements or social acceptance. Obviously such measuring criteria is context dependent, and should adapt to the creative process, as new products may introduce new desirable attributes that will further modify the metrics for value. Such measures closely resemble traditional assessment techniques for product effectiveness in product design.

Value is evaluated according to some established criteria related to the usefulness, performance or attractiveness of the design. Then it is computed similarly by clusterizing the value attributes and determining the distance to the closest cluster. As such, value computation may detect designs with a novel, more unique value, however it cannot infer whether its attributes might indicate a *better*value. It may serve, however, to reject new designs without a relevant difference in their value compared to existing designs.

Surprise deals with our expectations of how a new product could be. A surprising product entails novelty but also unexpectancy. That is, a surprising product does generally assumes novelty, but a novel product may not be surprising at all. Like Novelty, Surprise may be seen as a function of the new product attributes compared to previously ones, but also should measure how much a new product unexpectedly departs from the projected future values.

In [Maher, 2012] Maher proposes that a new design is surprising when the previously mentioned clustering algorithm would place it in a new cluster alone, thus changing expectations for future products.

Under such assumption, surprise is a binary attribute. Expectations in this model are implicitly modeled with the cluster structure, as future products are expected to belong to preexisting clusters. As new single-element clusters are promoted by outliers, this method mostly models categorical expectations, and is just triggered by events high high novelty values.

In a later paper [Grace et al., 2015] the author proposes a similarity metric between the cluster before and after the addition of the new product. Such distance indicates the level of perturbation it caused to the cluster structure. For the same reason as before, the method is expected to show correlation with the novelty metric, but it does allow for a finer granularity and may provide better insight into comprehension-based (structural) expectations[Grace and Maher, 2014]

Notice that the previous approaches do not explicitly include time to assess surprise, beyond the implicit transition before and after the incorporation of the new product. Indeed, it shows the same lack of temporal dimension that Maher notices in Baldi's bayesian framework[Itti and Baldi, 2009].

As user expectations may be highly dependent on the sequential nature of products and inferences we can do about the next ones, it makes sense to incorporate the temporal evolution to model,either to model the expectation of one attribute value over time, or to build up expectations on mutual attribute relationships.

In a later paper, Maher at al[Maher et al., 2013] employ time-series prediction to evaluate surprise. The authors propose different regression based methods to model user expectations. The greater the distance of the new model parameters to the prediction given by the regression model, the greater the surprise. The first model proved useful detection of outliers, the second one could apparently better detect historical “trendsetters” and the last one incorporated a sort of memory decay, which the authors emphasize should be carefully chosen accordingly to the users experience.

Nonlinear regression methods were incorporated in [Grace and Maher, 2014] to assess expectation based on relational and temporal trends.

6.4 Musical Creativity

We devoted the former sections of this chapter to formally introduce Creativity, reviewing the main contributions in its definition from the fields of cognitive psychology and artificial intelligence. As a final step into formal models of creativity, we surveyed the most relevant mathematical and computational models proposed to identify creative products and creative ideas, describing with detail Maher’s creativity metric.

While such computational models seek to be as generic and discipline-agnostic as possible, their actual application to specific domains is still incipient. Despite the fact that research in creativity has recently achieved promising results in design science (see [Gero and Maher, 2013]), with a body of research on visual arts and literature as well, music performance is still an emerging area in creativity research.

The next two introductory sections will highlight the situated and distributed nature of musical creativity: from the social perspective, and within the group in musical ensemble performance respectively.

We will then review the research on musical creativity in performance and will summarize the proposed cognitive or empirical models of creative processes both for score-based music performance and for improvisation.

Some of the proposed models have been computationally implemented, while many more implementations of generative systems for automated expressive musical performance and improvisation have been proposed, though many of them . It is out of the scope of this thesis reviewing the whole plethora of such systems, therefore we will only provide a few relevant examples, those more clearly grounded on performance analysis research and cognitive models of musical performance.

Finally, we will briefly review the incipient research on computational models of musical creativity, which will serve as an introduction to our own proposed model, to be fully developed in the following chapter.

6.4.1 The social context in musical creativity

Assessing creativity in music Clarke [Clarke, 2005b] puts Csikszentmihalyi's and Boden's generic models of creativity in the context of musical creativity, and advocates once again for the need or a more physically and socially informed perspective of musical creativity.

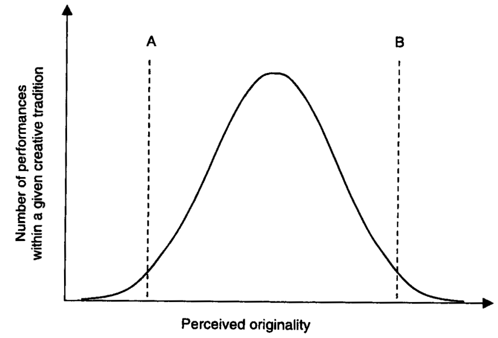
Indeed, the prospect for creativity assessment in art is not easy, as the intrinsic qualities commonly associated to a creative work (such as novelty or originality) and its social value are in fact deeply interrelated.

A temptative model of such interdependence is provided by Williamon [Williamon et al., 2006]. His model of creativity in music performance, originality and value are analyzed under the scope of their social and cultural drives (see Figures 6.1). Both figures place emphasis on the *situatedness* in performative originality and value ratings. According to the first figure (6.1a), highest originality ratings should be assigned to rare performances (close to B) within an accepted creative tradition or to radical and exceptional ones which would break the rules and setting new performative paradigms (after B), while being ultimately rejected as acceptably creative if they are too far away from that tradition (before A) . The second figure (6.1b) shows the acceptable degree of originality for stylistically well informed performances. When performers defy conventions, controversy arises and judgments on their artistic value start to be more negative. The added variance ranges represent graphically how consensus on artistic value is greater in original but uncontroversial performances, decreasing with very idiosyncratic proposals, beyond which excessive originality turns out to be less acceptable and valued, and perceived as eccentric.

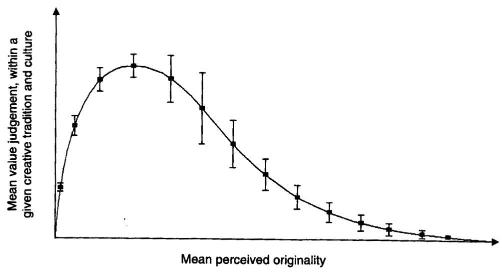
As the author points out, such analysis is only well suited to contexts with well-established performative traditions and clearly delimited stylistic boundaries, which are widely shared by their audience, such as the classical Western tradition. Accordingly, the lack of such stylistic constraints would implicitly make social driven evaluation much more problematic, as is the case in contemporary music genres.

6.4.2 Group creativity in musical ensembles

A musical ensemble constitutes a microsocial environment within which the aforementioned theories on collective and distributed creativity may be taken into consideration.



(a) A hypothetical distribution of perceived originality



(b) A hypothetical originality-value curve

Figure 6.1: Culturally mediated co-variation of originality and social value, redrawn for [Williamon et al., 2006].

Sawyer employs the term *Distributed creativity* to designate *situations where collaborating groups of individuals collectively generate a shared creative product* [Sawyer and DeZutter, 2009]. The author advocates for borrowing well-established methodologies from studies of distributed cognition to conduct real-time analysis of distributed creativity. His research on highly improvisational, unconstrained environments in theater, dance and free jazz show that the interactions among group members may be a source of creativity even more relevant than the individual inner processes of any participant.

Sawyer [Sawyer and Others, 2014] enumerates finally the key components of group creativity:

- **Process**: It has no external goals beyond the process itself: the process is the product.
- **Unpredictability**: As players choose among many possible consistent actions, the performance is essentially unpredictable.
- **Intersubjectivity**: To encourage collective flow, performer's contributions are open ended and their meaning is collectively assessed.

- **Emergence**: as a result of the former characteristics, a sort of group dynamics emerges and leads individual behaviors, a phenomenon that Sayers terms as *collaborative emergence*.

Group creativity in music ensembles has been addressed from different approaches and research disciplines, notably ethnography and conversation analysis. Most recent research focuses in the study of the interactional processes within musicians in performance (see, for example, [Hargreaves et al., 2012]).

6.4.3 Models of musical creativity

As we discussed in Chapter 1, we may identify three facets of musical creation: Composition, Performance and Improvisation, three closely related contexts with a notable degree of overlap and cross-fertilization between them.

Leaving aside the creative processes in musical composition, which are beyond the scope of this Thesis, we will focus on the study of musical creativity from the perspective of (score-based) music performance and improvisation.

For both contexts, cognitive and computational models of the creative processes involved have been proposed. As for music performance is concerned, such models focus on the characterization and of the expressive aspects of a rendition of the musical score. In the case of improvisation, research focuses in models which explain, and eventually implement the generative processes involved in the invention of musical materials in improvisational music genres.

We will first survey the contributions in the field of musical performance, followed by a more in-depth survey of the models of musical improvisation, finishing this section with a review of the generic computational models of musical creativity proposed in the literature.

Models of music performance

Cognitive models As Goebl states [Widmer and Goebl, 2004], the ultimate goal of performance analysis is to understand the relationships between the factors involved in music performance and formulate them as a model which generalizes the empirical findings and have both a descriptive and a predictive value. In regards to the outlook adopted by such generative models, Clarke observes that the majority of them are influenced by Chomsky's generative linguistics theory [Clarke, 2005a].

Clarke and Sloboda studied the issue of the cognitive processes involved in performance, improvisation and composition. Clarke, in particular, seeks to identify the characteristics generative principles for production and control of expressive aspects of performance [Clarke, 1988]. While Clarke does not intend to study the precise nature of such principles, he identifies a pattern of hierarchical mental representations or generative structures which part in an expressive performance: expressive gradients, discontinuities and contrasts. Such generative structures may be perfectly represented if all the information is available - as in performance of classical repertoire - , or incompletely represented - for example, when a performer is unfamiliar with a work. In the latter case, the generative representation is limited in depth and subject to further modification according to unfolding musical events. An extreme case of incomplete generative representation would be free improvisation, for which a performer has only a remote idea of how the overall shape of the piece will be realized.

Juslin [Juslin, 2003] proposes a psychologically grounded model of expressive music performance, the GERMS model, which identified five facets or factors which, together, contribute to the emotional and aesthetic impact of a given performance. Those factors are the Generative rules which parallel and clarify the musical structure, the Emotional expression used to convey emotional content to listeners, an unavoidable and unintentional Random variability, the Motion principles, which assume that performers intentionally re-create biological motor patterns and finally Stylistic unexpectedness, or unconventional variations in a music performance. A preliminary, rule-based implementation of the GERMS showed a good agreement between the different components and the predicted effects on listeners' ratings of the performance.

Computational models Models of expressive performance may ultimately be computationally implemented as expressive algorithmic performers. Such models may be of interest by themselves, as practical generative tools which enhance, humanize or otherwise add liveliness and realism to a software-based musical performance -indeed, there is even a contest on Performance Rendering Systems since 2002 [1]. They may also help to validate the proposed models, therefore contributing to the research of expressive music performance [Kirke and Miranda, 2009]. Bresin [Bresin and Friberg, 2013] provides an in-depth review of the methodologies for the evaluation of computer systems for expressive music performance.

The most prevalent modeling strategies resort to analysis-by-measurement, leading to empirically grounded models, and analysis-by-synthesis. Kirke and Miranda [Kirke and Miranda, 2009] review over 30 software systems for expressive music performance.

The KTH institute in Stockholm carried out a long term research on expressive performance modeling, based on actual performance measurements. Their Director Musices software [Sundberg et al., 1983] [Friberg, 1991], an ongoing project since 1982, incorporates over 30 user tweakable *rules* accounting for dynamics,

[1] `http://www.renconmusic.org/`

articulation, tempo but also harmonic context, phrasing and intonation rules and even ensemble tuning and synchronization. By fine-tuning the rules the user can model performances with different characters, either plausible or not. The team involved in the KTH performance model adapted their rule-system to the performance of contemporary atonal music [Friberg et al., 1991]. The DM model has been influential and similar rule systems have been incorporated into other models for expressive performance [Kirke and Miranda, 2009].

Todd proposed his Hierarchical Parabola Model [Todd, 1985] to model expressive timing in piano performances. It is a musicologically inspired model, resorting to a multilevel analysis of the score using Lerdahl and Jackendoff's Generative Theory of Tonal Music [Lerdahl and Jackendoff, 1985], and showed partial agreement with actual human performances of some classical music excerpts. On a later research [Todd, 1985], the author further developed a mathematical model of rubato by regression analysis of performances.

Models of musical improvisation

Cognitive models Several authors proposed cognitive models to explain the creative process involved in musical improvisation, notably in jazz. In the context of AI models for musical creativity modeling, several authors hypothesized cognitive models for improvisation which adopt a strong computational approach.

Pressing [Pressing, 1987] describes improvisation as a sequence of non-overlapping sections which the author calls event clusters. Those clusters are constituted by musically significant units such as melodic phrases, performative gestures, harmonic sequences. Musical events constitute cognitive or perceptual unities, and have some features which describe them as well as their similarity to other events, and are chained or *continued* according to process of changes of such features over time. Two methods of continuation are possible: associative generation and interrupt generation. With associative generation, a given musical direction is established and objects form a sequence in which features evolve according to the former constraints to effect continuity. In interrupt generation, the musician desires to explore a different musical direction, thus he resets the former set of components or sets of features. Pressing sees the whole process as a sort of feedback closed-loop system, in which a performer sets up goals and continuously re-adapts according to the actual output. While this model was not based on strong empirical evidence, and it was not formalized enough to be implemented and tested, is was historically relevant in providing a plausible theoretical model to understand the process of improvisation.

Ramalho [Ramalho and Ganascia, 1994] proposes a method intended to overcome the limitations of pure random-oriented choices or deterministic rule-based methods in improvisation. His model is based on the notion of Potential ACTions (henceforth, PACTs) as intentional units. Once again, PACTs encode a musically significative unit, though in this case emphasis is given in its action-oriented character: there can be procedural PACTs (what to play) and property-setting PACs (how to play). PACTs, in a given stylistic context, could be understood as gestural idioms. The system incorporates a Musical memory as well, that is, an accumulated catalog of musical resources, or actual realizations of PACTs.

Howoritz [Horowitz, 1995] proposes a a theoretical model for analysis and generation of jazz improvisations. Horowitz model draws on Minsky's model of K-lines[Minsky, 1980], and is based on a hierarchical model in which musical concepts, their contours and expectations, are modeled as separate, context-sensitive agents, coordinated by a further structural agent. Thus Horowitz proposes improvisation deals with multiple and concurrent representations of musical knowledge. Goals and intentions spread activation of such concepts (indeed, short scripts for certain musical behaviors, similar to Pressing's events or Ramalho's PACTs) downward in the agent hierarchy, while the currently active ones in the environment spread upwards activation to their related concepts.

Johnson-Laird [Johnson-Laird, 2002] advocates for a cognitively inspired models of computational improvisation. He argues against two possible algorithms formerly employed to model creative behavior, namely neo-Darwinist and neo-Lamarkist. Neo-Darwinist algorithms consist of random generation of ideas and natural selection for filtering them, and are hardly plausible given the real-time constraints of an improvisation, besides, they are based on lack of guiding criteria, while musicians improvising are guided by local goals. By the other hand, neo-Lamarkian algorithms, in which the generation stage is already constrained and guided by formerly acquired experience may be much more efficient but, assuming that predefined criteria provides valid choices, the creative process finishes there and does not cope well with multiple-choice selection and evaluation. Accordingly, Johnson-Laird suggest a third, hybrid method to model creative behavior, as a compromise between both approaches. Focusing on standard jazz improvisation, the author makes a clear distinction between the generative models for melodic improvisation and for the composition or improvisation of chord sequences. Melodic improvisation, which in jazz performers make take place at the fastest speed, must obey a neo-Lamarkian model of largely automated viable possibilities and a rapid arbitrary choice, without resorting to working memory at all. On the contrary, chordal improvisation must combine a neo-Lamarkian stage to generate viable chord sequences with a slower, reflective evaluation stage to select the better according to neo-Darwinian filtering processes.

Computational models Computational models of music improvisation are closely related to Interactive Music Systems [Rowe, 1992] and Algorithmic Composition Systems [Cope, 2005], and are mainly aimed at style modeling. Many of such models rely on traditional machine-learning approach based on a training phase with actual (human) performances, followed by the construction of a predictive/generative model of that musical style. That methodology shares similitudes with many algorithmic composition techniques, with the added constraints of interactive, on-the-fly real-time musical creation.

Systems for computer improvisation do resort to a plethora of generative procedures, some incorporate generative grammars [Quick and Hudak, 2013], grammar induction [Kitani and Koike, 2010], genetic algorithms [Weinberg et al., 2008] or machine learning techniques, such as unsupervised deep belief nets [Bickerman et al., 2010] or variable Markov models [Assayag and Dubnov, 2004], to mention just a few. The former cited model was eventually extended to cope with the requirements of mixed jam sessions with humans and software improvisers [Dubnov and Assayag, 2005]

While the systems cited above are tailored for real-time realization of improvised solos or accompaniments, modeling ensemble improvisation may be addressed by multi-agent systems. For example, the aforementioned Ramalho's model for improvisation[Ramalho and Ganascia, 1994] was eventually implemented to model a single improviser alone or interacting with more software agents as in a jazz ensemble [Ramalho, 1999]. The improvisational processes were modeled with a series of software modules: a Context module, which encodes contextual information such as the chord grid and external events, a Perception module, which listens to the context and puts events in a short-term memory, a Composer module which chooses the next playable PACT according to the context and an Execution module which executes the PACT by sending it to a MIDI synthesizer and notifying the Perception module.

Computational models of musical creativity

To our knowledge, very few researchers resorted to computational models of creativity for music creativity assessment.

In his study on tools for supporting collaborative online music creation with audio databases, Roma [Roma, 2015] proposes a creativity metric to evaluate the creative potential of sound clips. The metric is actually based on Maher's metric for product-based creativity assessment, though limited to content-based audio novelty and social relevance (i.e. value) of the exchanged sound samples, based on the amount of downloads, comments and ratings of the sound clips employed. Formally

$$creativity\_score = novelty * value \tag{6.10}$$

where novelty is evaluated as the distance to the closest element in a KD-tree data structure and value is computed as

$$value = \frac{\frac{f(downloads,comments,ratings)}{\Delta t}}{\#audio\_clips} \tag{6.11}$$

The measures seem to confirm the common hypothesis that small-world network structures favor creativity and innovation.

While the former research employed a creativity metric to relate compositional materials to their social relevance, Pointeau [Pointeau, 2013] carried out a computational creativity evaluation of actual musical performances with the Reactable.

The Reactable[2] was originally conceived as a tabletop tangible user interface for co-located, collaborative music making, though since 2010 it is also available as a mobile version. By placing physical objects called tangibles on the table (or their graphical counterparts in the mobile version), players may incrementally build and operate on a virtual modular synthesizer.

The reactable community website[3] hosts an online database of configurations and performances done with the Reactable mobile version which users may exchange and reuse. Pointeau retrieved the whole set of logged performance data (and, if available, transcribed performances). Such performance logs allow to generate statistics on the *tangibles* employed in a performance, and the parameters utilized in each tangible.

Pointeau proposed a creativity metric to quantitative evaluate the performances, inspired in attributes commonly employed by cognitive psychologists in the assessment of *divergent thinking*: namely *originality* and *elaboration*. A divergent thinking score[4] was computed as the average of both attributes. Formally

$$divergent\_thinking\_score = \frac{originality + elaboration}{2} \tag{6.12}$$

where originality and elaboration were operationalized with respect to the use of tangibles and their parameters respectively

$$originality = \frac{\sum_{tangibles} \frac{\#tangibles\_in\_dB}{\#tangibles}}{\#tangibles} \tag{6.13}$$

and

$$elaboration = \frac{\sum_{tangibles} \frac{\#used\_parameters}{\#available\_parameters}}{\#tangibles} \tag{6.14}$$

By comparing the divergent thinking metric with the amount of reuse of *tables* (which hold configurations of tangibles), the author confirmed that people that

[2] http://reactable.com/

[3] http://community.reactable.com/community/

[4] Divergent thinking is usually considered as a good indicator of the *potencial for creativy*, though it does not equate it[Runco, 2008]

rely more on the reuse of musical materials are more creative. This result provides quantitative evidence, therefore, of Csikszentmihalyi's notion of distributed creativity [Csikszentmihalyi, 2014].

6.5 Conclusions

In this chapter we reviewed the literature on Creativity from the perspectives of cognitive and social psychology. Other more recent approaches, such as the recent advances in the research of cognitive processes in the Neuroscience field, have not been addressed because they are not so closely related to our subject.

Once reviewed the research in Creativity as a cognitive or social process, we centered our review in the research in creativity assessment: the criteria to evaluate the creativity of artifacts, products and ideations. The literature on the subject provided an extensive listing of attributes, but three particularly emerged and were analyzed in depth: novelty, surprise and (social) value.

We presented the development of computational creativity metrics around these three attributes, particularly Maher's creativity model, which will be the basis of our own computational model for musical creativity assessment.

In the next section, we surveyed the research on musical creativity. Striking similarities arose when comparing the computational models of creativity just reviewed to the theoretical and computational models of both score-based music performance and, more particularly, musical improvisation.

Such evidence encourages further research into the explicit applicability of state-of-the-art computational creativity models for the assessment of musical creativity.

In this respect, we reviewed two pioneering approaches: Roma computes a potential of creativity metric to relate resource novelty and its social use, while Pointeau analyzes performance logs and extracts information on how the available resources are utilized to compute a creativity metric, though he apparently does not take into account the actual transcription of the performances, and therefore ignores the sequential nature of the performative actions.

While both authors do not apply computational models of creativity in the context of performance analysis to its full extent, their research is a major step forward.

Based on their contributions and the theoretical models previously discussed, the following Chapter will present and develop our own proposal for a Creativity Metric for musical performance.

Chapter 7

A CREATIVITY METRIC FOR COMPUTER ENSEMBLE PERFORMANCE

7.1 Introduction

This chapter presents our approach for the computation of a Creativity Metric for music performance, specifically addressed to multi-user instruments in Computer Ensembles.

The metric is based on real-time performance data capture and analysis from actual performances within a Network Music ensemble, and its main goal is to quantitatively evaluate the creativity of performers in improvisational settings.

The model we present is based on the methodology proposed by Maher and discussed in the previous Chapter, extended and adapted to the domain of music performance.

It might appear adventurous to employ Maher's metric, which is conceived to assess creativity in product design, for evaluating music ensemble performances, in which the creativity process has a very distinct nature: a musical performance -specifically, an ensemble improvisation, which will be the focus of our study- is not a monolithic product, but instead a live process which unfolds in time as a complex network of creative contributions from all performers. We may, however, consider that musical performance as the sum of the creative ideations contributed by performers, and evaluate the creativity of the *conceptual space(s)* defined by such set of ideations.

In the next sections we will provide the reasoning behind this decision, and the following ones will be devoted to derive our proposed methodology to compute musical creativity.

In the fist introductory section, therefore, we will justify why we may consider a musical performance, and specifically a free improvisation as a sequence of creative ideations. We will then introduce the concept of gestural ideations and its implication in the context of improvised electroacoustic music, as our *unit of ideation* of study.

The next section will fully develop our computational method, providing a RoadMap of the whole process and a detailed view of all the steps required to compute our creativity metrics.

7.2 Creativity as gestural improvisation

The goal of this section is to present the performative context we will utilize to derive our Creativity metric: free electroacoustic improvisation.

It will present as well the guiding principles of our computational metric: the assumption of atomic units of ideations in a free improvisation, its assessment through the analysis of performative gestures and a final measure of the gestural efficacy required to convey gestural creativity to an equivalent audible outcome.

7.2.1 Musical improvisation as a sequence of ideations

As we discussed in Chapter 5, a performer's creative contribution may be evaluated through music performance analysis. The precise nature of such contributions will differ according to the musical context, but, in any case, it may be largely assessed through measurements of performance.

An expressive rendition of a musical work, for example, displays systematic deviations from the score. Performance measurements do show to what extent they align to the structure of the musical work. Creative ideations, in such a context, expressively reinforce phrase boundaries and stylistically informed musical gestures as depicted in the score.

Without the need to align to an existing score, musical improvisation allows performers to make creative ideations far more explicit. As we have previously discussed in Chapter 6, cognitive psychologists proposed models for improvisational genres such as jazz in which creativity manifests as a series of clearly delimited, non overlapping musical ideations, events or actions. Given the stylistic constraints of the particular genre, we saw in Chapter 5 how a jazz improvisation closely aligns to the harmonic and formal structure, and how it largely manifests as a concatenation of formulaic patterns or *gestural idioms*.

In a free improvisational context, deprived of stylistic constraints, such ideations manifests as *performative gestures* in which performers try to avoid stylistic references or musical quotes, grounding creative ideations to moment to moment decisions, without any structural references at all. In free form, *Collective Improv* practices, for example, such sequence of performative gestures are mostly well isolated and non overlapped because of the stricter requirements for collective awareness in a joint improvisation, and the explicit desire to avoid salient individual roles in it.

7.2.2 Performative gesture as *unit of ideation*

The evidence of well delimited performative gestures in free improvisation scenarios suggested us to use such gestures as the discrete units upon which we would evaluate the creativity of a musical performance. By measuring their relative similarity and degree of unexpectedness, we may infer the degree of novelty and surprise depicted by a performance *from performer's gestures only*.

Our creativity metrics will be therefore optimally suited to assess musical creativity in freely improvisational scenarios in which performers resort to performative gestures to achieve their own musical voices.

As a final observation regarding our atomicity assumption for musical gestures: we do not claim that any particular gesture segmentation may act as the only level of ideation. Creativity researchers differ on this point, and ultimately, creativity unfolds at multiple levels and in multiple perspectives in improvised music[Sawyer and DeZutter, 2009]. Our choice of atomic gestures as *unit of ideation* provides a sufficiently low level of ideation to collect a significant number of discrete ideations per performance, while offering at the same time an adequate level of expressiveness.

However, assuming that there are no compound gestures (performative gestures made of several isolated gestures) or ignoring the inner structure of a single gesture is obviously an oversimplification.

7.2.3 Musical creativity assessment: gesture or audio?

There are several reasons why we will focus on performer's actions and not on their audible outcomes to assess creativity. First, the mapping between gestures and audio are usually multidimensional, even in digital musical instruments (see Chapter 3). Consequently by evaluating performer's contribution from audio we would require inferring the gestural contribution associated to a given *timbrical gesture*, which is a wide, emergent research topic by itself (see for example [Fran et al., 2013],[Caramiaux et al., 2010b], [Nymoen et al., 2012], [Jules Françoise, Baptiste Caramiaux, 2

Second, in the context of multi-user instruments which we seek to explore, a performer is only fully responsible of his own physical gestures, but the final audible result may typically be the result of the collaboration from multiple performers, which would defeat any attempt to evaluate individual contributions.

Last, but not the least, there is not such dichotomy between performer's actions and musical result: when experiencing a performance we are actually *listening through performer's gestures* and *feeling the gestures through the sound*. This is an evidence supported by the Embodied Music Cognition framework, developed by Leman [Leman, 2010] and Godøy [Godøy, 2006]

Even in the reduced listening, quasi acousmatic experience of a computer performance such affirmation still holds: sounds keep gestural affordances which listeners may attribute to human gestures[Paine, 2009] [Andean, 2012]. For the extreme scenario of a pure acousmatic listening, Dennis Smalley similarly refers to *surrogate gestures* as those gestures inferred or imagined when listening to music without a known source, originated by our natural tendency to relate sounds to sources and causes, drawing from our former experience[Smalley, 1997].

Of course, the actual audible result *matters* for the performance to be valuated as creative. Therefore, once what we should call *gestural creativity* is assessed, we should take into account its actual audible result. Here we state that creativity manifests, conversely, as a succession of *gestural-sonorous objects*, employing Godøy's terminology [Godøy, 2006].

Therefore our proposal is to infer the degree of *gestural efficacy* in the performative process, a measurement of the potential such gestural creativity has to be actually understood by others and valued accordingly, in order to validate the actual musical creative value of gestural ideations. We argue that, in a performance setting, the more related are the sonic gestures and the performer's gestures, the more successfully performer's creativity may be acknowledged and consequently the more valuable may them be considered.

7.3 Developing a Musical Creativity Metric

The goal of this chapter is to present our approach for the computation of a Musical Creativity Metric from performance measurements. Figure 7.1 provides a summarized road-map of the whole process, outlining the required steps. In the following sections we will provide a detailed explanation for each of the steps and the decisions adopted.

7.3.1 Data capture

Music Performance research, as we surveyed in Chapter 5, heavily relies on measurements of performance to perform quantitative analysis. In this respect, the adoption of the MIDI protocol by electronic musical instruments, particularly the piano, paved the way for a wave of music performance research in the nineties. Comparatively, measurements of performance with other musical instruments were

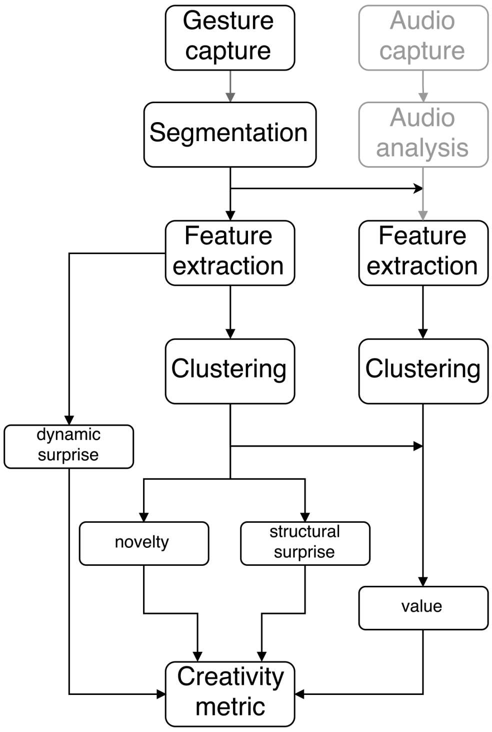


Figure 7.1: Block diagram for musical creativity computation

scarcer, largely due to the difficulties involved in the acquisition of performance data from such instruments.

Concerning computer music ensembles, our area of interest, retrieving performance data is greatly facilitated by the decoupled nature of digital musical instruments (see Chapter 3). In a similar way to a MIDI piano, computer-based musical instruments usually separate the gesture capture from the sound generation such that the manipulative gestures performed by the musician on the input device are already available for analysis, and are readily isolated from other non manipulative gestures (see Figure 7.2).

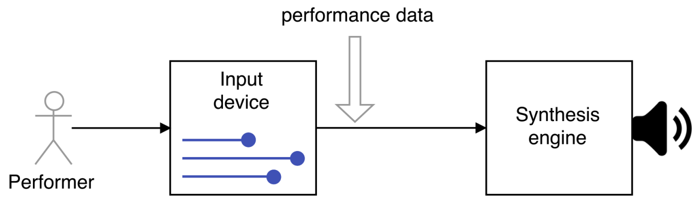


Figure 7.2: Performance data capture on a Digital Musical Instrument

We have seen in Chapter 3 that a number of computer music ensembles resort to a plethora of gestural interfaces for direct musical control, while some others make use of more standardized and readily available input interfaces such as keyboards and mouse pointers. In the latter case, a further mapping layer is often incorporated in the form of standard graphical control interfaces which allow for a precise control and monitoring of multiple parameters. Either way, performance data is often explicitly encoded as plain MIDI or Open Sound Control messages and may be easily logged.

Nowadays, Open Sound Control[Wright, 1997] is a *de facto* standard content format to convey such performance data to music software. It is a transport-independent protocol which allows data to be carried across different network technologies, such as Ethernet or Wifi based LANs, with flexible, user-defined namespaces and data types.

A typical OSC data stream consists of a sequence of OSC messages, which may be encapsulated into bundles incorporating NTP-based timestamps. An OSC message consist of a human readable label called address or path and a vector of primitive data types, whose type is explicitly indicated by a typetag. Table 7.1 shows a simple OSC message as captured from an actual performance, consisting of two integers and one floating point number. In this precise case it indicates that the performer 3 operated on slider 1 setting it to the position 0.23.

Unlike MIDI, the actual semantics of an OSC message is left to the designer, as well as the number and type of data contained. It is up to the interface designer as

Table 7.1: Structure of an OSC message

OSC message		
Address	Data	
	typetags	arguments
/blo/gen/perf	*iif*	*3,1,0.23*

well whether Open Sound Control messages incorporate high resolution time tags and whether they are sent at a fixed rate or asynchronously[Schmeder et al., 2010]. In the most generic and simple designs, OSC data is sent as soon as a change in a parameter is considered necessary to be reported.

Our performance analysis will be based on logging such OSC performance data. Upon selecting specific OSC paths, we may simply reconstruct the state of the original control interfaces or GUI controls by resampling the non-uniformly sampled OSC stream with zero-order hold interpolation (see Figure 7.3).

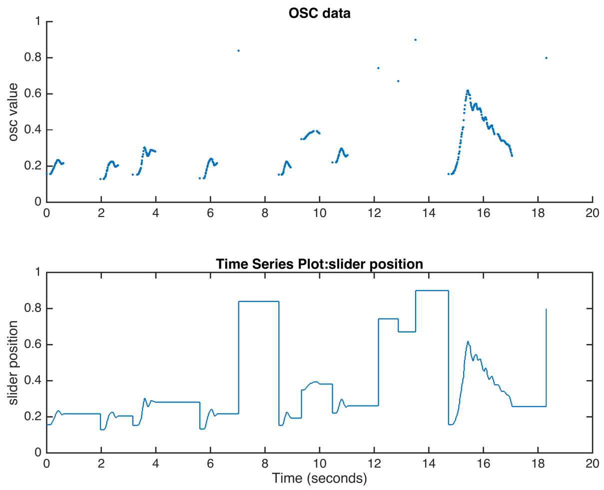


Figure 7.3: Captured OSC data and corresponding slider position

7.3.2 Segmentation

Once the performance data has been collected, and assuming that performers' ideations manifest as identifiable performance gestures -which is the basic assumption for our creativity metrics- we need a means to isolate such gestures from our continuous streams of OSC data.

From now on we will resort to unidimensional timeseries, which correspond to manipulative actions aimed at modifying a sound producing device either with a 1:1 or a 1:n mapping from gesture to sound parameters.

Beyond trivial equal spacing segmentation routines, or gestures clearly delimited by resting passages, segmenting gestural data meaningfully is a complex issue [Mitra and Acharya, 2007] because semantics of the gestures themselves should be taken into account in the segmentation procedure, and because of the hierarchical and multidimensional nature of gestural ideations in our context. As Kahol [Kahol and Rikakis, 2003] points out, gestural boundaries are entirely subjective and sequence dependent, i.e. the perceived gestural boundaries greatly depend on the ordering of the gestures themselves.

In order to simplify the procedure of gesture segmentation for our study, we may incorporate specific constrains in the control interfaces or in the musical task to perform, to enforce the production of an explicitly segmented performance. Those constrains may consist of turn taking routines, specific spatial and/or temporal constrains to the available control parameters or simply instruct players to perform with isolated musical gestures only.

While such constrains may severely impact on the generality of our conclusions, they allow the performer to clearly delimit their musical contributions in unambiguous gestural units. Besides, an explicitly well segmented performance *is* a good prospect for free improvisational practices as previously discussed.

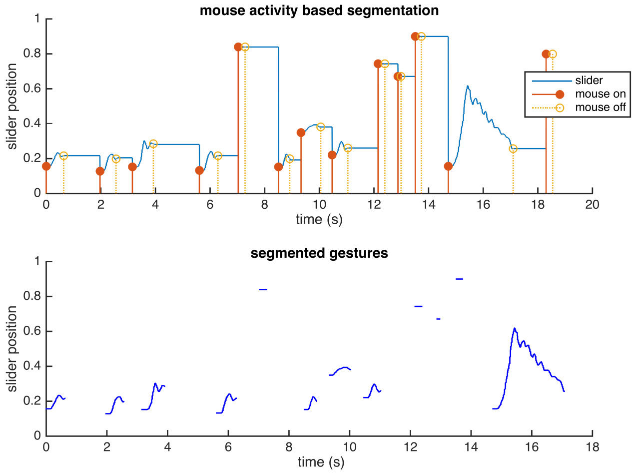


Figure 7.4: Segmented timeseries based on mouse state

As an example, Figure 7.4 shows gestural segmentation explicitly delimited by a user operating on a computer-based GUI for a sound-producing device. We simply employed mouse activity on the widget to segment slider data. In this case the control interface greatly facilitated isolating the meaningful gestural components, i.e. only the manipulative actions which actively operate on the sound-producing device.

7.3.3 Feature extraction

Once performing gestures are properly segmented, we need to extract relevant features to characterize them. That is, we will describe the morphology of the manipulative gestures exerted by a performer on a musical interface. Ultimately, the nature of such interface will determine what features may be the more relevant to extract.

With gestural interfaces, or even with full-body movement capture, it may be desirable to extract higher level gestural features with may be related to the expressive content conveyed by users through the combined movement of multiple limbs.

Hartmann et al. proposed a formal model for human gestures which has become a reference for research in gesture expressivity [Hartmann et al., 2005]. Human movements may be informally characterized with descriptive terms such as slow/fast, small/expansive, weak/energetic, small/large, and unpleasant/pleasant. Harmann summarized the most relevant attributes with six basic gesture descriptors: Overall activation, Spatial extent, Temporal extent, Fluidity, Power and Repetition.

Figure 7.2 summarizes the six gesture features proposed by Hartmann, some related terms and their definition. It should be emphasized that such gesture descriptors are often adapted to each particular context of study and formalized accordingly, therefore we did not attempt to provide a reference implementation.

Further gesture descriptors have been suggested in the literature, such as Kinetic Energy, Directedness, Impulsiveness or Symmetry (see for example [Piana et al., 2013]), many of which assume a three-dimensional space and are particularly oriented towards the analysis of human body movement, the relationship between bodily expression and emotion [Wallbott, 1998] and the analysis and eventually the synthesis of gestures for HCI research and body animation for expressive embodied conversational agents [Caridakis et al., 2006] [Bevacqua et al., 2007].

While such descriptors are suitable for summarizing highly dimensional data, for example when resorting to gestural interfaces for sound control, our studies focus instead on low dimensional interfaces, the ones we find as the computer native interfaces and in GUI-based control interfaces.

Table 7.2: The six basic gesture descriptors proposed by Hartmann et al.

Feature	Related terms	Definition	Dimension evaluated
Overall activation	Quantity of motion	Sum of motion	passive/static vs animated/engaged
Spatial extent	Contraction Index	Amplitude of movements	contraced vs wide, expansive
Temporal extent		Motion over a time period	slow/sustained vs fast
Fluidity	Smoothness	Continuity of movement	smooth, elegant vs sudden, jerky
Power	Energy		weak vs tense, powerful
Repetition		Presence of periodicities	

In this case simpler temporal features such as mean and dispersion for position, speed and possibly acceleration may be well enough to characterize simple gestures. Additionally, we may incorporate both the gesture duration and the time between consecutive gestures as two additional features, which as we have seen in Chapter 5, play an important role in the characterization of musical expressivity and coordination between players in a music ensemble. The final set of one-dimensional features which we will use and their meaning is summarized in Figure 7.3.

Finally, Figure 7.5 displays the analysis of the previously segmented perfor-

Table 7.3: A set of low level features for one-dimensional gesture analysis

feature	Definition
Duration	segment duration
IOI	inter-onset interval between segments
Position	Mean position
Range	Inter-quartile range. Related to spatial extent
Speed	Mean velocity. Related to temporal extent
Speed variation	Inter-quartile range for velocity. Related to fluidity

mance according to the selected set of low level features aforementioned. The value for each feature is averaged over the duration of the gesture and each feature vector is finally normalized, therefore each gesture will be represented with a single vector of 6 features.

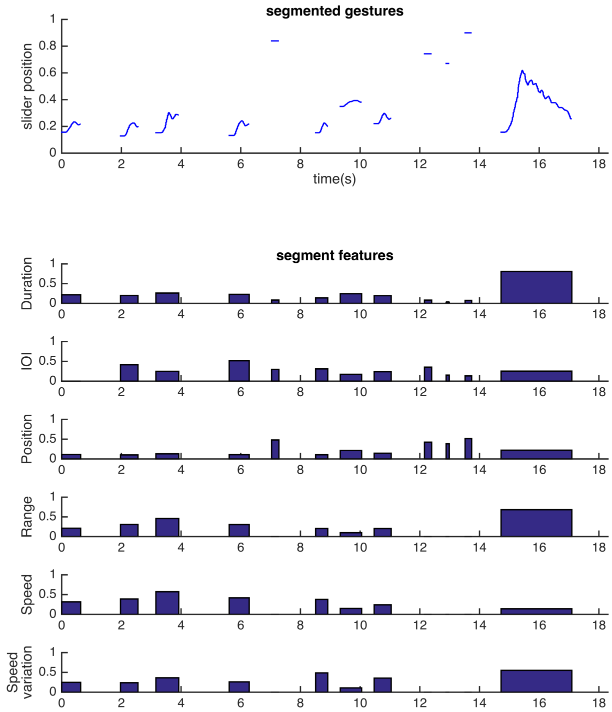


Figure 7.5: Mean feature value per segment, normalized

7.3.4 Clustering

The last step required in our methodology, before computing the creativity attributes for a musical performance, is to cluster the obtained performative gestures according to their similarity. The key assumption here is that in a creative performance, and specifically in an free improvisational context, such gestures are not expected to be neither uniformly monotone nor uniformly dissimilar. Instead, they unfold throughout the improvisation, either by processes of re-elaboration and variation of a given gesture, or by new and contrasting proposals.

This step is therefore required to compute both novelty, structural surprise and value, while dynamic surprise may be directly computed from the set of gesture features previously extracted. As the purpose of our clustering scheme is not to classify new gestures, but only to determine their relative similarity to former gestures as they are being produced, an unsupervised clustering scheme is adequate.

To cluster a set of creative products, Maher[Maher, 2010] suggests the use of k-means, possibly the simplest and most popular general-purpose partitional clustering algorithm[Jain, 2010]. With k-means, a set X of n-dimensional points $X = \{x_i, i = 1, ..., n\}$ is partitioned into K sets of points, $C = \{c_k, k = 1, ...K\}$with means $\mu_k, k = 1, ..., K$ such that the sum of the squared error over all K clusters

$$J(C_k)) = \sum_{k=1}^{K} \sum_{x_i \in c_k} \|x_i - \mu_k\|^2 \tag{7.1}$$

is minimized.

Each point x_i in the data set X corresponds to a single product, and its dimensionality n accounts for the number of attributes that represent it. In our case, each point would be associated to the set of features which represent a single performative gesture.

Being a multidimensional data set, dimensionality reduction techniques may be employed to represent the clusters graphically. Figure 7.6 illustrates this procedure, by performing Principal Component Analysis (PCA) on the clustered gesture set and then taking the two principal components to project the data set on a bidimensional plot. Each point in the figure corresponds to one of the gestures from the performance, grouped by similarity.

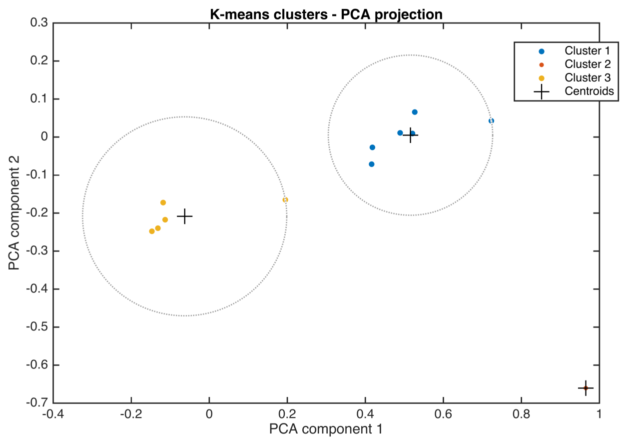


Figure 7.6: PCA projection for the K-means clustered set of gestures

K-means proved to be suitable enough to compute novelty and value metrics, because only a centroid model of the data was needed. It is an efficient algorithm which guarantees (local) convergence and as we will see the computation of creativity metrics with a centroid based clustering model is very simple. Even in a later research where hierarchical clustering algorithms were used to compute structural surprise, the authors still resorted to K-means clustering to compute novelty[Grace and Maher, 2014].

It has its drawbacks, however. A relevant disadvantage of k-means clustering is its assumption of separable, equal-sized spherical clusters with roughly equal number of points in it. A musical performance may consist of a relatively low number of expressive gestures from which we cannot take for granted equal variance and similar number of gestures per cluster as K-Means assumes.

Additionally, as we will incorporate a bayesian approach for structural surprise computation, a probabilistic model would be preferable in this case.

For the aforementioned reasons we will resort to Gaussian Mixture Models (GMM) to estimate the distribution of the data set. Gaussian Mixtures are multivariate distributions that consists of a mixture of multivariate Gaussian distributions (a generalization of the one-dimensional Gaussian distribution into multiple dimensions), each distribution being defined by a mean μ_k and a covariance Σ_k

$$f(x|\mu,\Sigma) = \frac{1}{(2\pi)^{D/2}\,|\Sigma|^{\frac{1}{2}}} e^{-\frac{(x-\mu)^T(x-\mu)}{2\Sigma}} \tag{7.2}$$

The mixture is defined as a weighted sum of K Gaussian components

$$p(x) = \sum_{k=1}^{K} \pi_k f(x|\mu_k, \Sigma_k) \tag{7.3}$$

Gaussian mixture models require, as with k-means, an *a-priori* assumption of the number of clusters. From them, an iterative Expectation-Maximization algorithm is used to maximize the log likelihood of the mixture model. Each point in the data set will be assigned a probability distribution instead of a hard cluster assignment. For a thorough discussion on GMM we refer the reader to [Bishop, 2006]

Figure 7.7 shows the computed density probability distributions for the same set of gestures segmented and analyzed in the previous sections. A two-component PCA projection was previously performed on the original descriptor set, with the first two components accounting for $> 95\%$ of the variance, while the number optimal of clusters was determined by visual and formal inspection of the silhouette as well as with the Davies–Bouldin index (see [Halkidi et al., 2001]). We employed diagonal instead of full covariance matrices because it is known that full covariance matrices are prone to over-fitting on small datasets [Magdon-Ismail and Purnell, 2010].

The distribution reasonably resembles the clusters obtained by the k-means algorithm.

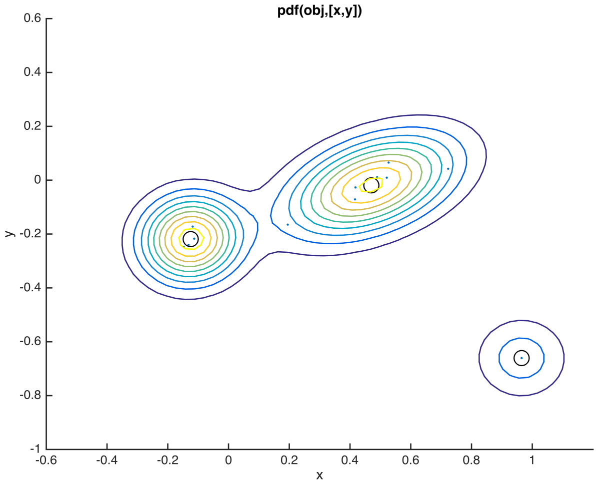


Figure 7.7: GMM clustering on the PCA projected descriptor set

From the probability distribution obtained with the Gaussian mixture model we may assign to each gesture a cluster corresponding to the mixture component with the highest posterior probability. As shown in Figure 7.8, the GMM based clustering correctly assigned distinct clusters to the three different gesture morphologies.

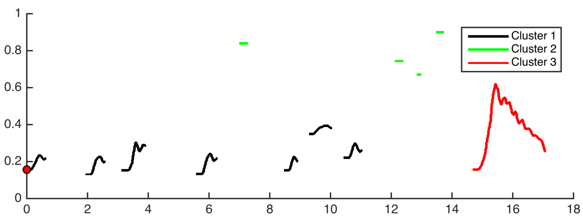


Figure 7.8: Clustered gestures according to the previous GMM clustering

An alternative procedure, as suggested in [Grace and Maher, 2014], would be to employ hierarchical clustering to analyze our data set. As the authors suggest, hierarchical clustering (see Figure 7.7) is expected to better match the multilevel nature of creative artifacts. Moreover, it does not require to make any assumptions

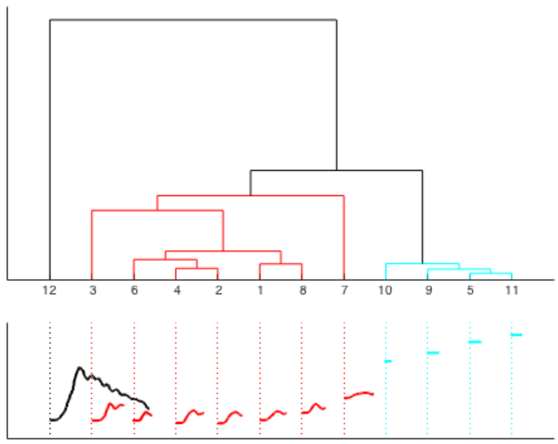


Figure 7.9: Hierarchical clustering of gestures. Notice the gestures have been reordered according to their similarity

about the number of clusters as in the former k-means and GMM based clustering methods.

We will not employ hierarchical clustering in our metrics for creativity computation, because both novelty and structural surprise assessment from hierarchical trees has been notably less explored in the literature surveyed. It is however, a promising line for further research.

7.3.5 Creativity metrics

The following sections will discuss in detail the procedure employed to compute the novelty, surprise and value for a musical performance, in order to provide a final summarizing creativity metric.

We will consider the analysis of performative gestures on a single performer basis, and later on we will discuss possible approaches to compute creativity metrics for the whole ensemble.

Novelty

To compute the novelty of a performance, we will compute the novelty value for each of the performative gestures just analyzed.

As we discussed in Chapter 6, in a k-means based clustering novelty is usually computed as the euclidean distance to the closest centroid. For example, Figure

7.10 maps a set of gestures according to their two most relevant attributes. The ideation labeled A, which is assumed to be the last one, is evaluated against the existing pool of ideations (encircled) to infer its creativity metrics. In this case, for example, if we assume the attributes define characteristics of the artifacts, A would be considered highly novel compared to former artifacts in the conceptual space.

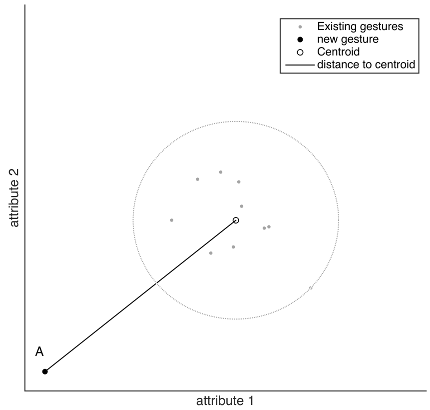


Figure 7.10: A new (and novel) gesture

As a performance develops and evolves through time, novelty evaluation must proceed similarly. Performative gestures are novel in reference to the previous ones (by the same performer as we focus on personal creativity by now).

We may accordingly evaluate the novelty of each gesture in such space either to analyze its individual creativity attributes, to study the evolution of such creativity attributes over the sequence of gestures or to infer a measure of the overall creativity involved in the generation of the whole gesture set. Such evaluation depends always on the order in which gestures are incorporated, because every new gesture is evaluated against the set of previously existing gestures.

Figure 7.11 illustrates this point: assuming an existing space defined by five gestures represented as a single data cluster, for each new gesture (labeled from 1 to 5), we calculate its distance to the cluster's centroid to assess its novelty. From all the collected distances we may compute a single, summarizing value which describes the overall novelty for the sequence of gestures performed so far.

We will compare now three distinct spaces to see how the novelty metric is affected by the distribution of ideations. Figure 7.12 displays, from to to bottom, a conceptual space with a single cluster with low variance, a conceptual space with two clusters with the same low variance and a conceptual space with two

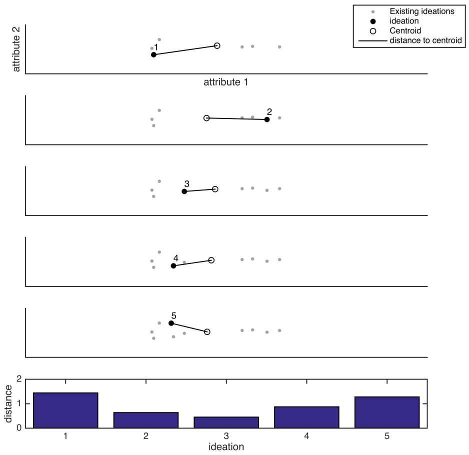


Figure 7.11: Computing the novelty for a sequence of new gestures

clusters with a larger variance. In the second and third spaces ideations have been incorporated on a cluster basis, so the 10th ideation will correspond to a change in cluster.

The second space displays a higher mean novelty compared to the first mainly because of the higher distance of the 10^{th} gesture to the former cluster. Once the new cluster has been established, the mean distance is similar. On the other side, the last space exhibits higher overall ratings because of its higher within-cluster variability.

The former examples showed how to compute the novelty of a sequence of gestures clustered with the k-means algorithm, and equating novelty to the euclidean distance from each new gesture to the closest existing centroid.

In our GMM based clustering scheme, novelty may be computed by equivalently measuring the distance to the closest component mean, giving virtually identical results (see Figure 7.13, second column).

Alternatively, novelty might be computed as the shortest Mahalanobis distance to the mean of each of the k components of the Gaussian mixture distribution. The Mahalanobis distance from a point x to the Gaussian component with mean μ and covariance matrix Σ is defined as

$d_m(x) = \sqrt{(x-\mu)^T \Sigma^{-1} * (x-\mu)}$

and has some interesting properties in our context, such as its scale invariance. Unfortunately, it is very sensitive to outliers -indeed, the Mahalanobis distance has been used for outlier detection in multivariate distributions [Rousseeuw and Van Zomeren, 1990]. The net effect would be a novelty metric which mainly accounts for outliers, providing a too coarse granularity (see Figure 7.13, third column). Therefore, we will

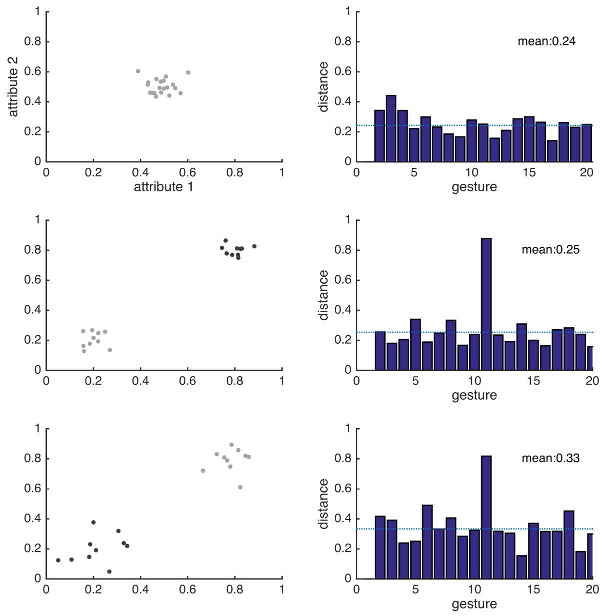


Figure 7.12: Individual and mean novelty for three distinct sequences of gestures, clustered using k-means.

keep using an Euclidian-based novelty metric.

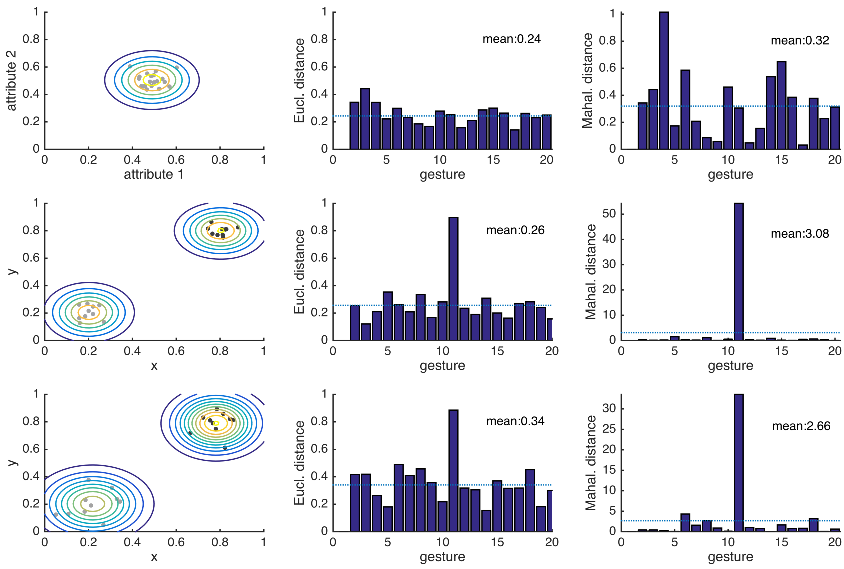


Figure 7.13: Individual and mean novelty for three distinct sequences of gestures, clustered using Gaussian Mixture Models. Left to right columns: probability distributions, novelty computed with Euclidean distance and novelty computed with Mahalanobis distance respectively.

Novelty for a whole performance may be computed as the average novelty for all the n gestures performed, as we have shown in Figures 7.12 and 7.13.

$$novelty = \frac{\sum_{i=1}^{n} novelty(i)}{n} \quad (7.4)$$

We have seen how a conceptual space is created as a sequence of ideations which may be iteratively evaluated to assess an overall creativity metric. The conceptual space may be employed to evaluate personal, group or historical creativity. Maher examples do focus in the later, we will similarly proceed to evaluate personal and group creativity.

Surprise

The conceptual framework employed to evaluate surprise is fully discussed by Grace and Maher in[Grace and Maher, 2014] and [Grace et al., 2015]. Among the four categories of expectations relevant for surprise assessment we will focus on category 4 (comprehensiveness expectations), which assesses *structural* surprise and category 2 (trend expectations) which assess *dynamic* surprise over time. Category 1 expectations are largely equivalent to novelty while category 3 (relational expectations) would be a specific dynamic surprise assessment which accounts for the evolution of more than one attribute over time.

Structural surprise We already discussed how structural surprise was originally assessed by Maher as a boolean attribute triggered by a redistribution of the clusters when the new event was incorporated into the conceptual space [Maher and Fisher, 2012]. Maher's proposal is in fact assigning (maximum) surprise to certain events with high novelty, as those events would eventually cause a reorganization of the cluster map. Further research (see [Grace and Maher, 2014]) show how surprise evaluation with hierarchical clusterings allow for a finer evaluation on the impact of a novel event in the conceptual map, though the values provided to weight each type of modification in the tree were rather arbitrary.

Another possibility could be to use a Bayesian approach to assess surprise, as described before [Itti and Baldi, 2009]. In our context, we must compute the probability densities before and after the new ideation, for example by fitting Gaussian distributions to the data. Once the prior and the posterior probability distributions have been computed, a suitable distance measure would indicate the degree of unexpectancy of the new ideation.

In order to incorporate surprise in our metric, it may be preferable to use the Hellinger distance instead of the Kullback-Leibler divergence which is used by the aforementioned authors, as it is a bounded metric, constrained to the interval (0,1). For continuous distributions, the Hellinger distance is defined as

$$H(P(M|D), P(M)) = \frac{1}{2}\int_M (\sqrt{P(M|D)} - \sqrt{P(M)})^2 dM \tag{7.5}$$

Such procedure would allow us to compute surprise as a continuous variable. Notice however than this measure of surprise neglects as well its sequential nature, but otherwise provides a statistically more robust measure of statistical *rarity*.

The following figure illustrates the process. When new data is assigned to existing clusters, their probability distribution is reevaluated and the Hellinger distance between both will be the measure of how surprising the new data is.

In situations when the new data is novel enough as to suggest to increase the number of clusters, we could compare the former map against the new map with an added cluster, or we might opt for keeping temporarily the same number of clusters and proceed with the Hellinger distance computation. Most possibly, keeping the number of clusters in this case will force anyway a notable reorganization of the cluster map and consequently a higher amount of surprise.

This second option, in a way, assumes that upon receiving new data, we first try to fit it in our preexisting conceptual map. Surprise arises from the fact that the new data is too unexpected as to seamlessly fit, and therefore asks for a re-organization of our beliefs. This initial *wow* effect is then resolved with a new conceptual map into which the surprising data is incorporated as another class of expected events.

To illustrate this, Figure 7.14 shows values for surprise computed immediately after the gestures B and C. Notice that surprise when gesture B has been received is very low, as the probability distributions after A and after B are very similar, as gestures A and B are, incidentally. However gesture C forces a cluster reorganization which results in a much higher surprise value, either by keeping the same number of clusters or by allowing for a new cluster centered at the surprising event (as in the last graph, labeled *Optimal clustering*).

This example shows how a sufficiently novel gesture exhibits higher surprise, but this is not always the case. If a gesture lays at a large distance from the closest centroid of a rather wide cluster, for example, it won't show such large degree of surprise, because the probability distribution won't be so drastically affected by the new gesture. In other words, this definition of surprise allows to distinguish *only rare* from *rare and surprising* events, in the spirit of the classic Weaver's Surprise Index [Weaver, 1948].

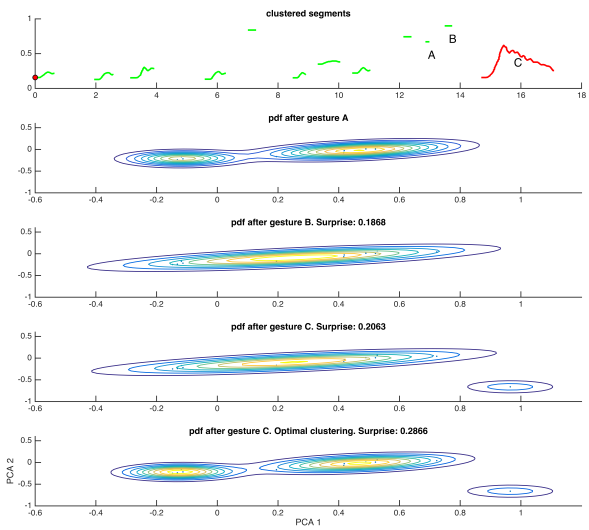


Figure 7.14: Surprise evaluation with Hellinger distance metric

As we will see, however, such structural surprise shows a high correlation with the novelty metrics, which confirms the notable overlap of both definitions.

Dynamic surprise As Maher notes out[Maher, 2010], surprise is inherently of sequential nature, and as such a pattern matching algorithm might be used to model listener expectations and the eventual mismatches that trigger surprise.

Surprise, therefore, may be as well caused by a hardly novel event, if it is perceived as unexpected. As an example, let's consider this sequence $e_i(n)$ of events e_i which are assigned to their respective clusters n

$$e_1(1), e_2(2), e_3(1), e_4(2), e_5(4), ... \quad (7.6)$$

Once a pattern in the cluster sequence is recognized, the last event will be considered as surprising no matter how much novelty might it convey.

As a first step towards a robust pattern based surprise assessment, we could incorporate some basic pattern matching routines in the cluster sequence, or maybe build simple regression models for it to detect trends.

But, beyond a sequence of clusters, even any trend over a particular sound attribute may trigger a surprise. This approach allows for finer detection of surprising events, even those which in our novelty metric would be assigned to the same cluster. Therefore, this will be the approach we will follow for dynamic surprise assessment.

We should emphasize that, while the methodology employed here is very similar to the one described in [Grace et al., 2015], we hare dealing with very different temporal scales: Grace studies the evolution of attributes in industrial designs to detect historical trends which span periods of several years, while we focus on the short-lived trends of a musical improvisation possibly spanning just a few minutes. The principle, however, is the same.

To detect temporal trends on a local scale, we perform a local regression or curve fitting for a sliding window of n events (as in the *third linear regression strategy* discussed by Maher [Maher et al., 2013]). The farther the next event n+1 will be of the predicted value according to the former n events, the most surprising it will be. However, we should take into account the local confidence of the prediction: only when a trend has been established with enough high confidence, an event outside the prediction bounds will be assumed to be surprising. Otherwise, the uncertainty of the prediction will prevent the event being considered as surprising.

We will illustrate both situations with an excerpt from a real improvisation (it corresponds to subject 1, scenario 1 from experiment 2), computing a local linear regression in a sliding window of 3 or 4 gestures. In Figure 7.15 we see a surprise triggered by an unexpectedly high mean value for the fourth gesture compared to the trend established by the former three gestures. Notably, this example shows how a high surprise can be triggered by a not so novel gesture.

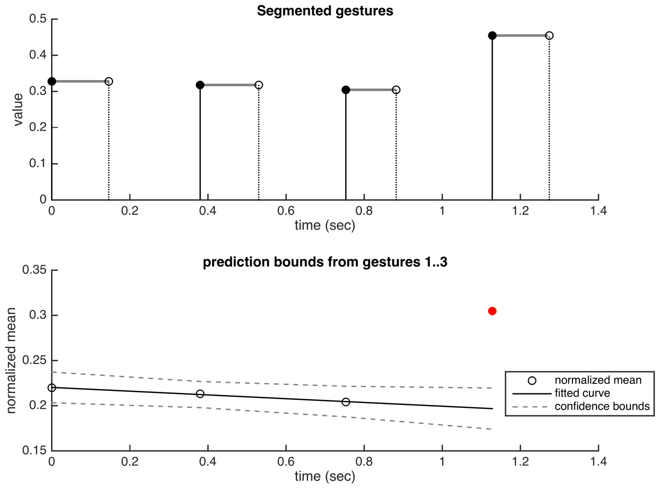


Figure 7.15: Dynamic surprise triggered by an unexpected mean gesture value. The solid red point in the second graph corresponds to the new surprising event.$CI = 0.95$ and $n = 3$ gestures

As a different example from the same musical excerpt, Figure 7.16 shows that the comparatively high inter-onset interval between gestures 4 and 5 does not trigger a surprise because of the large uncertainty of such a long term prediction.

The final dynamic surprise value for each gesture will be the maximum surprise value for all the attributes evaluated. Figure 7.17 shows how the final gesture-based dynamic surprise would be computed in the former example, assuming we only account for mean value and inter-onset interval.

The size of the window utilized to compute the local regression may have a significant effect on the surprise detector, an issue already discussed in Maher [Maher et al., 2013], who speculates about its possible application to better adapt to observers with different perceptions and memory. Figure 7.18 shows the same dynamic surprise metric as 7.17 but now using a window of $n = 4$ gestures for the regression. As we see, the surprise detector gives different results and in general is more insensitive now: some surprises were not triggered, while others showed a lower value.

In conclusion, a surprising event is an event which lies outside the prediction bounds, as computed by fitting a regression model to a local window of past events. Employing a sliding window effectively models short-term memory, so we assume listeners' expectations are mostly based on local trends. A surprising

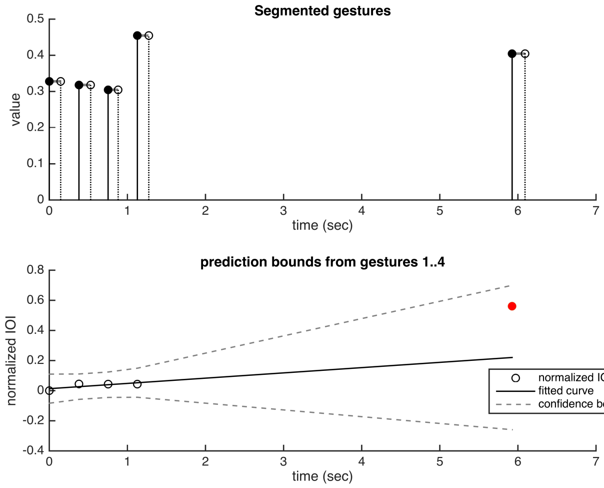


Figure 7.16: Dynamic surprise not triggered.The solid point in the second graph is far from the predicted value but still within the confidence bounds.$CI = 0.95$ and $n = 4$ gestures

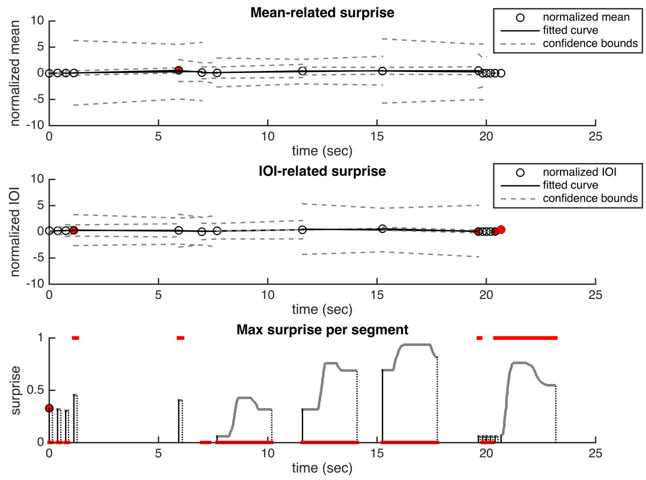


Figure 7.17: Dynamic surprise per segment, evaluated by assessing surprise over the trends established by two different gesture attributes. $CI = 0.95$ and $n = 3$ gestures

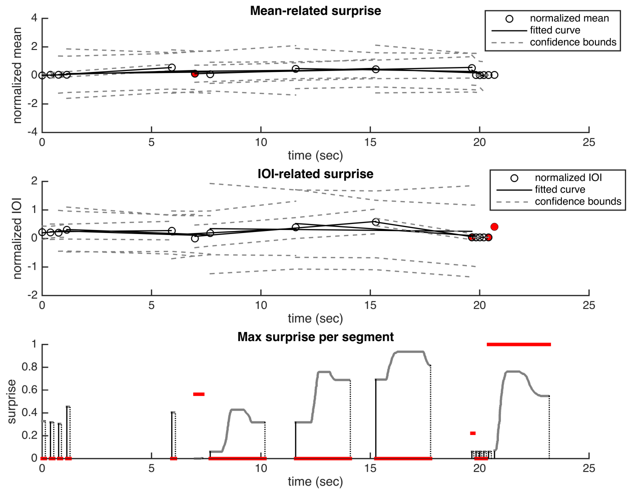


Figure 7.18: Dynamic surprise per segment, evaluated by assessing surprise over the trends established by two different gesture attributes. CI=0.95 and n=3

event is just an outlier according to local expectations, and it will be temporally incorporated in the model and subsequently forgotten. This sliding window therefore allows a similar sequence of events to re-trigger a surprise again once it has been forgotten.

The actual value for the surprise has no upper limit - a surprising event can always be more surprising just by moving farther from the expected value- therefore we arbitrarily set a normalized scale for the surprise metric as

- $Surprise = 0$ for any event within the prediction bounds
- $Surprise = 1$ for any event at a distance equal or greater to the prediction interval
- $0 < Surprise < 1$ for any event in between outside the prediction bounds but at a distance less than the prediction interval

As we have seen, surprise may arise by different factors. Given we have normalized metrics for all of them, we may just pick the highest observed surprise value to score each ideation, following Grace's methodology[Grace et al., 2015].

We proposed two expectation-based metrics: an holistic, atemporal measure of the impact of a new ideation into the preexisting conceptual map and a reductionist, temporal measure based on trend analysis for the attributes of gestures. We called the two metrics structural and dynamic surprise, respectively.

The dynamic surprise will for a given gesture i be just the highest surprise detected for all the k attributes evaluated

$$Dsurprise(i) = max(Dsurprise(i)..Dsurprise(k)) \tag{7.7}$$

therefore the final surprise value will be the maximum between the structural and the dynamic surprise values

$$surprise(i) = max(Ssurprise(i), Dsurprise(i)) \tag{7.8}$$

Surprise Finally, as with the novelty metric, the overall surprise metric for the whole performance will be the average surprise for all n gestures

$$surprise = \frac{\sum_{i=1}^{n} surprise(i)}{n} \tag{7.9}$$

Value

In goal-oriented designs, or in products which must fulfill a set of quantifiable requirements, social value can be assessed by evaluating their performance attributes with respect to industrial design requirements or social expectations.

This is hardly the case with non-utilitarian artifacts such as artistic products as a musical composition, which are not socially evaluated in those terms [Williamon et al., 2006] - except possibly those with strong utilitarian constraints. It is even more questionable to pretend to assign such attributes to a live performance: even if we could be able to monitor the impact of the performance on the audience, it would be an extremely situated and contingent criteria.

Consequently Maher's proposal, in which value is measured by defining a *product performance space* and a distance metric, *is clearly non applicable in our context*.

Note, however, that we are not evaluating the creativity of an audible musical outcome but the creativity of the performative actions involved. In this context, while we cannot reliably assign social-related value attributes to a performance, what we can do is to determine when a performance can hardly exhibit valuable contributions from a performer because the casual relationship between performer and sound is challenged. There are mainly two reasons why a creative gestural contribution might be hard to be valued by an external observer:

- Performative manipulations (what we call gestures) and/or their effects (sonic outcomes) are not fully revealed, leading to what Reeves call secretive, magical or suspenseful interfaces [Reeves et al., 2005].

- Manipulations and effects are revealed but they don't define the same conceptual spaces.

The first issue is mainly addressed at the design stage of the instrument. We discussed in Chapter 3, for example, what strategies do follow computer music ensembles to enhance the visibility of performers' actions though gestural interfaces or shared projection screens. As we cannot evaluate this issue from the performance measurements themselves, from now on we will assume that both gestures and their outcomes are adequately revealed to the audience. It is however a key consideration: an interface completely hiding performer's gestures might almost negate any creative value to a musical performance at all, as long as there would be no way for a listener to relate the musical outcome to the performer.

As for the second issue, a number of reasons may cause a mismatch between performative gestures and audible gestures. They are ultimately related to Fels' concept of *Transparency* [Fels et al., 2002]: an expressive interface heavily relies on the cognitive understanding and physical proficiency by the performer, and on the audience understanding, mainly acquired through the perception of physical causality relationships,

Let's illustrate this with an example. We discussed in Chapter 3 the role of the mapping layer in a DMI to adapt performative gestures into parameters for the sound generating device, which ultimately will impart a specific *gestural signature* to the sound. Apart from strictly linear mappings (of any dimensionality), which should only require scaling properly the gestures, any other more complex mapping do ask for a process of learning and adaptation for the performer to take full advantage of the musical possibilities offered by the interface.

Should the performer ignore the interface mappings, his performative gestures will lead to corresponding sonic gestures depicting a different creative process, in which conceptual clusters are seemingly unrelated to performative gestures. This mismatch might be even more exacerbated in scenarios of shared control, for example.

We can indeed quantitatively evaluate this relationship between the conceptual space defined by gestures and the one defined by their sonic outcomes, and this will be the basis of our value metric -or *value potential*, as we will often call it.

To summarize, our purpose is to compute a simple, holistic metric which measures the matching between gestural and sonic ideations, to evaluate how a performer efficiently translates *gesturally creative artifacts* into *sonically creative gestures*.

One possible procedure may therefore consist of the following steps -already presented in Figure 7.1:

1. Segment the audio according to the performative gestures.

2. Analyze the audio to extract timbrical descriptors.
3. Compute gestural descriptors for the aforementioned set of audio descriptors.
4. Cluster the audio segments in the space defined by their timbrical descriptors.
5. Measure the similarity between the gesture-based clustering and the audio-based clustering.

The higher such similarity, the higher the potential for the performance to be effectively valuable. In other words, the better the conceptual map set up by the gestures translates to an equivalent map of audible gestures.

The selection of timbrical descriptors is far from trivial. In the general case, with an a priory unknown mapping from gesture to audio, we might apply some dimensionality reduction techniques on a possibly large set of audio descriptors to focus the analysis in a set of attributes with the largest variance, or even better, to those that are possibly more correlated to performer's gestures. Several techniques have been proposed to identify the implicit mapping between gesture and sound, such as Hidden Markov Models[Fran et al., 2013] and [Caramiaux et al., 2010a] and Canonical correlation analysis [Caramiaux et al., 2010b].

If the mapping from gesture to timbre is explicit, this additional step may be omitted and we may focus on the relevant mapped timbrical parameters directly. This is the case when we map gesture to a *meaningful sound parameter* which already corresponds to a salient timbrical feature, as would be the case in a two or three-layer DMI mapping model [Hunt and Wanderley, 2003]. For example, an instrument could offer a control to modify the RMS gain, the frequency or the brightness of the sound, in which case we would not need to analyze the audio.

For illustrative purposes, we will resort to a very simple 1:1 mapping from a single control to a single timbrical parameter: a slider controlling the frequency of an oscillator. Figure 7.19 shows the timbrical gestures as frequency curves, given three distinct possible mappings: a linear mapping, a sigmoid compressed frequency range and another with a step quantized frequency curve. We might assume that they correspond to three distinct performances, playing with identical gestures but with instruments incorporating those three different mappings.

By visual inspection we can see that the original frequency mapping follows more closely the performer gestures than the second mapping, which completely flattens the low frequency gestures, while the quantized mapping distorts the gestures but still keeps a bit of the distinctive details from the short gestures.

From the frequency value we clusterize the timbrical space by employing the same descriptors we used to clusterize performative gestures, summarized in Table

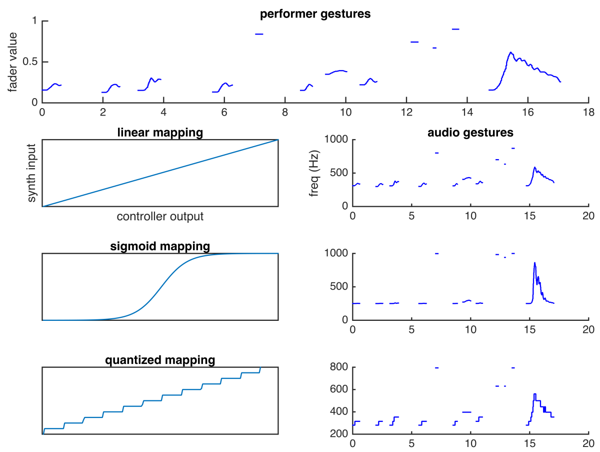


Figure 7.19: Audio gestures from three distinct parameter mappings

7.3. We will perform the same GMM based clustering scheme and then we will perform a hard clustering based on the closest component mean, as we did in the novelty computation stage.

Once we have both gesture-based and audio-based clusterings, we must evaluate their similarity. Given two (hard) clusterings $C = \{C_1, C2, ..., C_m\}$ and $D = \{D_1, D2, ..., D_m\}$, the similarity between C and D may be computed as the average Jaccard Similarity Coefficient between pair-wise cluster sets, as suggested in [Torres et al., 2008].

$$Sim(C, D) = \frac{\sum_{i<=m, j<=m} J(C_i, C_j)}{max(m, n)} \tag{7.10}$$

being the Jaccard coefficient for two sets A,B the percentage of differing elements over all different elements in two sets.

$$J(A, B) = \frac{\mid A \cap B \mid}{\mid A \cup B \mid} \tag{7.11}$$

We may now proceed with the evaluation of the value potential for the three hypothetical performances. Let's assume the performer used the same gestures for instruments incorporating three different mappings from gesture to the frequency value of the sound generator.

The results are shown in Figure 7.20. As expected, the linear mapping keeps a 1:1 correspondence between gestural clusters and audio clusters, and its value potential is consequently 1. The sigmoid mapping has the lowest value potential

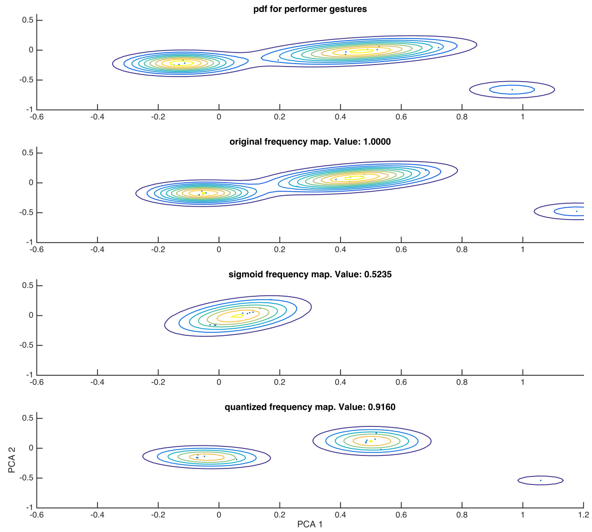


Figure 7.20: Gesture to audio clustering similarity with three different mappings

(0.52) because the subtle gestures performed on the low zone of the slider were mapped to a single frequency value, fusing the two clusters -which accounted for small differences between the gestures- into one. In a way this particular mapping is partially hiding the effects of gestures, otherwise we might say that the performer was not taking into account the characteristics of the mapping while performing, elaborating on subtle differences which were eventually negated by the audible outcome. The quantized mapping, on the other side, only partially removes the detail in the performative gestures when mapping them to frequency: possibly only a few gestures have been misplaced to different clusters. Therefore the value potential is close to 1.

This metric tries to be mapping agnostic, though it is optimally suited to *deterministic* and *static* mappings (see [Arfib et al., 2002] for a quick summary of mapping categories). We should emphasize that *we are not advocating for simple and/or linear mappings* to assess higher value potentials. Indeed complex, nonlinear mappings and nonlinear couplings are inherent in many acoustic instruments and actually several authors consider nonlinearity as an essential feature to achieve expressive control and expert gestures on a musical instrument [Hunt and Wanderley, 2003] [Jorda, 2005].

Higher value potentials are achieved when the performer takes into account the affordances provided by the instrument through such mappings from gestures to sound and, if this is the case, the intended grouping and evolution of expressive gestures should, to a large extend, be successfully projected into the audible outcome.

More significantly from the perspective of our research, mappings which exhibit some kind of non-deterministic or dynamic behavior -from the perspective of the performer- will inevitably display a lower value potential no matter what the performer does. One possible source of non-determinism is *concurrent actuation* that is, several performers operating on a shared parameter. This issue will have profound consequences in the evaluation of creativity metrics in multi-user instruments, as we will see in Chapter 9. Let's illustrate this fact with two more value potential assessments from the same performance, both adding a continuous sinusoidal drift in the performer's control value, the second one of higher frequency and greater amplitude (see Figure 7.21).

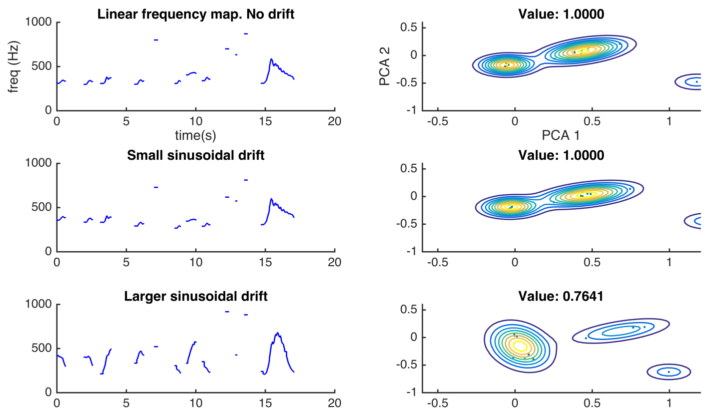


Figure 7.21: Gesture to audio clustering similarity with different amounts of sinusoidal drift on the mapped value

As expected, a sufficiently large drift has a noticeable impact in the value potential, while the smaller one does not alter the cluster assignations and therefore does not modify the value. Still, the temporal descriptors for our audio gestures are not affected in this scenario, therefore the third rating is still not as low as could be expected.

7.3.6 Performance creativity metric

We may now define the overall creativity metric, incorporating all the aforementioned considerations to display the evolution of the novelty, surprise and value ratings throughout the whole performance.

Figure 7.22 plots the ratings for novelty, averaged surprise and value potential for all the gestures in a performance, and their corresponding mean values. While it may be of interest to study creativity attributes on a gesture basis, in the experiments in Chapter 9 we will resort to mean values to globally compare different performances.

Notice that both novelty and the two surprise ratings are computed on a gesture basis: each surprise is evaluated against the former ones. The value rating, on the other hand, is computed from the beginning of the performance up to that current gesture, by comparing the clusterings defined by the performative gestures and audio gestures up to then.

Additional observations may be done regarding surprise and value potential. Both metrics depict a small number of discrete values for most of the timeseries. In the case of surprise, many events trigger maximum surprise and its value is therefore arbitrarily set to 1, as explained before. In the case of the value potential, it is a consequence of the Jaccard metric employed and the low number of clusters generated (typically less that 7 in this scenario), giving only a small number of possible values for this metric.

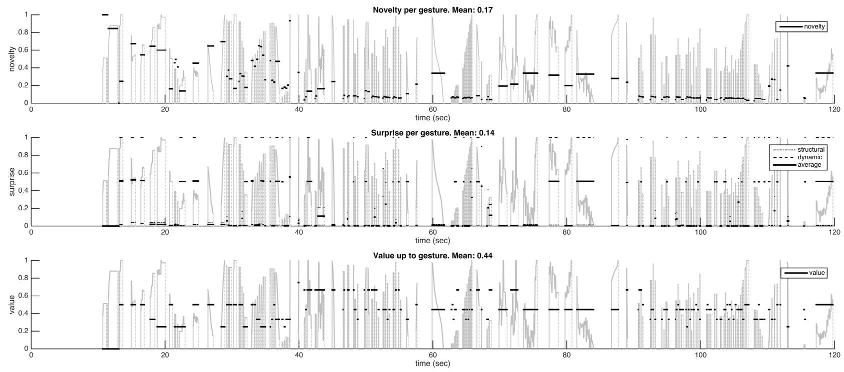


Figure 7.22: Creativity attributes for a whole performance. The data corresponds to Experiment 1, Group 2, Scenario 2.

A more useful visualization may be obtained by computing the running average and the cumulative average of creativity attributes. By comparing both values we may locally assess the creativity of a performer, detect periods of exceptional creative/uncreative activity and identify global trends. We may do so for each creativity attribute and for the whole summarizing creativity metric, which is simply computed as the product of the three attributes.

The final results are shown in Figure 7.23. We do not advocate, however, the use of a single summarizing creativity metric, because it provides little information on what factors contributed to its actual value and because it is yet to be assessed what would be the proper weighting, if any, for each one of the creativity attributes proposed.

Some global trends may be observed regarding the creativity attributes displayed in the last plot:

- **Novelty** shows a trend downwards after some initial time (less than one minute). This result implies that new gestures after that initial time begin to be more similar to former ones. This could effectively indicate a decrease in the novelty of newer gestures or could be a side effect of the computational procedure involved. Indeed, Grace et al. suggest de-trending the novelty timeseries by normalizing by its running average[Grace et al., 2014], though they work on a much larger temporal scale.

- **Structural surprise** shows the same global trend and local trends than novelty, confirming the correlation between both metrics.

- **Dynamic surprise** displays a high amount of activity. Triggered surprises account for any departure of well stated trends. We used a short window of $n = 3$ gestures which may make the surprise detector very sensitive to short term trends, possibly more related to processes of exploratory gesture variability than to purposefully set local trends. Longer windows would be more selective in this respect.

- The **Value potential** metric tends to converge to a fixed value. This makes sense as the sets of performative gestures and sonic gestures are expected to show a similarity which accounts for the adaptation of a performer to the instrument. With a static and linear mapping as it is the case, such similarity should tend to be rather invariable on a global scale.

- Finally, we may see that the **Creativity** metric shows the complex interplay between the three attributes involved. Highest creative gestures are only those in which novelty, surprise and value are highly rated.

Summarizing ensemble creativity

Evaluating ensemble creativity may be addressed from two perspectives. By one hand, we may compute the creativity for each performer, possibly adding or averaging performer's metrics as we did before to summarize their overall creativity. This approach does not take into account the emergence of creative behavior through social interaction between performers, as it ultimately equals to the measure of three isolated performances. But it may prove useful to assess and compare

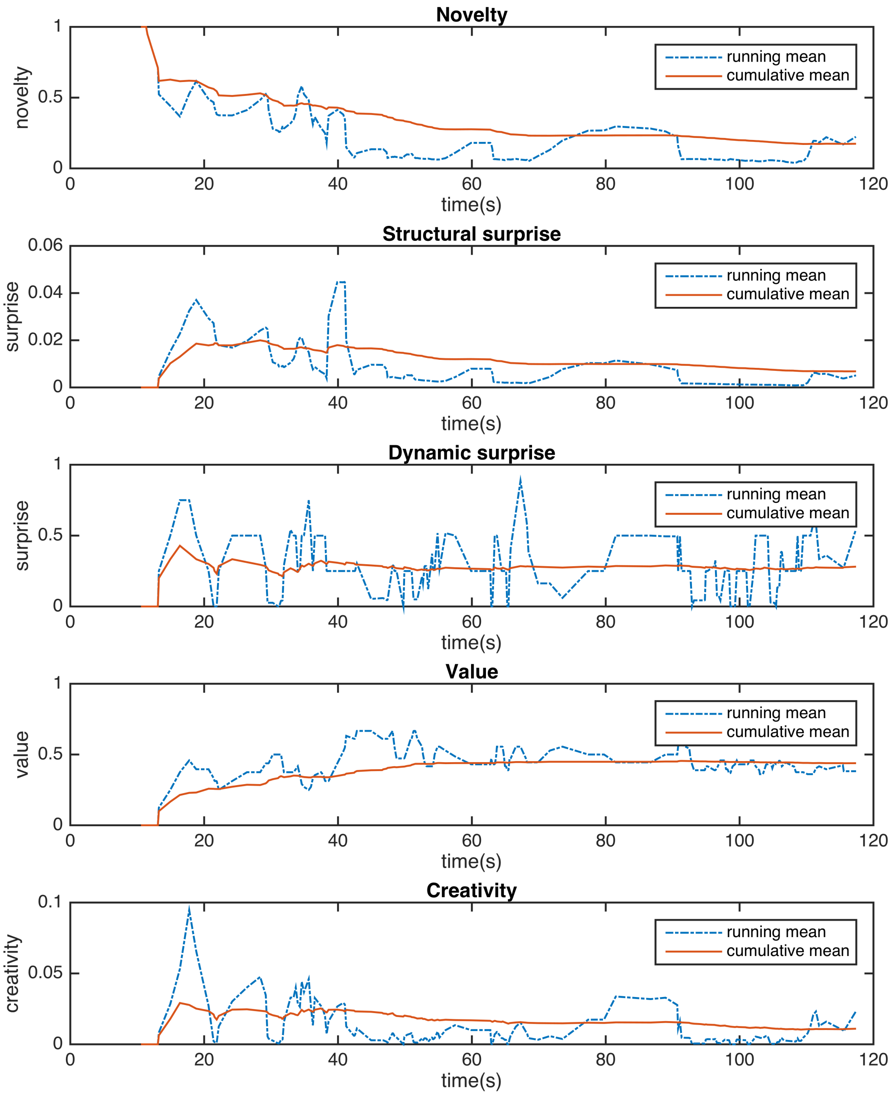


Figure 7.23: Cumulative and moving average for creativity attributes and creativity metric. The data corresponds to Experiment 1, Group 2, Scenario 2.

individual creativity in an isolated setting or in a group performance. The computation in this case resorts to the aforementioned methodology, extended to every member of the ensemble.

Alternatively, we may compute the overall creativity of all the contributed ideations regardless the performer involved in their production. This second perspective assumes that, in an ensemble performance, a shared ideational framework emerges and performers explore and transform together such global conceptual space. The methodology in this case is the same, but analyzing to the whole set of gestures performed by all the members of the ensemble.

While this is still a very crude approach, despite its simplicity it already takes into account the interplay between performers. In Chapter 9 we will provide individual and ensemble creativity measurements for the scenarios evaluated in our first experiment,

7.3.7 Elements for further discussion

Context and limitations The metrics developed are intended to be as generic as possible, but some assumptions had to be made in order to implement them in actual performative environments.

The main assumption is that performers perform *gesturally*. Their creative outcome is constrained to successions of raw instrumental gestures, unfolded through improvisational performances.

Additionally, our approach will involve constraining body motion to direct control of a low dimensional, even one-dimensional control interface as those natively provided by computer devices, similarly reducing the timbral complexity to a small number of salient timbral features, ideally those which more easily afford gestural responses [Godøy, 2006]. Next, as we deal with multi-user instruments, interdependence will be achieved in the form of shared controls or sound generating devices.

While this scenario may be easily extended to incorporate richer controllers and more complex mappings, it won't easily adapt to some multi-user instrumental paradigms which are commonplace in computer ensembles. Those include live-coding or instrument incorporating autonomous processes (be it live algorithmics, interactive systems or autonomous agents for example), the first because the additional layer of high level language commands defers any eventual gesturality to autonomous processes, the second because we assume that the musical outcome is a result of human decisions. Also, for obvious reasons, we exclude from this methodology the analysis of process-based or score-based musical repertoire, as well as explicitly conducted improvisations.

Gesture The atomicity assumption for performative gestures, as previously stated, is an oversimplification which does not take into account the complex relationships that may be established between gestural ideations. A first step towards a more realistic analysis of actual gestural improvisations which still resorts to the atomicity assumption could be employing hierarchical clustering to characterize the conceptual space defined by gestures. That would require to redefine both novelty and structural surprise detection algorithms. On this respect we may refer to the work of Kazjon Grace, who employed the COBWEB algorithm to build hierarchical conceptual clusterings and compute structural surprise with them [Grace et al., 2015] and Gerard Roma, who employed KD-trees to obtain novelty values from audio clips to be used in his own creativity metrics[Roma, 2015].

Novelty We discarded the Mahalanobis distance metric to measure novelty, but scale invariance might be relevant when assessing it. We should value as equally novel a small difference in a context of subtle variations than larger variability in a context of more varied ideations. Usually the conceptual space expands throughout the performance as new gestures are ideated, but when measuring the novelty of each ideation we do so against the former ones, not against an hypothetically larger conceptual space which at that moment does not yet manifests.

Related to that, we observed that for longer performances novelty decreases over time. It remains to verify whether this is a side effect of our computational methodology or performers effectively run out of ideas after playing for a while with a single slider -the environment we used to perform our preliminary computations. Another possibility would be to employ a sliding window to construct the conceptual space, effectively letting the model *forget* older ideations. Indeed, long term relationships should not be expected in free improvisational contexts, otherwise long term planning should be more relevant only when evaluating composed music. An additional benefit of a sliding window to construct the conceptual space would be its ability to better cope with dynamic mappings when assessing the value potential, for example.

This simple equation raises a second issue: the duration of gestures may have an unexpected impact on novelty evaluation. Leaving our definition of performance novelty score as it is, we will rate as more novel those performances with many short, novel ideations, while a single, highly novel but much longer ideation will just count as much as the shortest one, which should not be the case.

Ultimately, however, a minimum temporal span is required for an ideation to be identified, while much longer ideations are expected to display a hierarchical internal structure themselves which would defeat our atomicity assumption for gestures, or conversely require a multilevel analysis.

We might handle this situation by incorporating some heuristics, such as weighting each gestural novelty by the gesture duration or resorting to more sophisticated weighting factors to simulate the decay of attention and perceived novelty for longer ideations.

Surprise Both Maher's definition and our bayesian implementation for structural surprise suggest a high correlation with novelty. This evidence could ultimately restrict our surprise metric to dynamic surprise only.

Regarding dynamic surprise, we discussed the impact that the choice of the analysis window has in the detection of unexpectedness. Even if we could resort to multiscale analysis to detect surprise being triggered on different time scales, we believe that dynamic surprise should be based on more empirical evidence in order to determine the optimal temporal spans and a plausible relationship between pattern mismatch and surprise amount. In this respect, relating our metrics to the recent developments in cognitive Neuroscience, specifically on auditory scene modelling and auditory stream segregation, would surely provide a more reliable metric. We must cite here the experimental research carried out by Istvan Winkler on predictive processing and the mismatch negativity (MMN) to assess predictions based on representations of predictable patterns -see, for example, [Winkler et al., 2009][Winkler et al., 2012][Winkler and Czigler, 2012].

Value The value metrics, as it is, does not actually provide a rating for musical value of a performance, instead it tries to quantitatively assess the relationship between manipulative actions and audible outcomes, assuming that the more similar are the spaces defined by both, the more creative may be a performance acknowledged.

The Jaccard distance metric utilized to evaluate such similarity only compares the assignations of both manipulative and audible gestures into their respective clusters. A possibly less coarse evaluation could consist of computing both creativity and surprise metrics on the audio itself, and define the value in terms of their similarity to the metrics computed on the performative gestures.

It is however debatable to what extent perceiving a different organization between the gestures performed and the gestures listened may lead to a diminished sense of causality and therefore negatively impact on the value rating on a performance.

For example, complex mappings incorporating nonlinearity and even unpredictability may be a path to expressivity and virtuosity [Jordà, 2004b]. Moreover, unexpected and novel gesture to sound relationships are not to only regarded as less valuable, they may indeed be highly appreciated and considered more engaging by performers [Chadabe, 1997],[Jensenius, 2007]. It is worth noting that such perceptual mappings between gestures and sounds is highly cultural dependent as well. Gestures may not only suggest the nature of sounds but their instrumental identity: a sonic affordance may afford gestures of actual musical instruments [Tanaka et al., 2011] and vice-versa. To what extent the gestural and sonic affordances of our instrument may positively reinforce our expectations regarding a instrumental identity or, by the contrary, may frustrate them, would also contribute to the perceived value of a performance. Finally, as we restricted our metrics to the data obtained by performance measurements and such measurements usually deal with manipulative gestures only, we neglected embodied gestural cues which in a musical performance may contribute to a large degree to the semantics of performative gestures [Iazzetta, 2000].

For all the reasons above we suggest to restrict the proposed value potential metric to comparative studies in simplified and controlled environments only, as the ones we will employ in the next chapter to correlate value potential to interdependence. In the context of an actual performance, assessing value should be still expected to be inextricably linked to social evaluation.

Creativity Ultimately, there are more subtle dependencies between the three creativity attributes, beyond the aforementioned correlation between novelty and structural surprise, for example

- A very high novelty triggers a dynamic surprise because it departs from any eventual trend.

- But we may resort to slowly varying, low novel ideations to set up such a trend, therefore low novelty values may be previously required to elicit a surprise.

- Adapting the performative gestures to better match audible results and therefore achieve a higher value potential may consequently alter novelty and surprise ratings because the performer modifies the conceptual space defined by his gestures.

- An instrument which incorporates high unpredictability may show a low value potential, but such unexpected behavior may trigger more novel ideations by the performer.

This complex interplay between the creativity attributes suggests to analyze them separately instead of resorting to a single creativity metric. This is the approach we will follow in the following Chapter.

7.4 Conclusion

We presented our methodology to computationally assess creativity in music performance, with a specific focus in computer music performance. The model we presented is closely based on the methodology proposed by Maher and her collaborators, extended and adapted to the domain of music performance.

We first presented our performative context: free improvisation on digital musical instruments, consisting of a sequence of gestural ideations. Taking performative gestures as our unit of ideation, we provided a RoadMap of the whole process and a detailed view of each step required to compute the novelty, the surprise and the value potential for the sequence of gestures which constitute a musical performance.

In the next chapter we will test our creativity metrics on experiments specifically tailored to computer music ensembles. By analyzing the ratings of the different creativity attributes under a number of distinct performative scenarios, we expect to gain insight into the differences between individual and group creativity and the challenges and opportunities that multi-user instruments provide in terms of creative engagement.

Chapter 8

EXPERIMENTS

8.1 Introduction

In this Chapter we will present two Experiments which will incorporate the Creativity metrics presented in Chapter 8 7.3 to quantitatively evaluate improvisational performances with computer music ensembles, with a focus on multi-user instruments.

By analyzing the ratings of the different creativity attributes under a number of distinct performative scenarios, we expect to gain insight into the differences between individual and group creativity and the challenges and opportunities that multi-user instruments provide in terms of creative engagement.

We surveyed in Chapter 3 the research in this topic around Collaborative Musical Interfaces and multi-user instruments in Computer Music Ensembles, and preliminary hypothesis were drawn from our artistic research with the Barcelona Laptop Orchestra, as discussed in Chapter 4. This Chapter seeks to provide quantitative, experimental evidence to validate our hypothesis on the subject.

Our first attempts to reuse the performative environments designed by the Barcelona Laptop Orchestra in our experiments were discarded for a number of reasons. We will quickly justify why we did not find them suitable for our research:

- Score-based repertoire (as well as conducted scenarios) would require evaluating creativity in terms of deviations or ornamentations from the commands indicated in the score (or, respectively, by a conductor). This would translate to expressive performance analysis (surveyed in Chapter 5) and ask to evaluate creativity in reference to the constraints provided. Additionally, notational approaches in computer music repertoire go far beyond the paradigms of traditional Western Music Notation (see Chapter 3 for a survey of such approaches in the repertoire for Laptop Orchestras) and would ask for a creativity assessment exclusively tailored to a given piece.

- Our repertoire based on collective sequencing does not fit in the performative context developed to assess creativity, because performative actions do not consist of sequences of gestural ideations, but of discrete inputs which activate autonomous processes. Besides, we believe that the asynchronous nature of such environments make them less suitable to evaluate spontaneous, reactive improvisation but instead it better accounts for a mixture of improvisational and compositional procedures.

- Our two gestural interfaces are rather suitable to assess creativity. There are two issues to take into account though: the interfaces are multidimensional (up to 6 and 3 dimensions, respectively) and they explore discrete and discontinuous spaces, which in their turn may challenge our assumption that a performative gesture is entirely meaningful -the performer often just jumps between active zones and such transitional movements are not manipulative at all. Multidimensional gestures, by the other hand, should be entirely feasible to be analyzed in terms of their creativity attributes.

- Finally, some of out multi-user instruments partially fulfill the requirements to assess creativity from performance analysis. Particularly, both La Roda (the unconducted version) and the proxy-Hub, in which performers resort to low dimensional gestural manipulations, would be suitable. However both environments have evolved into complex instruments which for different reasons are not optimally suited for a primary evaluation of our metrics: in the case of La Roda, each performer faces a new sonic material which may become the primary reference to ideate new gestures, and not the former performer's ideations, challenging the study of collective creativity. Besides, its strictly circular turn-based scheme restricts short-time interaction to adjacent users. In the case of the proxy-Hub, there is an autonomous software process (the generative algorithm implemented in the server) which behaves as an additional actor in the performance, and performers mainly ornament concurrently that higher level shared process through their gestures. In both cases, but just for opposite reasons, a conversational interaction between performers, a desirable scenario for the study of group creativity, is therefore hard to achieve.

For the aforementioned reasons, we resorted to develop performative environments specifically tailored to the needs of each experiment. Some decisions had to be taken regarding the design of the required instruments:

- **Interdependence topology**: Design multi-user instruments which explore some relevant topologies of interdependence and use of shared resources which have been already addressed in the repertoire of Laptop Orchestras (surveyed in Chapter 3) and particularly in our own repertoire (surveyed in Chapter 4). The two particular axes analyzed refer to the amount of simultaneous access to a shared resource (Experiment 1) and to the kind and amount of shared resources provided (Experiment 2). Other multi-instrument related axes will not be studied and therefore will be kept unmodified in our experiments: both will be synchronous, centralized, exhibiting low physicality, medium-high homogeneity and assuming equal roles among performers.

- **Performative paradigm**: Our performative context should be aimed at fostering sustained individual and ensemble creativity, as much unconstrained as possible beyond the limitations inherent in our control interfaces. This requirement suggested addressing the performances as free improvisations, avoiding both the use of scores or *a priori* procedural schemes, getting rid of any kind of human or computer based conduction and trying to avoid stylistic references to familiar improvisational genres.

- **Architecture**: Resort to a centralized scheme to simplify the logging of performative gestures and audio generated. Our metric will be entirely based on performance measurements just as we have discussed in Chapter 4, and we will resort to OSC logs and, if required, audio captures.

- *Ecological validity*: Design the experiment to be as much similar as possible, if simplified, to an actual performance with a computer music ensemble. We hope that, by making the experiments closely resembling an actual performance,the results may be more significant and may shed light on challenges and opportunities encountered in real scenarios. By the other hand, we believe that designing instruments as if they were addressed to a real performance may provide a more positive attitude and ultimately will guarantee a more engaging and more creative experience.

In the following sections we will discuss the two controlled experiments conducted.

8.2 Experiment 1 : Creativity and interdependence

8.2.1 Introduction

A computer ensemble offers us an exceptionally well suited environment to investigate creativity in musical performance: measurements of performance are readily available and the performative environment may be flexibly reconfigured to suit the needs of a controlled experiment.

This experiment is focused on the study of musical creativity in a free improvisation with an instrument capable of partially sharing its control among several performers.

We will first introduce the aims of the experiment and the hypotheses we seek to confirm. We will next explain what experimental scenarios will be evaluated.

After carrying out the performances with each group of performers, we will perform both qualitative and quantitative analysis of the performance. Finally, we will compute the creativity attributes for every scenario and will analyze them comparatively.

8.2.2 Aims

This study seeks to get better understanding of the relationship between mutual interdependence and musical creativity through an analysis of performances with multi-user instruments.

The performative scenarios analyzed form a continuum from a purely set of isolated performers playing with their respective instruments to a collective performance with a single shared sound entity. Intermediate scenarios account for a joint performance with independent instruments and a joint performance with partially shared instruments respectively.

The experiments consist of short free improvisations with a set of Theremin-like instruments. Performers can only control the onset and offset and the frequency of such sound generators.

The goal of this study is to characterize ensemble performance in terms of performers' experience and their creative outcome, and the role of mutual interdependence in promoting or inhibiting creativity.

8.2.3 Hypotheses

Hypotheses to be assessed through user questionnaires

- H_1 Ensemble playing is more both socially and musically rewarding than playing alone
- H_2 Collective awareness is higher in isolated or fully shared instruments than in partially shared instruments

Hypotheses to be assessed through performance analysis and creativity metrics

- H_3 Playing together increases individual creativity, compared to playing alone
- H_4 Playing a partially shared instrument may inhibit individual creativity because of a decreased significance of player's actions
- H_5 A fully shared instrument increases the significance of player's actions and that leads to an increased individual and collective creativity

8.2.4 Configurations evaluated

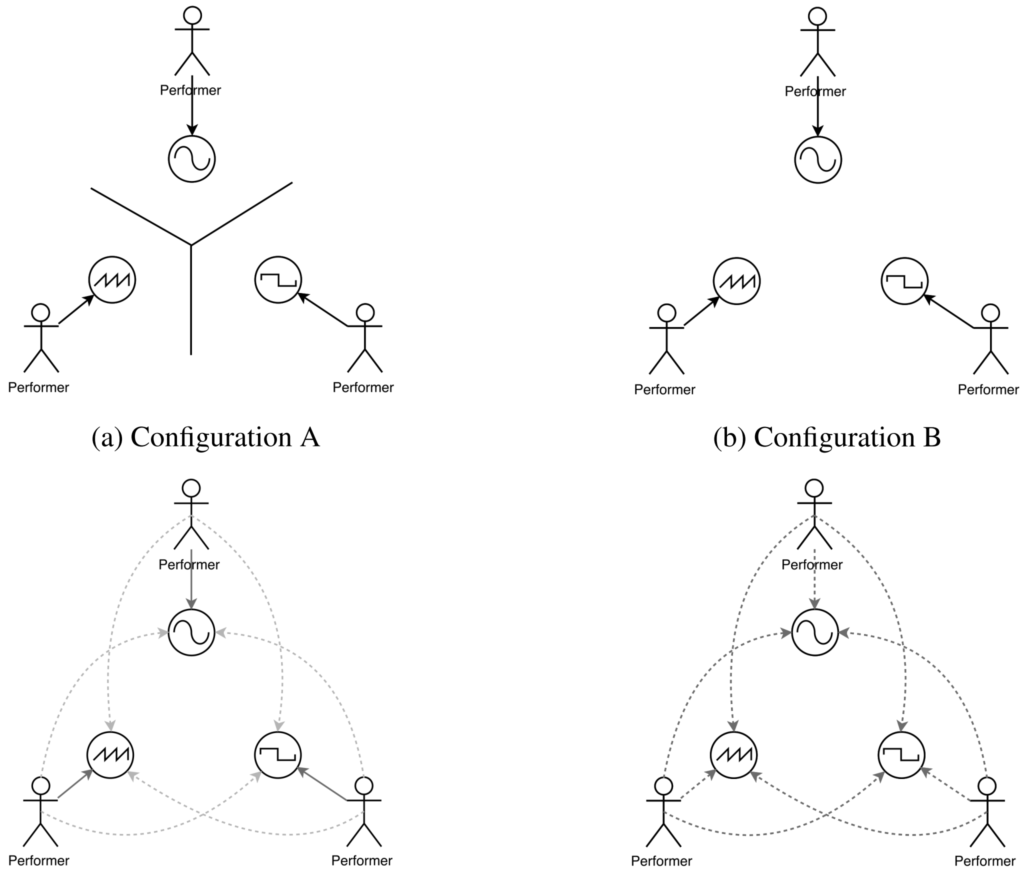
Four scenarios were evaluated, allowing us to compare solo and ensemble performances, with ensemble performances as a continuum from individual instruments to a single collective instrument.

The scenarios are, respectively:

- Configuration **A**. Playing alone, with no aural or visual feedback from any other performer.
- Configuration **B**. Playing together, with independent instruments.
- Configuration **C**. Playing together, with individual instruments slightly coupled.
- Configuration **D**. Playing together with a fully shared multi-user instrument.

Figure 8.2 summarizes the use of shared resources for each configuration, we might consider all configurations equivalent (except for the first scenario where performers are deprived of feedback from the other players) but with a different degree of *leakage* from others' controls into a performer's own control. Therefore, the contribution of a performer to the final parameter value of another performer is null in scenarios A and B, a 12.5% in scenario C and a 33% in scenario D, while a performer owns' contribution to his own instrument decreases from 100% to 75% and finally 33% respectively. The interconnection matrices and their corresponding interdependence metrics for each scenario are displayed in Tables 8.1.

The consequence of this additive interdependence is that the whole range of the control parameter (in this case, the frequency range of the instrument) is shared in different amounts between performers. To make this range expansion symmetric the contribution was bipolar, centered at the mid position of the slider: above it performers added a certain frequency amount to others' frequency control, while below it frequencies were subtracted. A balance had to be found between a desirable minimum range for individual performance and the expanded range which should not be too different between configurations (see Figure 8.3).



(a) Configuration A (b) Configuration B

(c) Configuration C (d) Configuration D

Figure 8.1: Interdependence diagrams for the four configurations evaluated. Experiment 1.

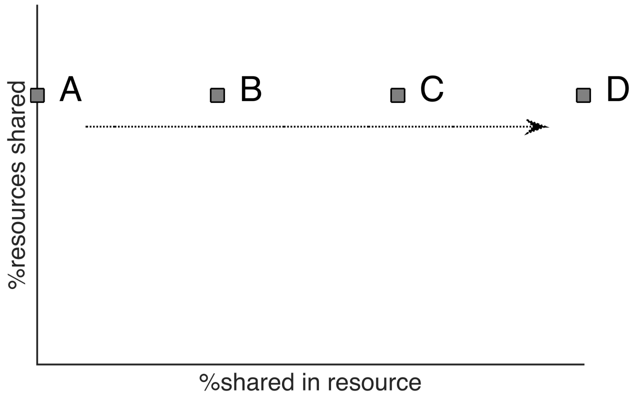


Figure 8.2: Interdependence topology for the four configurations evaluated. Experiment 1.

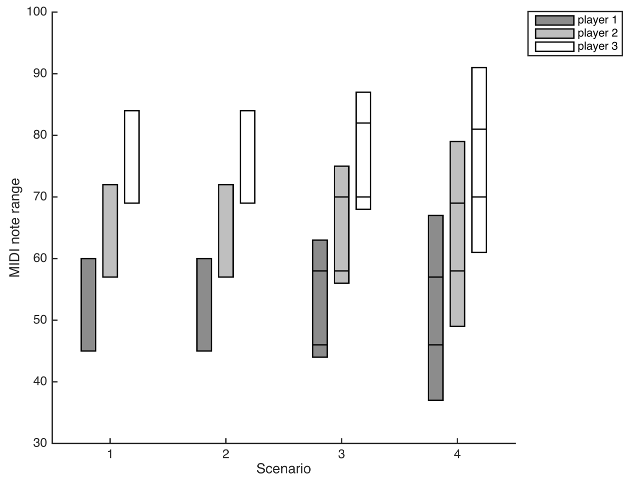


Figure 8.3: Frequency ranges for experiment 1, 4 scenarios. For scenarios 3 and 4 and additional extended range is provided by others' performers additive contribution (upped and lower regions in the ranges)

Table 8.1: Interconnection matrices and interdependence metrics for the scenarios evaluated in experiment 1

(a) A & B. $im = 0$

2	0	0
0	2	0
0	0	2

(b) C. $im = 1/4$

.75	.125	.125
.125	.75	.125
.125	.125	.75

(c) D. $im = 2/3$

.33	.33	.33
.33	.33	.33
.33	.33	.33

8.2.5 Participants

Twelve participants took part in the experiment (all male, mean age 23.4). They were all music students with good familiarity with computer music environments, 9 of them Sonology students and the rest students of music composition. None of them reported former experience in free ensemble improvisation.

8.2.6 Instrument design

A specific collective instrument was designed, consisting of three Theremin-like monophonic synthesizers.

The three sound generators, implemented on a centralized sound server, consisted of simple two-operator FM oscillators, with c:m ratios of 2.013:1, 3.013:1 and 4.013:1 respectively, which imparted distinct harmonically related partials and subtle beating as well, with a modulation index of 0.1 and an added sinusoidal AM (tremolo) at the output of the oscillator with depth 0.0085, 0.015 and 0.023 respectively, and frequencies 3.5Hz, 5.12Hz and 6.35Hz. With such settings they resemble a sort of Theremin trio, with the minimal timbrical differences to make the three oscillators more easily distinguishable.

The sliders shift the base frequency for the carrier and modulator in the FM oscillators. The frequencies are quantized to a Dorian mode when the user jumps to a value, to guarantee a properly tuned, fretless operation, but allowing for free vibrato and *portamento* if the user drags the slider.

The graphical interface is shown in Figure 8.4. This interface both sends the performance data to the server (mouse clicks and slider values) and displays the performance gestures of the other players to guarantee proper collective awareness. Each performer has one *performance slider* assigned (the biggest ones, numbered 1 to 3) and, attached to it, a smaller additional slider displays the actual mapped value to drive the sound generator. In shared configurations, this value will display the combined activity of other performers on the shared frequency parameter. Finally, the small black vu-meter on the right of each performance slider works as a countdown, displaying the time available to perform a *gesture*, which was set to 2.5s.

The instrument plays a sound whenever a user clicks the mouse and keeps dragging the slider, when the mouse is off the sound is blocked. If the performer runs out of time, the sound is blocked as well and he has to free the mouse and click again to start a new gesture.

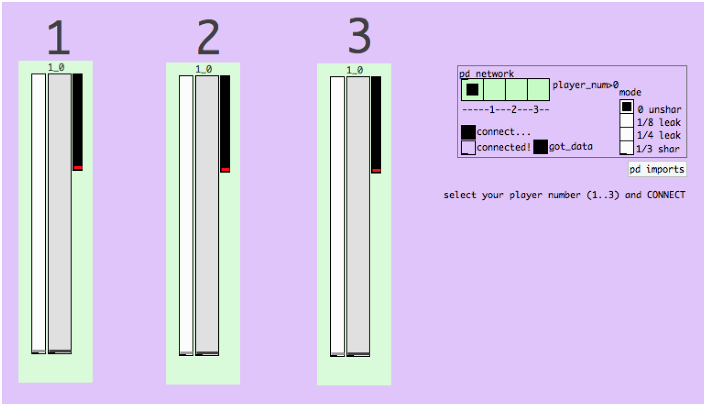


Figure 8.4: Graphical interface for performers. Experiment 1

8.2.7 Procedure and apparatus

The setup consisted on a set of three Windows 7 computers, one per subject, running PureData-based clients and a MacBook Pro running a PureData server. All the computers were sharing the same network through a Switch. Additionally, the server accessed a multichannel sound-card which delivered specific mixes to three headphones, one per player. Performers were not facing each other and the audio mix was provided by headphones, as an attempt to restrict interpersonal interaction to computer mediated feedback as much as possible (see Figure 8.5).

The performance session proceeded as follows. The four configurations were explained and briefly illustrated, and then the user interface was briefly explained. Performers were told the purpose of the restricted time available for single gestures: the need to continuously harvest new and interesting musical ideas, clearly delimited just like spoken words.

Performers were encouraged to freely improvise but, in scenarios B, C and D, they were explicitly asked to try to listen to each other and try to establish a musical conversation, avoiding a too vertical performance, discouraging chordal, homophonic passages.

Performers were given a few minutes to familiarize with the interface and adapt their gestures to the aforementioned temporal constraints. Once finished,

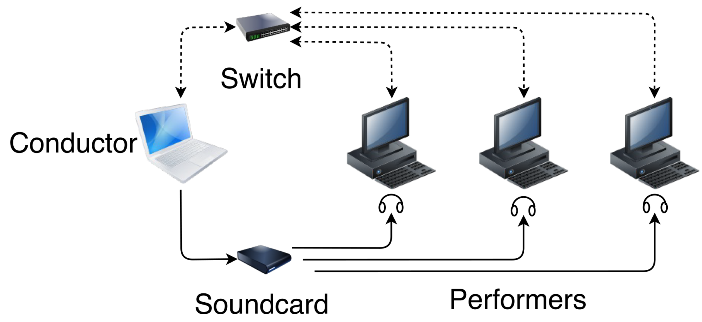


Figure 8.5: Schematic overview of the experimental setup

the experiment consisted on an interrupted eight-minute free improvisation where scenarios were changed automatically every two minutes, in random order.

After the performance an informal conversation with the ensemble took place and the post-test questionnaires were filled.

8.2.8 Data collection

Performance data consisted on mouse clicks and drags over the only available slider. Both actions were collected as OSC data in the server for further analysis. The Windows 7 computers provided a rather low mouse polling rate of 125Hz but the slider location was interpolated within the PureData client patch, therefore to achieve a good temporal resolution the maximum data rate for mouse dragging was 200Hz.

Survey

Players carried out a questionnaire which gathered information about their experience with the four scenarios. The questionnaire contained with multiple choice questions in a 5-point Likert scales, as displayed. The main purpose of it was to evaluate collective engagement and creativity and whether there was a relationship between scenarios and degree of personal and collective agency.

The test was based on the *Mutual Engagement Questionnaire* proposed by N.Bryan-Kinns and F.Hamilton [Bryan-Kinns and Hamilton, 2012] with added questions aimed at gathering preliminary information of the creative behavior in distinct scenarios (see table 8.2). The questions were presented in random order and proper balancing of positive and negative statements to avoid bias in the answer. We choose an odd numbered Likert scale to allow performers the possibility of a neutral evaluation of statements in certain scenarios.

Table 8.2: Questions included in the questionnaire for Experiment 1

Question number	Question(s)
Q1	I like the musical result
Q2	I don't feel involved in the group
Q3	I enjoyed playing it
Q4	What I do is meaningful in the end result
Q5	I don't know what others do
Q6	It differs a lot from an ensemble performance with acoustic instruments
Q7	What others do impedes my goals
Q8	I can clearly notice what I do
Q9	I feel I am not in control
Q10	I understand what happens
Q11	I adopted others' ideas
Q12	It reminds me playing an instrument
Q13	Unexpected things happened
Q14	The result was banal and predictable

8.2.9 Results

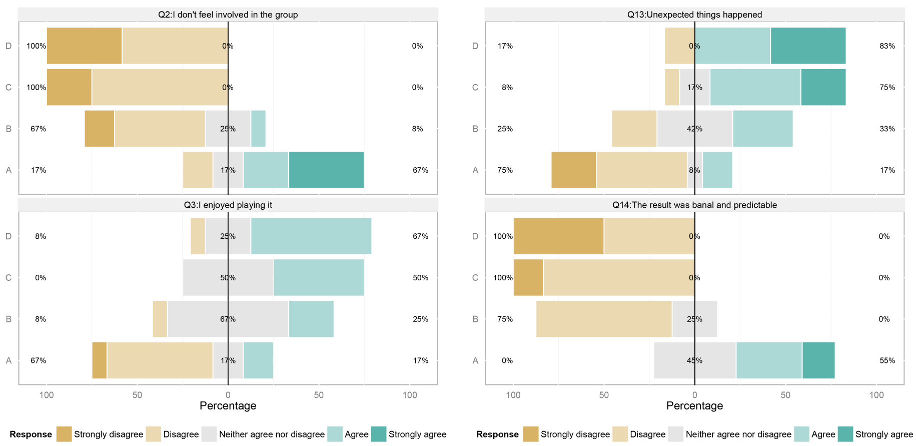
Post-test questionnaire

To compare participants' responses in the four scenarios, we used two-sided Fisher's exact test instead of Chi-square test due to a low sample size of 12 participants. We could identify the following results with statistical significance ($p < 0.05$).

A first, unanimous agreement concerning both user enjoyment and engagement was confirmed for all collective scenarios (B,C,D), while individual performance gets lower scores. This result confirms that ensemble performance, no matter the level of interdependence, provide opportunities for active engagement. Interestingly, the most enjoyable scenarios were 4,3,2,1 in this precise order, suggesting that interdependence is positively regarded (see Figure 8.6a).

Regarding the perceived similarity between the provided scenarios and ordinary musical instruments and instrumental ensembles, subjects significantly identified scenarios C and D as the more distinct both in terms of the characteristics of the instrument and of the overall ensemble, though the answers were not unanimous. Accordingly, and with a similar level of agreement, performers identified conflicts in the last two scenarios.

Subjects reported that the more interdependent, the less banal and predictable the results are. Additionally, performers reported that unexpected things happened in more interdependent scenarios (see Figure 8.6b. This conclusion suggests that



(a) Engagement related answers. (b) Expectancy related answers

Figure 8.6: Summary plots for the post-test questionnaire, grouped by scenario. (I)

collective surprise is expected to be higher in scenarios 3 and 4.

There is no conclusive evidence that subjects' understanding of others' activities change in different scenarios. Indeed, they reported active exchange of ideas in the three collective scenarios. The results might suggest that greater novelty is expected in collective scenarios, being the fourth scenario (D) the one with less exchange of ideas. There is, however, significant evidence that personal agency and feeling in control is higher in not interdependent scenarios (A,B) compared to interdependent scenarios (C,D) (see Figure 8.6a). These results might suggest that the value potential for a performance might be higher in the first two scenarios.

The results of the questionnaire confirm the hypothesis H_1 (*Ensemble playing is more both socially and musically rewarding than playing alone* but don't provide sufficient evidence to confirm the hypothesis H_2 (*Collective awareness is higher in isolated or fully shared instruments than in partially shared instruments*).

Performance analysis

We will proceed now with some preliminary quantitative analysis based on the performance logs.

Co-activity A preliminary observation has to be made concerning this experiment. An ideal scenario for collective creativity measurement assumes that performers actively listen, suggest and react to each other in a free, conversational

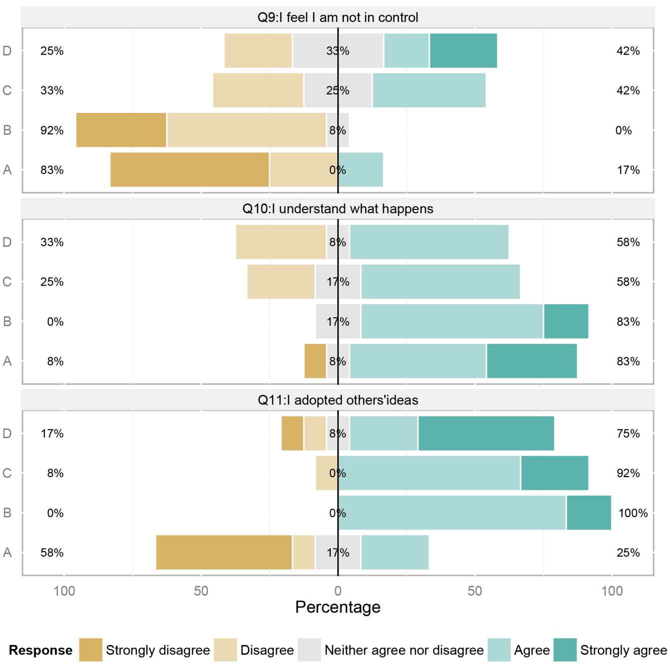


Figure 8.7: Summary plots for the post-test questionnaire, grouped by scenario (II).Agency related answers.

Table 8.3: Mean player's activity and ensemble co-activity for each configuration.

configuration	**A**	**B**	**C**	**D**
mean activity	$.50 \pm 0.12$	$.47 \pm 0.11$	$.49 \pm 0.13$	0.45 ± 0.10
mean co-activity	0.12 ± 0.02	0.12 ± 0.05	$0.15\pm, 0.06$	0.10 ± 0.04

improvisation [Sutton, 2001] which avoids overlaps and collisions between players to maximize dialogic interaction in the performance, a typical practice in the British free *improv* scene[Bailey, 1993], primarily oriented toward the perception-reaction model[Neeman, 2014]. Playing simultaneously decreases collective agency because attentional resources cannot be focused on a single performer.

As none of the subjects were professional improvisers, despite the provided instructions some passages exhibited an excessive level of onset synchronization, suggesting the presence of chordal passages or rhythmically entrainment. We hypothesize that these passages display a decreased collective creativity, because adherence to a collectively established shape or pattern constraints individual exploration and transformation of gestures.

Table 8.3 summarizes the mean time of performer's activity for each configuration, and the mean time of concurrent activity (the three performers playing together). Ideally the first value should not exceed 0.33 for a trio, and our experiments show mean times close to 0.50, indicating that performers don't give enough space for other players or that concurrence was intentional in some passages. Co-activity, by the other hand, should be as close to zero as possible, the means between $.10$ and $.15$ should be, ideally, even lower. A further observation regarding users' activity is that the different configurations do not seem to have any significant effect in it. We hypothesize that an experiment with musicians well trained on free improvisation techniques would exhibit much less overlap and more distinct behaviors in the four scenarios.

Amount of gestures The actual number of gestures in the first three scenarios was rather similar, except for the fourth one, suggesting a different performative behavior took place in that scenario. Indeed, most performers spontaneously switched to a noticeably much quicker sequence of gestures in scenario D (see Figure 8.8).

Gesture sets An additional, preliminary observation can be drawn by comparing the sets of gestures performed by each participant in the different scenarios. As we equate gestures to ideations, such sets indeed define conceptual spaces on a performer basis. By extracting the mean descriptors of such gestures and plot-

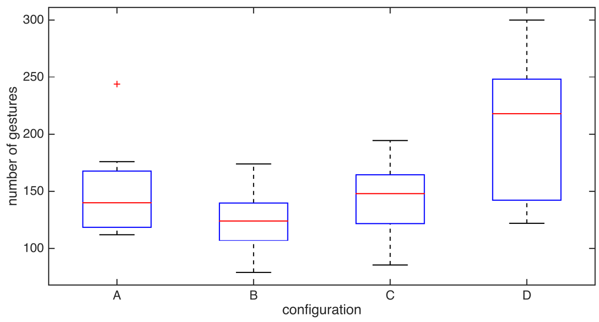
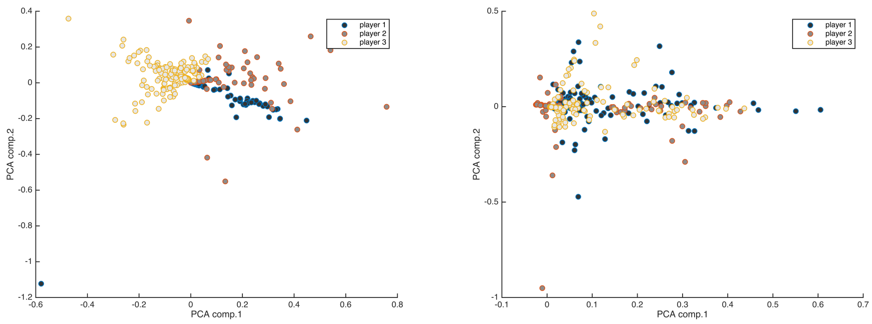


Figure 8.8: Number of gestures per scenario. Experiment 1.



(a) Scenario A (b) Scenario C

Figure 8.9: PCA projections for two sets of gestures. Group 1.

ting them on the plane defined by their two principal components (see Figure 8.9), we may easily visualize the conceptual space explored by each performer and by the whole ensemble. Interestingly, we may easily appreciate how such conceptual spaces are much more well defined in scenario 1 (as in sub-figure 8.9a, where performers play alone, without feedback from the other players) and become more scattered and overlapped in scenario 3 (as in sub-figure 8.9b, where performers play together with interdependent parameters).

While, due to the limited scope of the experiment, this preliminary evidence is not consistent enough to allow us to characterize performers according to their creative signature -as in [Goebl et al., 2004]- or possibly identify performers within a joint performance -see for example [Swift, 2012]-, it suggests that ensemble performances, including interdependent scenarios, promote an active process of exchange and transformation of creative ideations between performers, and apparently confirms the existence of consistent shared conceptual spaces. The convergence of musical ideas, in its turn, is a good prospect for mutual engagement assessment [Bryan-Kinns, 2012].

Creativity metrics

This section evaluates quantitatively the performances in the four configurations studied by computing metrics for the novelty, surprise and value potential for each performer, according to the procedure detailed in Chapter 7.

Even if we could compute a single creativity metric as a product of the three creativity attributes, it may be more illustrative to study them separately, as suggested in 7.3.6. Finally, although our creativity attributes are guaranteed to be in the range $[0, 1]$, more than evaluating their actual absolute values and ranges in the experiment, we seek to compare them between the four scenarios. Therefore, for clarity we will post-normalize the computed metrics, in such a way that the highest novelty, the highest surprise and the highest value potential among all the performances will be all 1.

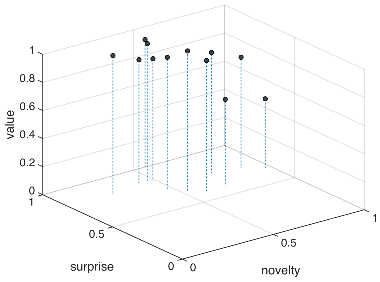
Finally, we may display the three normalized creativity attributes for the four scenarios in three-dimensional plots (see Figure 8.10), each plot belongs to a single performer. The most obvious trend observed, despite the notable dispersion, is a decrease in value potential in configurations C and D compared to configurations A and B. This was an expected result, because of the disturbance added to the frequency output by the additive contribution of the other performers.

We may tentatively seek additional trends by averaging all the performances and summarizing the creativity attributes for each scenario, though we should take into account that the sample size is too low as to draw rather meaningful conclusions (only 12 individual performances per scenario). This final plot (see Figure 8.11) displays the novelty, surprise and value potential for each configuration evaluated.

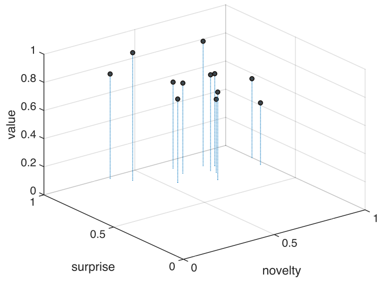
The scenario promoting the most creative behavior is a classical ensemble of independent musical instruments (scenario B). As reported by the performers in the questionnaires, it allows for an optimal collective awareness which maximizes the transfer of musical ideas, increasing the rate of novel gestures and their magnitude throughout a performance.

Additionally, it shows the highest ratios for surprise as well. High surprise, according to our metrics, may imply not only higher novelty ratings but the ability to set up gestural trends *purposefully* as well, and it seems scenario B provides the best conditions for it. The comparative ratings for both novelty and surprise between scenarios A and B seem to confirm our hypothesis H_3 - *playing together increases individual creativity, compared to playing alone.*

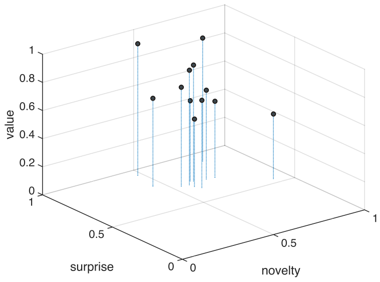
Performances under scenarios A (individual performance) and C (slightly shared performance) are overall less creative but for different reasons. In an individual performance we appreciate less novelty, presumably because of the lack of the additional stream of ideations coming from other performers which might be adopted. On the other side, the slightly shared performance clearly shows a lower



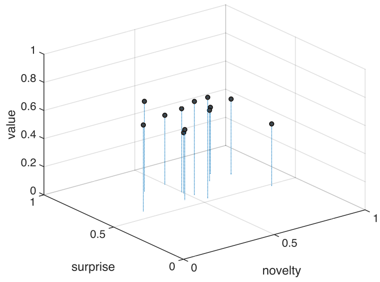
(a) Configuration A



(b) Configuration B



(c) Configuration C



(d) Configuration D

Figure 8.10: Scatter plots showing the normalized creativity metrics for the four configurations evaluated in Experiment 1. Each plot indicates the mean novelty, surprise and value for every performer taking part in the experiment

[Scatter plots showing the normalized creativity metrics for the four configurations evaluated in Experiment 1]

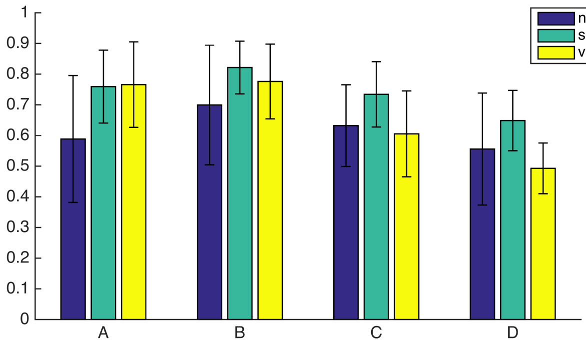


Figure 8.11: Novelty, surprise and value for each configuration in Experiment 1. The metrics have been averaged on a performer's basis and post normalized.

value potential. This result was expected as the audible realization of individual gestures is partially mixed with others' contributions. The low ratings for the value metric in scenario C would confirm hypothesis H_4 *-playing a partially shared instrument may inhibit individual creativity because of a decreased significance of player's actions*.

Finally, performances in scenario D are rated as the less creative of all, with a further decrease both in surprise and value ratings compared to scenario C. Note, however, than novelty is approximately the same that in scenario A. Low novelty and surprise ratings are expected given the comparatively fast performances which were unexpectedly observed for this scenario in virtually all groups. These results seem to reject our last hypothesisH_5, at least in terms of individual creativity - *a fully shared instrument increases the significance of player's actions and that leads to an increased individual and collective creativity.*

It was expected that scenario D would provide a more coherent, unified environment in which the loss of exclusive control in one performer's instrument was compensated by a simultaneous gain in others'. In other words, the decrease in value potential should be compensated by higher novelty and surprise ratings. To clarify this, notice that, as long as in our fully shared environment any individual action symmetrically affects all three sound generators with the same amount and in the same parameter, and this effect is further emphasized by mapping all three frequencies to the same pitch in different octave ranges, any gesture exerted by one performer on his interface will be projected with the same pitch gesture to the three sound generators. Unfortunately, as we discussed before, the quick collective dynamics observed in this scenario did not allow to take benefit of the sharing of ideas between performers.

As for collective creativity, we may evaluate the sets of gestures regardless of their authorship as described in Chapter 7. The metrics for the scenarios in which performers played together, that is, scenarios B, C and D, are shown in Figure 8.12.

The results are sensibly different, specially in the surprise metric of scenario D, which is relevant. Given a comparatively low novelty rating for this scenario, the high rating for the surprise is to be attributed to dynamic surprise alone, as structural surprise is mostly correlated to novelty. This was effectively confirmed in the video recordings: despite the notoriously quick activity which prevented performers from elaborate sets of distinctive gestures, both from an individual and from a collective perspective, all groups resorted to a dynamics in which performers followed former gestures, imitating them and clearly setting collective trends. Whenever such collective trends are defeated collective dynamic surprise arises, as is clearly seen in this last plot. Therefore we might partially support H_5, at least in terms of dynamic collective surprise.

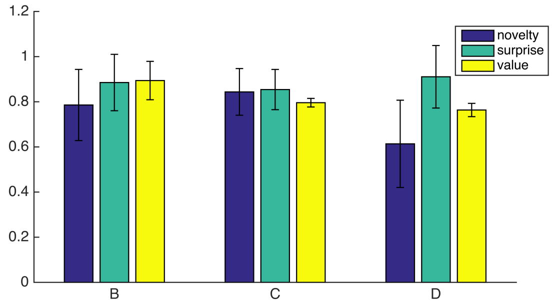


Figure 8.12: Collective novelty, surprise and value for each configuration in Experiment 1. The metrics have been averaged on an ensemble basis and post normalized.

8.2.10 Preliminary conclusion

This section concludes the first experiment, in which we comparatively studied musical creativity on individual, ensemble and interdependent scenarios with variable degree of interdependence.

We will now proceed with the description of our second experiment. First, we will introduce our motivations to conduct an additional experiment and then we will present our performative context and the aims and hypothesis formulated. The performance analysis will proceed in a similar fashion to Experiment 1.

After the second experiment we will summarize the results obtained for both experiments, and will provide a final comparative discussion of our findings.

8.3 Experiment 2 : Creativity and resource ownership

8.3.1 Introduction

In Experiment 1 we evaluated creativity in multi-user instruments by seeing interdependence as a step beyond group performance. It served us to compare musical creativity in individual, ensemble and interdependent performances.

Our controlled variable was the amount of concurrent control on a shared parameter. Through the study of the different shared scenarios, designed to allow additive concurrent control, we could appreciate for example how unbalanced shared controls could negatively affect performer's creative outcome.

However, the conversational-like design of the interface prevented users from fully exploring concurrent access: only in periods of concurrent activity several performers could influence each other, but both the instructions given to performers and the timer mechanism conceived to enforce a segmented performance prevented users from exploring the possibilities of intense interplay provided by multi-user environments. To a great extent shared access was not fully understood by performers.

While this approach is possibly the most suitable for the assessment of our creativity metrics, due to its simplicity, it does not yet fully address one of the central questions of our research: how creativity manifests in contexts of full interdependence between performers. Therefore, our control variable in this experiment will not be the amount of interdependence through a single shared parameter but the number and typology of the parameters shared.

Experiment 2 will provide a performative context intended to maximize the ensemble interplay and let them fully appreciate the effects of mutual interdependence. We expect with this experiment to gain further insight into the relationship between exclusive or shared access to a number of instrumental resources and the creative outcome of the ensemble.

Table 8.4 summarizes the differences between the instruments designed for each experiment. As we may appreciate, the second instrument is designed to provide more opportunities for concurrent interplay, being at the same time an environment prone to scenarios of conflict and competition for the shared resources.

8.3.2 Aims

This study seeks to measure individual creativity in a musical performance with a multi-user instrument which fully shares several resources (namely parameter controls and audio channels), and evaluate how may it be affected by the poten-

Table 8.4: Comparison between the multi-user instruments used in Experiments 1 & 2

Experiment 1	**Experiment 2**
one shared parameter	multiple shared parameters
balanced parameters	unbalanced (one parameter may negate the other)
parameters may be partially shared	parameters are either unshared of fully shared
segmented by default	unsegmented by default
concurrent acces through mix	concurrent access through replacement
stops sounding if not playing	keeps sounding if not playing
one sound channel per performer	sound channels may be shared

tially disruptive effect of concurrent and competitive access to shared resources.

The performative scenarios (detailed below) address typical configurations of sound awareness and control ownership in a computer network (see Chapter 3). The experiments themselves consists of short collective, free exploratory improvisations with iconic *Musique Concrète* sound materials processed in realtime by the performers.

In contrast to Experiment 1, this experiment places a less restricted improvisational scenario where performers can interfere and mask other performers through the access of shared resources. We have surveyed in Chapter 3 a number of different approaches and relevant repertoire which addresses the issues of shared resources, and particularly shared sound sources and shared controls. This experiment will investigate to what extend the exclusive or shared access to those resources may impact in the creative performance on a computer music ensemble.

On a first stage, we will study performers' attitude and behavior regarding different combinations of shared controls and shared sound channels. We want to know whether such configurations have a noticeable impact in the level of engagement and collective awareness throughout the performance, and whether performers resort to specific strategies regarding the potentially conflicting nature of shared resources.

On a second stage we will compute creativity metrics for the different configurations in order to determine their relationship with the configurations evaluated.

8.3.3 Hypotheses

Hypotheses to be assessed through user questionnaires

- H_1 Players feel more involved when they *own* specific instrumental resources, may them be timbrical controls or sound output channels.

- H_2 Personal and collective awareness increase with audio and control ownership.
- H_3 Playing with shared controls might be more enjoyable but ultimately less musically rewarding.

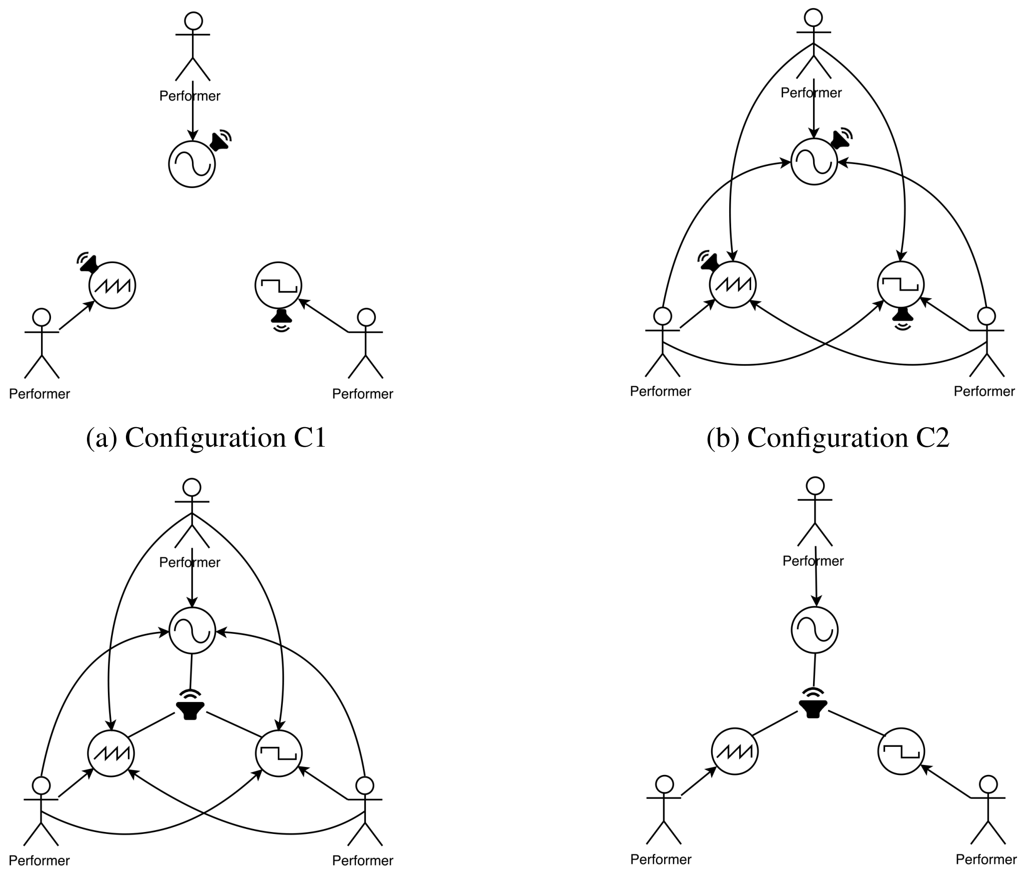
Hypotheses to be assessed through performance analysis and creativity metrics

- H_4 Sharing controls may lead to higher individual novelty ratings because performers are more closely exposed to others' contributions, promoting a more active exchange of ideations.
- H_5 Sharing controls on non shared audio channels may lead to a lower value potential due to the amount of time a performer plays other instrument or other performers play his instrument, because of the decreased significance of his gestural actions on his own audio channel.
- H_6 We don't expect that a single audio output inhibits the ideation of novel and surprising gestures, but the increased mutual masking in these scenarios might encourage resorting to more exaggerated gestures and more salient timbrical manipulations.
- H_7 Value potential with a single audio output, for any configuration of controls, will be lower because the audio output incorporates the gestures of all the performers.

8.3.4 Configurations evaluated

Four different configurations were presented to the participants (see Figure 8.13)

- Configuration **C1**. Sound and control ownership. In this configuration each performer could control only one assigned sound source and the sound output was routed only to his loudspeaker.
- Configuration **C2**. Sound ownership but distributed control. In this configuration each performer kept a sound source but anyone could freely operate on the controls of all the instruments.
- Configuration **C3**. Control ownership but distributed sound. In this configuration each performer could control only his sound source but the whole mix was equally routed to all the loudspeakers.
- Configuration **C4**. Distributed sound and control. In this configuration any performer could freely operate on all the instruments and the whole mix was routed to all the loudspeakers.



(a) Configuration C1

(b) Configuration C2

(c) Configuration C3

(d) Configuration C4

Figure 8.13: Interdependence diagrams for the four configurations evaluated. Experiment 2.

(a) C1. $im = 0$

4	0	0
0	4	0
0	0	4

(b) C2. $im = 1/3$

4	1	1
1	4	1
1	1	4

(c) C3. $im = 1/6$

3.33	.33	.33
.33	3.33	.33
.33	.33	3.33

(d) C4 $im = 2/3$

1.33	1.33	1.33
1.33	1.33	1.33
1.33	1.33	1.33

Figure 8.14: Interconnection matrices and interdependence metrics for the scenarios evaluated in experiment 2

The interconnection matrices and their corresponding interdependence metrics for each scenario are displayed in Tables 8.14, though it is debatable whether the actual weight of each shared parameter, most notably the audio channel, are equivalent. As in Experiment 2, the first scenario consist of independent, single-user instruments and the interdependence index gradually increases in the following scenarios.

8.3.5 Participants

Twelve participants took part in the experiment, in four groups of three players each (all male, mean age = 24.6). They were Sonology students, musicians with proficient technical background and instrumental training in different musical styles (classical, folk and jazz respectively). All of them were familiar with computer music tools but had only limited experience in improvisation. Most had no previous experience with collaborative musical interfaces.

8.3.6 Instrument design

A specific collective instrument was designed, consisting of three phase-vocoders with flexible routing of parameter controls and audio channels. Distinct sound sources were loaded in each of the phase-vocoders (Freesound[1] samples of bells, a squeaking door and a steam train respectively).

Each individual sound source offered three controls, available as individual sliders. As such, players could only operate with a single slider at any time. To guarantee collective awareness, all players could visualize those nine sliders and

[1] `http://freesound.org/`

who was operating on them, as displayed in Figure 8.15. Those controls were, from more generic to more instrument specific: a volume control, a sample offset control and an instrument-specific timbrical control (respectively, an FFT based partial threshold, an spectral compander and a low frequency shelving filter).

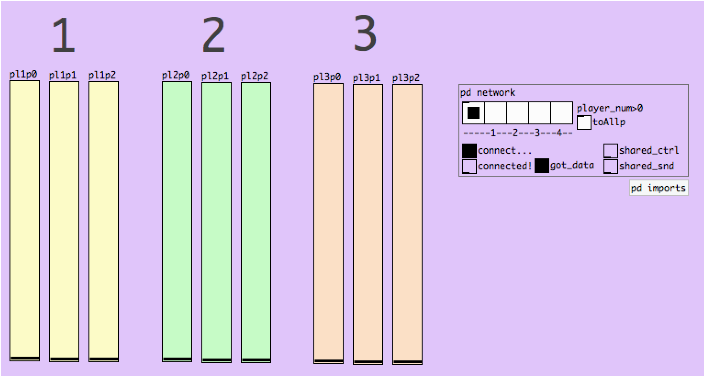


Figure 8.15: Graphical interface for performers. Experiment 2

Both the controls and sound output of N sound synthesizers were distributed among N players either in N:1 or N:N topologies. As opposed to experiment 1, where there was a 1:1 relationship between control and timbrical parameter (frequency) and where such parameter was actually pre-computed and could be readily incorporated into our analysis, this scenario deals with a more complex mapping: parameter 1 modifies the volume in a 1:1 relationship, but the volume may be influenced by parameters 2 as well (because it essentially depends on the soundfile contents and may be further altered by the spectral effects incorporated in parameter 3) , parameter 2 is mapped to a sound file pointer, therefore an arbitrarily 1:M mapping is established between its value and the audio content of the sound file, while parameter 3 is essentially an arbitrary sound modifier, which performs another 1:M mapping but to the audio, modifying its spectral content on multiple frequency bands.

The instrument was implemented in Puredata with a server patch acting as a centralized DSP and control routing, ans a client patch providing the graphical interface mentioned above. A multichannel soundcard was attached to the laptop running the server patch, providing individual audio channels to loudspeakers located next to each performer.

8.3.7 Procedure and apparatus

Each rehearsal started with an introduction by the researcher. Participants had a short training period (two minutes, with headphones) with each of the three instruments, being assigned to a random one after the training. Participants were then given between 90 and 120 seconds to play with each experimental condition with speakers. Once finished, participants filled a post-test questionnaire and a final group discussion was held.

The setup consisted of three Windows computers on a LAN, with 19” monitors, placed one next to the other at a distance of 80cm. Small X-mini II speakers were placed in front of each computer to provide localized audio. The room provided sufficient sound isolation for the performance despite the low power rating of the speakers (2.5W RMS).

The provided minimalistic graphical control interface (see Figure 8.15) was similar to Experiment 1, but in this case the feedback of the final value for each parameter was displayed as an actual change in the slider position instead of an adjacent fader. The purpose is to clearly inform performers that there is effectively a single shared parameter control, and its value is at any time the last value received by any of the performers manipulating it.

8.3.8 Data collection

Both the performers action on the faders and the configuration changes done by the conductor were logged as time-tagged OSC data. Event logs contained information on the performer, the instrument, the fader and its value and was sent asynchronously whenever a fader was moved. In this experiment we didn’t capture the actual performer actions needed to operate with the sliders (mouse drags or clicks).

Survey

Players carried out a questionnaire which gathered information about their experience with the four scenarios. The questionnaire contained with multiple choice questions in a 5-point Likert scales, and was similarly inspired on the *Mutual Engagement Questionnaire* [Bryan-Kinns and Hamilton, 2012] though with a reduced subset of questions, as listed in table 8.5.

The main purpose of the questionnaire was to evaluate collective engagement in each scenario and determine whether such engagement was either more socially or musically centered.

The questions were presented in random order and proper balancing of positive and negative statements to avoid bias in the answer. The choice of an odd numbered Likert scale was the need to allow performers the possibility of a neutral evaluation of statements in certain scenarios.

Table 8.5: Questions included in the questionnaire for Experiment 2

Question number	Question(s)
Q1	I like the musical result
Q2	I don't feel involved in the group
Q3	I enjoyed playing it
Q4	I feel I am not in control
Q5	I understand what happens
Q6	I don't know what others do
Q7	I can clearly notice what I do

Video recording

The sessions were video recorded. A final group discussion served to elicit user preferences and perceived differences in in the four scenarios regarding awareness, focus and interaction strategies.

8.3.9 Results

Post-test interview

An informal talk was carried out after the tests to discuss and compare the user experience in the different scenarios.

Participants raised mixed and somehow unexpected comments regarding the audio configurations.

The shared audio channel didn't seem to affect the ability to identify themselves and other players, but we might question here the experimental setup: participants were placed in a row, therefore the relative volumes of other performers' instruments were affected by their relative distance. Indeed, two performers reported it was harder (and not easier, as it should be expected) for them to identify some players in non shared audio configurations, those which were placed farther apart.

Additionally, when sharing a single audio channel and radiating it through all the three loudspeakers simultaneously, an unexpected spatialization effect dramatically changed the perceived acoustic space, which turned out to be broader and more reverberant. Most performers were surprised when the audio configuration was changed to a shared channel, and two of them reported this significant change was a bit distracting at first, and it took them some seconds to get used to it.

All these issues might be addressed by resorting to headphones as we did in experiment 1, and applying a suitable panning and gain to every sound source, but this setting would loose the intuitive physicality of location-dependent sound sources. Otherwise, providing a quasi acoustic experience with shared instruments deserves further study regarding the distribution of audio sources.

More interestingly, shared controls proved to be intriguing for a number of performers, and a source of social interaction. One of them reported that *"I was playing to counterfeit others' movements, whenever one performer was doing something very obvious, I lowered his volume"*. Similarly, another player told that *"it is funny to interfere with other players' actions, fighting for the same control"*.

Post-test questionnaire

To compare participants' responses in the four scenarios, we used once again two-sided Fisher's exact test instead of Chi-square test due to our low sample size of 12 participants. We could identify the following results with statistical significance ($p < 0.05$)

When analyzed in pairs, scenarios with non-shared controls -C1 and C3- provided better agency (ie answer to the question Q5:*I understand what happens*) than shared ones -C2 and C4, with Fisher's one-tailed Exact portability test $p < 0.01$. Comparatively, self-awareness was not significantly different between both sets of scenarios (it was evaluated with Q7:*I can clearly notice what I do*).

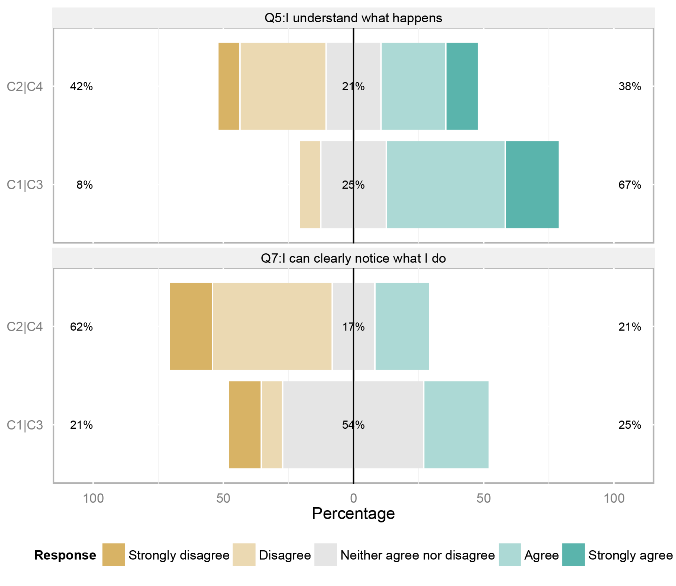


Figure 8.16: Summary plot for the post-test questionnaire, grouped by shared/unshared controls scenarios. Collective agency and self-awareness related answers.

Participants felt most involved in the group in scenario 3 and the minimum involvement was in scenario 1, though significance was just below $p < 0.05$. That is, controlling your own instrument under a single shared audio channel was perceived as the most collectively involving scenario, while controlling your own instrument with independent audio channels was perceived as the least involving one.

The hypotheses that participants might find scenarios with shared controls funnier to play even though the musical result could be better with individual controls was rejected. Participants reported that all the scenarios were virtually equally engaging. Indeed, more people tended to find playing with unshared controls more musically satisfactory but this was not statistically significant.

The main conclusions regarding the questionnaire results are:

- Control ownership positively contribute to self-awareness and group involvement. This evidence partially confirms H_1 -players feel more involved when they *own* specific instrumental resources.

- Audio configurations don't significantly contribute to personal or collective awareness, but a configuration consisting of personal, unshared controls is regarded as the most collectively involving scenario, whereas the opposite, a collective instrument with shared control and audio, is considered the least involving one. This conclusions rejects H_2 in relation to audio ownership and is only in partial agreement with Fencott's study on audio delivery and territoriality, in which the author advocates for unshared audio configurations to encourage group involvement and user awareness [Fencott and Bryan-Kinns, 2012], therefore we suggest that performers partially compensated the lack of location dependent aural cues by efficient source separation, even if it required an additional effort which might be detrimental to group involvement.

- Finally, there was no significant evidence, either from the interviews nor from the questionnaire, to confirm or reject H_3 -playing with shared controls might be more enjoyable but ultimately less musically rewarding.

Performance Analysis

Before addressing the quantitative assessment of musical creativity, some preliminary observations may be done regarding the activity and behavior of performers in each scenario.

Table 8.6: Fraction of time devoted to own and others' controls in each scenario.

configuration	**C1**	**C2**	**C3**	**C4**
time on own controls	1.0	0.36 ± 0.21	1.0	0.30 ± 0.14
time on others' controls	0.0	0.64 ± 0.21	0.0	0.70 ± 0.14

Distribution of activity By comparing the fractions of time each performer devotes to his own controls and to others' controls in shared configurations (scenarios C2 and C4), we notice they tend, on average, to be evenly distributed among all three synthesizers. Indeed, we may see in Table 8.6 when controls are shared the whole set of controls available are actually accessed quite uniformly, suggesting that ownership shifts accordingly from single sound generators to the whole multi-user instrument.

While this evidence could suggest that performers immediately perceive the whole setup as a monolithic and uniform entity when controls where shared, this was not always the case.

For some groups and shared configurations, we observed than players tended to operate *on an instrument basis*, that is, once they started exploring an instrument, it was more probable to stick on another slider from the same instrument than to switch to a different one. Figure 8.17 exemplifies such clear distribution of roles. In terms of creativity, such experience should be rather comparable to playing an individual instrument.

In some other groups, players resorted to rapid alternating manipulations of faders from all the instruments. As observed in the recordings of the performances, such dynamics spontaneously emerged whenever one performer unexpectedly modified a parameter of an instrument already being operated by another performer, triggering a cascade of similar actions. Eventually such activity lead to a kind of social game in which relationships between gestures and therefore any kind of gestural creativity were secondary to the social experience itself. Due to the presence of such behaviors, lower creativity ratings are therefore expected in shared scenarios on average. This behavior was mostly observed in scenario C2 but was also present in scenario C4 (see Figure 8.18).

In sharing control scenarios performers systematically avoided conflicting on the same slider. Only one performer in one of the experiments used it to counterfeit what another performer was trying to do, an attitude that was depicted as too intrusive by the other performer, and subsequently rejected. For the rest, only occasionally several players interacted with the same control simultaneously. Table 8.7 illustrates this fact, showing an extremely low proportion of time devoted to concurrent edition of the same slider by two or more players.

Both behaviors, conflict avoidance, which was generalized, and instrument-

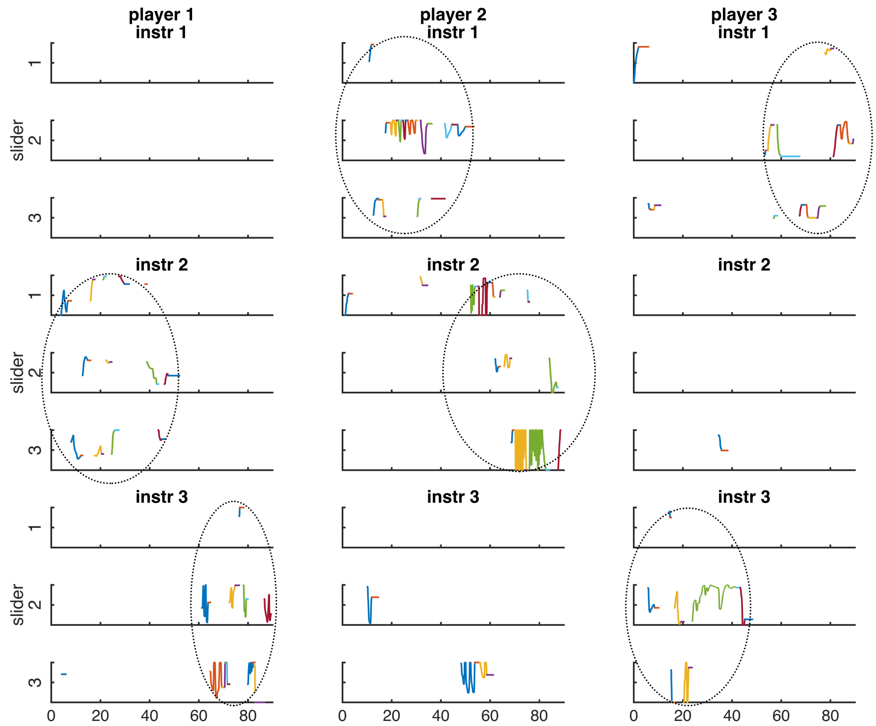


Figure 8.17: Distribution of user activity in a shared control scenario (Experiment 2, Scenario C2). Notice how players focus activity on an instrument basis (encircled areas), mostly avoiding simultaneous operation on the same resource.

Table 8.7: Fraction of time with concurrent activity (2 or more players acting on the same slider)

configuration	**C1**	**C2**	**C3**	**C4**
time (fraction of total duration)	0.0	0.013 ± 0.022	0.0	0.018 ± 0.026

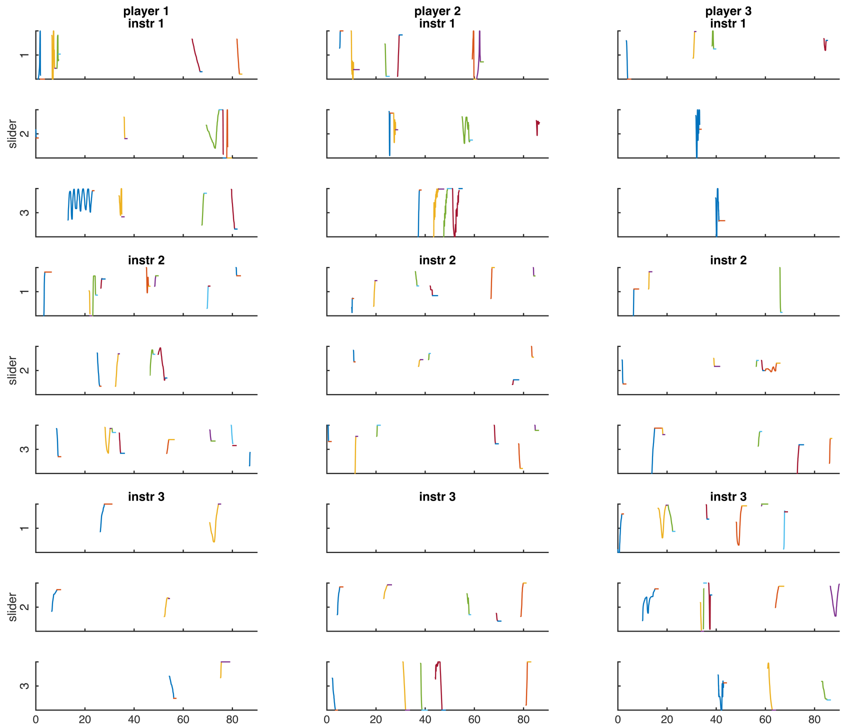


Figure 8.18: Distribution of user activity in a shared control scenario (Experiment 4, Scenario C4). Notice the rapid alternation of brief and fast gestures in all performers.

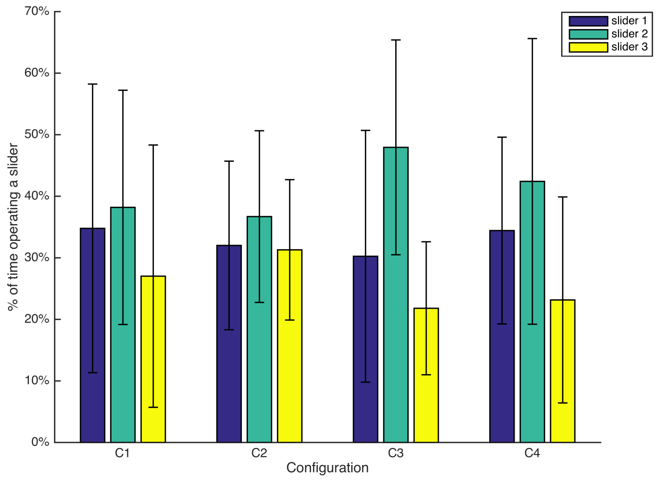
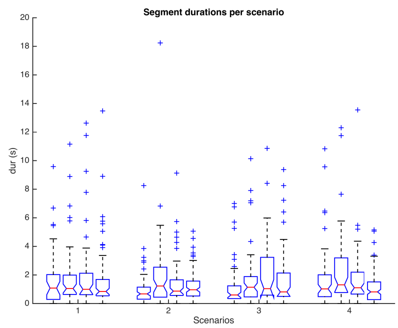


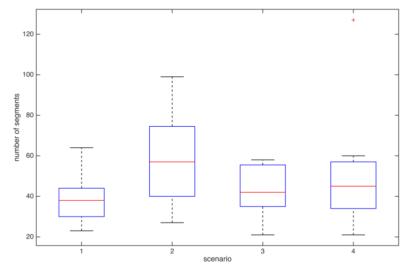
Figure 8.19: Fraction of time devoted to each slider.

based activity, which was found occasionally, suggest an spontaneous, unnegotiated distribution of *performative spaces*. Such territoriality has indeed been observed both in co-located settings with collective tangible interfaces where performers actually negotiate a physical space [Xambó, 2015] and in computer-mediated collaborative environments [Fencott and Bryan-Kinns, 2012].

Use of resources As for what timbrical modifications performers do prefer, we could analyze the time spent performing with each type of slider. The second slider is always designed to give the most salient timbrical parameter (the temporal offset in the sample being played) and it requires to be actively operated to recreate the original sample. The 1st slider is also relevant as it controls the volume of the sound generator, which otherwise is permanently active, while the 3rd slider performs a more subtle spectral processing. Figure 8.19 shows high variance in the proportion of time devoted to each sliders, but suggests otherwise a tendency to favor sliders in order of timbrical saliency. This trend is more evident in configurations C3 and c4, the ones with a shared audio channel. We hypothesize that, when performers share a single audio channel, resorting to more salient timbrical modifications helps to make them contributions more distinguishable despite the loss of spatial cues provided by separate audio channels. The aforementioned results are in accordance with our hypothesis H_6, which states that the increased mutual masking in single audio channel scenarios might encourage resorting to more exaggerated gestures and more salient timbrical manipulations.



(a) Duration of segments for each scenario and group. Experiment 2.



(b) Segments for each scenario. Experiment 2.

Figure 8.20: Summarizing plots for the segmentation process in Experiment 2.

Amount and duration of gestures As previously stated, this experiment did not enforce a segmented performance. The motivation was twofold: by one side we wanted to explore a performative environment more prone to concurrent activity, and, additionally, we noticed in preliminary tests with the interface that performers' activity could be reasonably segmented according to a pair of rather simple heuristics, which made an enforced segmentation unnecessary:

- A *rest period* was considered to signal the end of a segment (the rest period was set to 0.75secs).

- Switching to a different slider also indicated a new segment

The segmentation process showed a rather consistent tendency to perform with short gestures between 0-2secs of duration, though there was a notable presence of much longer gestures which could not be properly segmented (see Figure 8.20a). The number of gestures per player and scenario are therefore around 30 and 60 for our performances which lasted 90secs (see Figure 8.20b). The notable exception is scenario C2, which shows a lower number of high duration gestures and a higher number of gestures on average due to the rapid switching dynamics observed in performances in this scenario.

Finally, by performing feature extraction in the segmented performances we may appreciate the distinctive behavior in shared control configurations (C2 and C4). Figure 8.21 shows how in both configurations the variation of speed was higher. While the average speed is always zero because we deal with sliders, performers achieved higher absolute speeds in both scenarios as well.

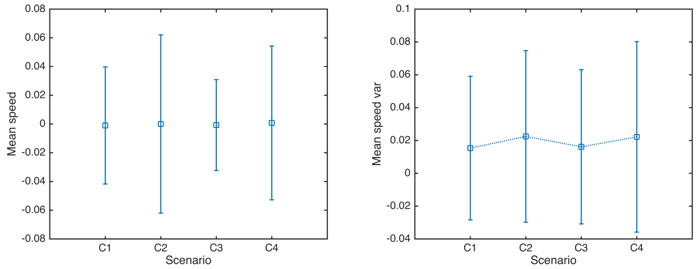


Figure 8.21: Average speed and speed variation for each scenario in Experiment 2.

Creativity metrics

We will now compute the individual creativity attributes and plot them on three-dimensional plots for a quick preliminary visualization, following the same procedure described in Experiment 1.

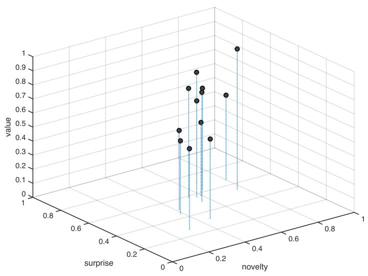
It is worth mentioning that, to compute the value potential for all configurations, the clusterized set of gestural ideations from each performer was compared to the clusters derived from the analysis of his own audio channel, no matter which performer was actually controlling what instrument. In the shared audio configurations the procedure was the same, though in that case the audio channel was the mix of the three instruments.

The results, compared to the first experiment, show more clear defined trends (see Figs. 8.22), notably in the value potential, but with marked differences in the other attributes as well.

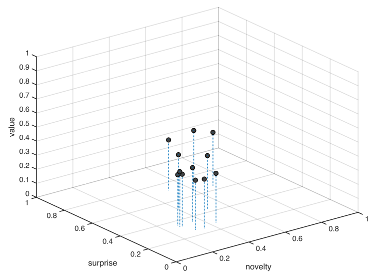
As before, we will provide a summarizing bar plot to better analyze the creativity for each scenario. The plot (see Figure 8.23). The results show a higher dispersion than Experiment 2, a result which accounts for a larger variability in the strategies adopted by performers in shared scenarios, as discussed before.

The results, in contrast to experiment 1, don't confirm the expected higher novelty ratings in shared scenarios because of a more active exchange of ideations, proposed as hypothesis H_4. However, by comparing creativity and surprise in scenarios with a single audio output, we may confirm our hypothesis that they are not inhibited by a lack of isolated sound sources (H_6).

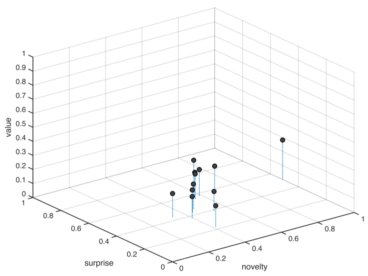
Both novelty and surprise follow a similar pattern: the least creative scenario with respect to such attributes is C2 (shared controls but individual sound channels). Interestingly, just by unifying the sound sources (C4) higher ratings seem to be achieved. Though the results were unexpected and should be taken with caution given the low sample size and its high dispersion, we could hypothesize that sharing a single sound source actually *improved* collective awareness, because performers were able to successfully isolate the individual sound streams coming



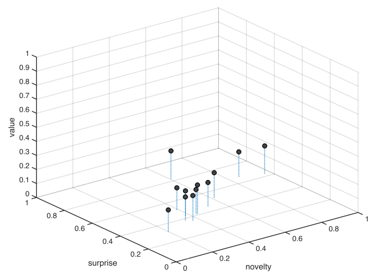
(a) Configuration C1



(b) Configuration C2



(c) Configuration C3



(d) Configuration C4

Figure 8.22: Scatter plots showing the normalized creativity metrics for the four configurations evaluated in Experiment 2. Each plot indicates the mean novelty, surprise and value for every performer taking part in the experiment

[Scatter plots showing the normalized creativity metrics for the four configurations evaluated in Experiment 2]

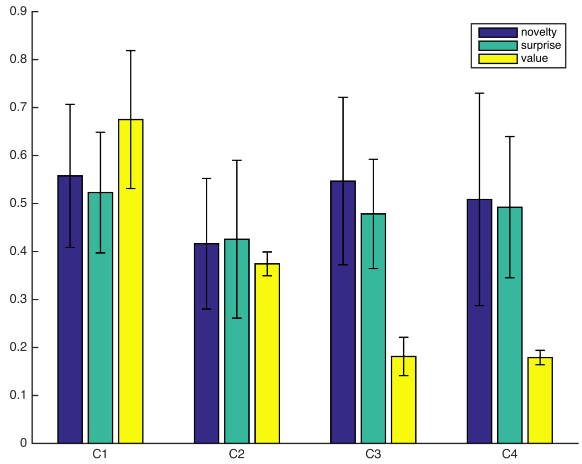


Figure 8.23: Creativity metrics for Experiment 2

out from a single loudspeaker. This result is in accordance with some comments from participants regarding the unequal aural feedback from other players because of the experimental setup, which located performers in a row.

With respect to the value potential, the differences are much more evident but, once again, the results should be interpreted with caution. The lower rating for C2 compared to C1 actually indicates that the audio coming out from a performer is less related to his performative manipulations, either because the performer was operating on a different sound generator or because other performers were operating on his own. On the other side, the even lower ratings for C3 and C4 indicate that, according to our metric, a shared audio channel is hardly capable of projecting the individual gestural creativity as an audible outcome, even in a scenario with unshared controls (C3). This results are in partial accordance with out hypotheses H_5 and H_7.

This result is however in contradiction with the former novelty and surprise ratings and the participant feedback after the experiments. Performers were actually perfectly able to identify their own contributions from a shared sound source. In a way it clearly shows the limitations of our value metric, which is far too coarse compared to our abilities to segregate sound sources from monoaural sound mixtures.

8.3.10 Preliminary conclusion

This section concludes the second experiment, in which we comparatively studied musical creativity on scenarios with a different typology and amount of resources

Table 8.8: Experiment 1. Summary of the hypothesis testing results

Experiment 1	
H_1	accepted
H_2	rejected
H_3	accepted
H_4	accepted
H_5	rejected

shared.

This second experiment was based on a performative scenario which, unlike experiment 1, required to incorporate both the audio capture and analysis stages in our creativity computation (see 7.1).

We will now summarize the results obtained for both experiments, and will provide a final comparative discussion of our findings.

8.4 Discussion

In this section we will summarize and discuss the results obtained in the two experiments conducted. We will first review the main findings, and then we will discuss the validity of the results. Finally, we will discuss the suitability and reliability of the proposed Creativity metrics in the context of music performance analysis.

8.4.1 Findings

Summary of the Hypothesis

Tables 8.8 and summarize, respectively, our decisions regarding the hypothesis questions stated for Experiments 1 and 2.

Inconclusive results were both due to non significance of the user ratings -in which case the hypothesis was rejected- and to the low reliability of our value potential metric - in which case we *partially accepted* the hypothesis. This issue will be further addressed in (8.4.2).

We should emphasize, however, that these preliminary conclusions should be confirmed in future studies, as we will discuss later on (8.4.2).

Table 8.9: Experiment 2. Summary of the hypothesis testing results

	Experiment 2
H_1	partially accepted
H_2	rejected
H_3	rejected
H_4	rejected
H_5	partially accepted
H_6	accepted
H_7	partially accepted

User experience

Regarding the analysis of post-test interviews and questionnaires, we draw conclusions on the user experience on our performative environments which may complement the quantitative results we will extract from performance analysis. The main findings are

- **Ensemble playing is more both socially and musically rewarding than playing alone**. While this evidence is well known in respect to single-user instruments (see for example [Sawyer and Others, 2014]), we proved that playing multi-user instruments, no matter their degree of interdependence, is also more highly regarded that individual performance, and performers do not only appreciate their social context, but acknowledge the superior musical possibilities provided by those environments.

- **Keeping exclusive access** (or ownership) **to a parameter control is positively related to performative engagement**. This evidence has implications for design. Exclusive access to controls may be granted (as in Scenarios 1 and 2 for Experiment 1, and Scenarios 1 and 3 for Experiment 2), otherwise it must be negotiated during the performance.

- **Neither personal nor collective awareness seem to be affected by partial or exclusive access to a parameter control nor to exclusive sound channels per instrument**. Regarding shared access to parameter controls, it is known that interdependence scenarios may challenge the level of awareness required for a proper creative engagement in ensemble performances [Weinberg, 2005], therefore our interfaces proved to be sufficiently consistent and unambiguous as to provide enough identity and mutual awareness cues to performers. Once again, this finding proves that by providing enough collective feedback it is possible to compensate for the diminished perceptual strength of players in an interdependent scenario [Booth, 2010].

Performers' behavior

Our preliminary analysis of the performance logs, which compared the activity, co-activity and distribution and negotiation of shared resources in the different scenarios, we may summarize the following relevant findings

- **Sharing resources leads to social interaction**. In particular in Experiment 2 we observed an spontaneous, unnegotiated distribution of *performative resources*, mostly emerging from conflict avoidance in scenarios with competitive access to shared resources.

  Such spontaneous distribution of resources may lead to unexpected strategies: territoriality based in sound generators in fully shared environments, and social games unrelated to the actual musical outcome. Both behaviors do not take advantage of the possibilities for intense interplay provided by fully shared environments and suggest that to achieve a deeper collaboration, concurrent and collaborative operation on the same processes may be enforced (as in Experiment 1) or much longer-term practice may be required, possibly with the assistance of human and or software based conduction strategies.

- Finally, we observed that **performers resort to more salient parameters in shared sound configurations**. This evidence may be considered an example of the natural tendency of performers to keep their own musical voices thoughout the performance. Whenever one particular instrumental feature which is taken for granted to be exclusive of an instrument (such as its own sound emission) becomes shared making the individual contribution less identifiable, the performer resorts to compensating strategies, in this case the use of more explicit timbral modifications which bring back its individual presence.

Creativity

Finally, regarding the analysis of the creativity attributes for the performances, we might summarize the following findings:

- We confirmed that **ensemble performance increases individual creativity when playing with autonomous instruments**. This evidence confirms the transfer of musical ideas, manifested in our scenarios as an increased rate of novel gestures and their magnitude throughout the performance. Such convergence of musical ideas, which suggest the emergence of a shared mental model [Fuller et al., 2010], is at the same time a valuable cue for mutual engagement assessment [Bryan-Kinns, 2012].

- We proved that a **partially shared instrument may benefit less from the increase in collective creativity seen in ensembles with single-user instruments**, but nonetheless it still can provide better opportunities for individual novelty and surprise than individual performances. Lower creativity ratings in our metrics were mainly attributed to the decreased significance of player's actions in terms of its audible outcome (assessed with the value potential metric).

- **In case of competitive access to shared resources** (as in Experiment 2), **we observed a marked decrease in all the creativity attributes**. This results makes us warn once again against such strategy in multi-user instruments.

- **However a fully shared multi-user instrument without exclusive access to resources promotes novel creative behaviors**. This was, at least, the case with the topology explored in Experiment 1 (synchronous, homogeneous, with symmetric control of the same control parameter). This topology results in a unified environment which behaves as a single musical voice. Here the concurrent activity of performers speeds up the pace of musical ideations, detracting users from a thoughtful, purposeful individual performance. However it significantly promotes the emergence of creative collective behaviors, most notably assessed by dynamic collective surprise ratings.

- Our results are inconclusive in respect to the relationship between ensemble creativity and sound ownership (that is, delivering our own generated sound in a specific, closely located speaker), but seem to suggest that **the importance of audio distribution terms of creative engagement might be overestimated**. This evidence comes both from performer's feedback and from the novelty and surprise metrics evaluated in scenarios with and without user-specific audio channels in experiment 2. We will not deny the usefulness of extended meta-instruments in computer and mobile ensembles which resort to localized sound sources. Actually they may help to provide useful cues to the conductor, to the listener and to the performers themselves, and give at the same time a sense of physical presence which may so positively contribute to a more intimate individual and ensemble performance [Trueman and Cook, 2000b]. However, performers are perfectly able to identify individual contributions even from a single shared monoaural sound source, at least for a small sized computer ensemble with a not too homogeneous set of instruments. Indeed, localized sound sources proved to be more challenging in terms of performative involvement because of the risk of unbalance, masking, delays and reverberation related issues inherent in this setup.

1. What is the nature of performers' creative behavior with interdependent multi-user instruments?

2. What are the commonalities and differences, in terms of musical creativity, between a free improvisation with individual instruments and a performance with a multi-user instrument, in the context of a network ensemble?

8.4.2 Validity

Ecological validity

We claim that the two performative scenarios utilized have sufficient ecological validity as to draw conclusions relevant for network music repertoire aimed at computer ensemble performances. Actually, Experiment 1 is a simplified implementation of the principles of weighted interdependence behind Weinberg's *Squeezables* and explore, on an homogeneous setup, the same contexts of interdependence that G. Booth investigates with *Inclusive Interconnections*, a framework employed in performances with the Huddersfield Experimental Laptop Orchestra. Similarly, Experiment 2 explores the very same interdependence topology -but without gestural interfaces- as S.Kotts' *XYZ*, a piece for laptop ensembles which has been performer by the BiLE and Benoit and the Mandelbrots, and which in its turn is based on the HUB's classic *StuckNote*. Well reviewed all this repertoire in Chapter 3.

Generalizability

We conducted two experiments which were carried out by four ensembles of three performers each. This is a too small sample as to draw conclusions with sufficient significance, and should be considered as pilot studies developed to test the reliability of our creativity metrics, but which nonetheless provided useful preliminary findings. Additionally, our performative scenarios studied are limited in scope compared to the large diversity of multi-user topologies available. Further studies on a larger scale and extended to more performative scenarios should be therefore needed.

In particular, we believe a combination of both experiments would provide further insight into the creative opportunities of concurrent actuation in multi-user instruments. Experiment 1 addressed a scenario which minimized concurrent activity. While interdependence extended beyond gestures by holding the state of the sliders and the mapping allowed for a collaborative addition of individual contributions, it was not fully promoting the exploration of concurrent actuation. We partially remedied this shortcoming in Experiment 2, by allowing for concurrent activity in complementary, interrelated parameters of the same sound generator.

However, concurrent actuation on the same slider was not collaborative but competitive, and performers clearly avoided it.

We therefore suggest an additional performative environment which should be addressed to investigate how performers manage to keep distinctive contributions and mutual awareness in an scenario of concurrent actuation on shared controls. Such environment should balance the need for sufficiently long periods of continuous performance to let performers appreciate the combined result of concurrent actuation with the need for a well segmented performance which our creativity metrics do require.

Suitability of the Creativity Metric

One of the aims of the Experiments conducted was to incorporate our Creativity Metric in the analysis of musical performances and evaluate its suitability by complementing it with additional performance measurements and qualitative user based performance evaluation.

In general terms, we found our metric complements well with the other two evaluation techniques employed in our experiments: qualitative assessment though user-based questionnaires and quantitative assessment through preliminary performance analysis.

Both the novelty and the surprise metrics provided useful insight into the creative behavior of the performers both at individual and ensemble level, confirming in many cases their feedback and in some others giving additional clues on the internal processes taking place throughout the ensemble improvisation.

In this respect it is remarkable the assessment of higher individual novelty and surprise ratings in ensemble performances, a result in accordance with Sawyer's theoretical framework on distributed creativity [Sawyer and DeZutter, 2009]. Another relevant finding was the detection of increased inter-player coordination by dynamic surprise assessment in Experiment 1/Scenario 4, showing that our creativity metrics may provide indirect insight into the complex phenomena of ensemble improvisation.

On the other side, by seeing the limitations of our proposed value potential metric, it is debatable its usefulness. The main advantage of the metric in this respect is that it sees the instrument as a *black box*, without requiring any knowledge of the internal mappings and topology of interdependence to evaluate the correspondence between manipulative and audible ideations. The ratings it provide in performances with interdependent multi-user instruments, however, seem to be closely related to their interdependence metric, which would be equally reliable in this context, provided we have an accurate knowledge of the topology of interdependence of each configuration. When audio sources are mixed.

For example, value in Experiment 1 was constant in the two first independent scenarios and then gradually decreased in scenarios three and four. In an inverse fashion, the interdependence metric goes from 0 in the first scenarios to $1/4$ and $1/3$ in the other two scenarios (see 8.1). However the relationship is less clear in Experiment 2 because of the poor performance of the value metric with mixed sound sources (see 8.14).

8.5 Conclusion

In this chapter, we incorporated our proposed creativity metrics in the analysis of musical performances with computer ensembles. Two distinct multi-user instruments were specifically developed to evaluate the musical creativity of performers in an improvisational context.

We first chose our performative context to be a free, unconducted improvisation on a typical network-music environment, as an optimal context to assess musical creativity within a computer ensemble environment.

We then defined the requirements for the multi-user instruments to be designed for the experiments, which should consist of a centralized setup which hosted the sound generating devices, controlled through graphical user interfaces and default input devices (only the mouse) and aimed at a co-located and synchronous interaction.

Given those requirements, the two experiments explored two complementary designs which explored the remaining dimensions of the Collaborative Dimension Space (see [Hattwick and Wanderley, 2012]) not yet determined: for Experiment 1, we addressed the assessment of creativity in a conversational and collaborative improvisation with a rather homogeneous set of instruments, by varying the degree of interdependence of a single control parameter, while Experiment 2 was focused on overlapping, competitive improvisation in a more heterogeneous instrument set, by varying the amount of multiple interdependent parameters.

We finally summarized our findings in both experiments and discussed their implications. This experimental study served us to delimit the strong points and limitations of our creativity metric, and provided us valuable insight into the performer's creative behavior in contexts of interdependence typically found in multi-user instruments.

In the next chapter, we will conclude this dissertation by summarizing our findings in the literature, in our artistic research and in our experimental research throughout this dissertation. Finally, we will summarize the most relevant contributions of this research and will close with proposals for future work.

Chapter 9

CONCLUSIONS

In this Chapter, we will discuss the challenges and opportunities provided by digital multi-user instruments for creative performance in computer music ensembles. We will summarize our key findings related to the literature review on the subject, our artistic praxis and our empirical work. We will then list our contributions and we will finally close this thesis providing proposals for future work.

9.1 Introduction

This dissertation was based in our assumption that computer music ensembles provide an environment particularly well suited both for artistic research and for research on musical creativity in general. We showed how both approaches are mutually complementary:

- Through our research into the artistic process, we conceived, developed and discussed novel performative paradigms, identifying their unique creative opportunities 4.2.1.

- Through the research into musical creativity, we gained further insight into the requirements, challenges and opportunities provided by such performative paradigms 8.4.

Our research into musical creativity was informed by a creativity metric we developed for this purpose. We focused our research in multi-user instruments for computer music ensembles, the most paradigmatic contribution of Network Music (3.2.3), the most explored environment in our artistic praxis (see 4.1) and the subject of our recent research [Comajuncosas et al., 2011], [Comajuncosas and Guaus, 2014].

9.2 Research questions revisited

This thesis was aimed at answering the following research question:

What are the challenges and opportunities provided by digital multi-user instruments for creative performance in computer music ensembles?

Additionally, we sought to answer these two subsidiary research questions as well:

1. What is the nature of performers' creative behavior with interdependent multi-user instruments?

2. What are the commonalities and differences, in terms of musical creativity, between a free improvisation with individual instruments and a performance with a multi-user instrument, in the context of a network ensemble?

We addressed the main research question through Chapters 3,4 and 8. We have studied both subsidiary research questions in the two Experiments conducted and presented in Chapter 8.

We will now summarize our findings related to the main research question, aimed at characterizing the creative challenges and opportunities of multi-user instruments in computer music ensembles. Our findings arose from the literature review and from our artistic praxis.

After that, we will summarize the answers to the subsidiary research questions by collecting together the relevant findings from Chapter 8.

9.2.1 Findings related to the literature review

Our review on computer ensembles provided insight the role of scientific and technical research in the genesis of a new ensemble identity. We summarize here the most relevant findings.

- Through a survey of the NIME literature concerning computer music ensembles, we quantitatively determined a peak in the 2012-2014 period (3.3.2), with research topics centered on compositional strategies (3.3.3), infrastructure and design methodologies (3.3.2) and analysis and frameworks for collaborative interfaces (3.3.2).

- Through an analysis of the performative approaches and repertoire from computer ensembles, we identified two well delimited, though often overlapping ensemble models, which we tentatively called HUB and LOrk models respectively (3.3.2), and provided a characterization for both of them (see 3.3.2 and 3.3.2)

- Through the study of multi-user instruments in the repertoire of computer ensembles, we identified three approaches related to the musical resources shared (3.3.3): sharing musical structures, control parameters, and audio objects. The former literature on the subject did not make explicit distinction on the resources actually shared on an interdependent network, and were mostly concerned on parameter control sharing (3.2.3).

9.2.2 Findings related to the artistic praxis

We provide here a summary of findings related to our artistic practice and research within the Barcelona Laptop Orchestra

- With respect to score-based repertoire, we found a notable advantage, in terms of creative engagement, of open notation forms loosely framing the performance while at the same time provisioning space for incremental learning and development of improvisational strategies (4.2.2). By the other side, delegating unambiguous manipulative commands to software automation allows performers to focus on their creative contribution (4.2.2). We found, in our real-time scoring environment (4.2.2), that the ability to precise control musical features hardly compensates the lack of immediacy of direct manipulative performances. Taken to the extreme, offline musical scoring in a performance setting is best to be avoided for the very same reasons (4.2.2).

- With respect to collective sequencing environments, we found that multi-threaded sequencing environments benefit from strategies aimed to provide better collective awareness, such as physical placement, individual audio channels and individual spaces in shared projection screens. Projecting others' contributions let performers share musical contents more efficiently (4.2.3 and 4.2.3). Resorting only to real-time and public manipulations allows for better immediacy and mutual awareness (4.2.3). *Interleaved multi-threading* is a novel collective sequencing based performative context which allows to provide space for individualized contributions on a single shared musical sequence (4.2.3).

- Regarding interdependent networks, we sought how such topologies may give rise to emergent, unexpected collective behaviors (4.2.5 and 4.2.5). A balance is to be found between the unplanned unexpectedness which arises from such dynamics, which in its turn may encourage creative practice through continuous reinterpretation of the musical context, and the decreased collective agency that such dynamics could cause, which may have actually inhibit creativity.

- Additionally, caution should be exercised in the use of strategies that enhance social interaction which do not (only) provide competence for shared resources, as it may promote a *game-like behavior* (4.2.3). While incorporating game-like strategies to promote engagement in collective musical interfaces have been explored [Blaine and Forlines, 2002], we must determine which is the main driving force of the performance, whether the music itself or the visuals and the game dynamics (3.3.3)

- Further findings related to our implementation of software-mediated or software-assisted conduction techniques for multi-user instruments:

  - **Predictable constraints** help to frame and stimulate creativity, while dynamically changing constraints promote a passive tracking behavior (4.2.4)
  - A **loosely conducted scheme** addressed to a particular musical dimension combined with additional degrees of freedom in other dimensions let performers to establish parallel, complementary creative dynamics (4.2.4).
  - Software agents for **conductor assistance** in multi-user environments let coordinate a performance while providing a novel bidirectional creative flow between conductor and performers (4.2.5).

9.2.3 Experimental findings

Regarding the first subsidiary research question

1. What is the nature of performers' creative behavior with interdependent multi-user instruments?

We provided evidence that

- Performers resort to more salient parameters in shared sound configurations. Sharing resources, notably if there is competitive access to resources, may lead to socially-centered interaction.

- Exclusive access to parametric controls is positively related in terms of performative engagement.

- A fully shared multi-user instrument without exclusive access to resources promotes novel creative behaviors.

Regarding the second subsidiary research question

2. What are the commonalities and differences, in terms of musical creativity, between a free improvisation with individual instruments and a performance with a multi-user instrument, in the context of a network ensemble?

We provided evidence that

- Social and musical interplay provided by multi-user instruments is regarded, at least, as much rewarding that individual instruments in terms of creative engagement.
- Ensemble performance increases individual creativity when playing with autonomous instruments.
- Partially shared instrument benefit less from the increase in collective creativity seen in ensembles with single-user instruments.

9.3 Contributions

Related to the practical outcomes provided within the context of this research, we summarize now our contributions

9.3.1 Interdependence metric

We proposed a further classification criteria for interdependent networks accounting for the number and typology of resources shared and the amount of interdependence in such resources, utilized both in our survey of Laptop Orchestra repertoire (3.5) and the survey (4.1) and analysis (4.27) of our own works and multi-user instruments.

A **interdependence metric** was proposed (3.1) to summarize both attributes, based on Booth's matrix representation for interdependent systems [Booth, 2010], which will be suggested as an alternative metric to assess value potential in our experiments (8.4.2).

9.3.2 Network Music repertoire

An exhaustive list of practical contributions related to the repertoire developed through these years of artistic research would too long to be summarized here. We will provide the most relevant and, to our knowledge, unique performative paradigms which we developed and explored through our artistic practice:

- An in-depth exploration of a co-located asynchronous environment for ensemble performance, La Roda (4.2.5). This specific multi-user configuration has hardly been utilized in performative contexts, as noted by Barbosa [Barbosa, 2003] and Xambó [Xambó, 2015].

- The mapping techniques to facilitate target-based gestural navigation in concatenative synthesis (4.2.4).

- The use of machine listening and machine learning tools to assist in the conduction of multi-user instruments (see 4.2.5) is, to our knowledge, a novel contribution, which extends the proposals of R.Friebrink -see [Fiebrink et al., 2009a] and [Fiebrink, 2011]- to the context of interdependent networks.

- The novel paradigms of time-interleaved multi-threaded collective sequencer (4.2.3) and homogeneous environment developed for Experiment 1 (8.2.6).

9.3.3 Creativity metrics

- The whole concept is new and is based on the assumption that a musical improvisation can be characterized as a sequence of gestural ideations (7.2.1) based on the literature on the subject (6.4.3), and taking a performative gesture as the unit of ideation (7.2.2) upon which we built our creativity metrics.

- The definition of mean creativity attributes for a conceptual space (7.3.5) and the visualization of the temporal evolution of creativity attributes in a performance (7.3.6) as well as the visualization of the mean creativity metrics for whole performances on three-dimensional space plots (8.2.9 and 8.3.9).

- The use of dimensionality reduction techniques (PCA) to build the conceptual spaces (7.3.4)

- The use of GMM based novelty assessment in the context of creativity studies (7.3.4 and 7.3.5).

- An structural surprise assessment with a Hellinger distance metric (7.5) between two GMM probability density functions (7.3.5)

- The extension of the linear regression strategy proposed by Maher [Maher et al., 2013] for local (dynamic) surprise assessment incorporating confidence intervals (7.3.5)

- An heuristic-based assessment of dynamical surprise for outliers (7.3.5)

- A comparison of creativity ratings for individual, ensemble and interdependent contexts (8.2)
- A comparison of creativity ratings for contexts of variable interdependence (8.2 and 8.3)

9.4 Critique

In this dissertation we presented a novel methodology to assess musical creativity in a performance. While our creativity metric was already used in two conducted experiments -together with some other well established evaluation techniques- to shed light on the research questions addressed, there are still opportunities for enhancing it.

We provide here a summary of the most relevant shortcomings that should be addressed to draw more reliable and significant conclusions from experimental research utilizing our creativity metrics.

9.4.1 Small sample size for the experiments

As discussed in the previous Chapter (8.4.2), the sample size for the two experiments was too small to guarantee sufficient significance for our experimental results. While we took this into account when evaluating the user-based questionnaires, we believe that carrying out similar experiments on a larger scale would greatly help to validate our preliminary findings.

9.4.2 Weight of creativity attributes

Our Creativity metrics is based on Maher's computational methodology, which does not take into consideration the study of the relative importance of the three creativity attributes (novelty, surprise and value) in the provision of a summarizing creativity rating [Maher and Fisher, 2012]. This is the reason why we evaluated our creativity attributes separately in our experiments, a decision justified in 7.3.6. One possible methodology to determine the importance of our creativity attributes in the evaluation of a creative performance would be the one applied by Jordanous [Jordanous, 2011] in her evaluation of improvisation systems, where the different creativity attributes where weighted according to user ratings.

9.4.3 Orthogonality of the creativity attributes

We assumed our three creativity attributes (novelty, surprise and value potential) to be independent, and indeed we used three-dimensional plots to visualize the creativity of individual performances, an approach already followed by Maher in her comparative evaluation of creative artifacts [Maher and Fisher, 2012]. However, as we noticed in, novelty and structural surprise seem to be strongly correlated.

9.4.4 Reliability of the value potential metric

As pointed out in 7.3.7 and in 8.4.2, our proposed metric to assess the potential for value has severe limitations. Besides, value assessment will always be inextricably linked to social evaluation. While it proved to be valuable to determine the relationship between degree of interdependence and dissimilarity between performative and audible gestures, notably in Experiment 1, it is debatable its validity in other contexts and should be most possibly redefined.

9.4.5 Narrow scope

Finally, our assumption that a free improvisation (discussed in 2.4.2) may be characterized of a sequence of discrete gestures addressable as semantic units lies at the basis of our gesture-based creativity metric, as discussed in 7.2. While this assumption was based on our literature survey on cognitive models of musical improvisation (6.4.3), it is not only an oversimplification of such a complex phenomena, in fact other improvisational practices commonly seen in computer ensembles, such as live-coding or higher level control of interactive systems, cannot be addressed by our Creativity metric because they are just not based on our assumption of direct gestural control at all. We already discussed that limitation in 7.3.7.

9.5 Future Work

We will finally provide a RoadMap suggesting the immediate future work related to the Computational Creativity Metric proposed.

As discussed in 7.3.7, there is much room for improvement. Going back to the issues discussed there, we might tentatively propose the immediate lines of action to be centered on

- Extending the performance analysis to multidimensional gestural data capture
- Reformulating or getting rid of the structural surprise attribute
- Redefining the value metric

- Resorting to user-based (maybe expert-based) evaluation to calibrate the creativity attributes.

Beyond improving the reliability of the metric proposed, we believe further research should consider the possibility of extending it to areas which have not been covered in this research:

- **Extend the study to our own repertory** The logical next step of this research is to incorporate our proposed Creativity metrics into the study of our own repertory. As already discussed (8.1), some of our performative environments are already well suited in this regard, particularly those resorting to continuous, gestural control such as the Nuvolet (4.2.4) and La Roda (4.2.5).

- **Use the metrics as a feedback within the performance itself** Finally, in the line of the recent contributions by Charles Martin et al. [Martin et al., 2015], computational creativity might assessment be incorporated to monitor the performance in real-time, feeding back the analysis of the creativity of a performance in the performance itself. Indeed our creativity metric is conceived to allow for a continuous evaluation of the performance, and it might serve as a tool aiding the conductor in the monitoring of the performance, or as a valuable feedback for performers to encourage a more active engagement in the performance.

Bibliography

[Amabile, 1983] Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of personality and social psychology*, 45(2):357.

[Andean, 2012] Andean, J. (2012). An Embodied approach to acousmatic music. *Monitieteinen musiikintutkimus*, pages 80–85.

[Arfib et al., 2002] Arfib, D., Couturier, J. M., Kessous, L., and Verfaille, V. (2002). Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Organised Sound*, 7(02):127–144.

[Ashley, 2002] Ashley, R. (2002). Do [n't] change a hair for me: The art of jazz rubato. *Music Perception*, 19(3):311–332.

[Assayag and Dubnov, 2004] Assayag, G. and Dubnov, S. (2004). Using factor oracles for machine improvisation. *Soft Computing*, 8(9):604–610.

[Babbitt and Peles, 2003] Babbitt, M. and Peles, S. (2003). *The collected essays of Milton Babbitt*. Princeton University Press.

[Baer and McKool, 2009] Baer, J. and McKool, S. S. (2009). Assessing creativity using the consensual assessment. *Handbook of Research on Assessment Technologies, Methods, and Applications in Higher Education*, pages 1–13.

[Baguyos, 2014] Baguyos, J. C. (2014). Contemporary Practices in the Performance and Sustainability of Computer Music Repertoire. (September):99–102.

[Bailey, 1993] Bailey, D. (1993). *Improvisation: Its nature and practice in music*. Da Capo Press.

[Barbosa, 2003] Barbosa, Á. (2003). Displaced Soundscapes: A Survey of Network Systems for Music and Sonic Art Creation. *Leonardo Music Journal*, 13:53–59.

[Barbosa et al., 2011] Barbosa, J., Calegario, F., Magalhães, F., and Cabral, G. (2011). Towards an evaluation methodology for digital music instruments considering performer ' s view : a case study. In *Proceedings of 13th Brazilian Symposium on Computer Music*.

[Barbosa et al., 2012] Barbosa, J., Calegario, F., Teichrieb, V., Ramalho, G., and Mcglynn, P. (2012). Considering Audience ' s View Towards an Evaluation Methodology for Digital Musical Instruments. *NIME 2012 Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 403–408.

[Barbosa et al., 2015] Barbosa, J., Malloch, J., Huot, S., and Wanderley, M. M. (2015). What does 'Evaluation' mean for the NIME community? *New Interfaces for Musical Expression (NIME)*, pages 156–161.

[Beaton and Tech, 2010] Beaton, B. and Tech, V. (2010). Digital Drumming : A Study of Co-located , Highly Coordinated , Dyadic Collaboration. *Science*, pages 1417–1426.

[Beck et al., 2011] Beck, S., Branton, C., Maddineni, S., Ullmer, B., and Jha, S. (2011). Tangible Performance Management of Grid-based Laptop Orchestras. *Proceedings of the International Conference on New Interfaces for Musical Expression*, (June):207–210.

[Benson, 2003] Benson, B. E. (2003). *The improvisation of musical dialogue: a phenomenology of music*. Cambridge University Press.

[Benson, 2008] Benson, D. (2008). Music : A Mathematical Offering. (December).

[Bergee, 1997] Bergee, M. J. (1997). Relationships among faculty, peer, and self-evaluations of applied performances. *Journal of Research in Music Education*, 45(4):601–612.

[Bergee, 2007] Bergee, M. J. (2007). Performer, rater, occasion, and sequence as sources of variability in music performance assessment. *Journal of research in music education*, 55(4):344–358.

[Berlin, 2014] Berlin, R. S. (2014). Free Improvisation and Nothing : From the Tactics of Escape to a Bastard Science.

[Berliner, 2009] Berliner, P. F. (2009). *Thinking in jazz: The infinite art of improvisation*. University of Chicago Press.

[Besemer and O'Quin, 1987] Besemer, S. P. and O'Quin, K. (1987). Creative product analysis: Testing a model by developing a judging instrument. *Frontiers of creativity research: Beyond the basics*, pages 367–389.

[Bevacqua et al., 2007] Bevacqua, E., Mancini, M., Niewiadomski, R., and Pelachaud, C. (2007). An expressive ECA showing complex emotions. In *Proceedings of the AISB annual convention, Newcastle, UK*, pages 208–216.

[Bevilacqua et al., 2005] Bevilacqua, F., Müller, R., and Schnell, N. (2005). MnM: a Max/MSP mapping toolbox. In *Proceedings of the 2005 conference on New interfaces for musical expression*, pages 85–88. National University of Singapore.

[Bickerman et al., 2010] Bickerman, G., Bosley, S., Swire, P., and Keller, R. (2010). Learning to create jazz melodies using deep belief nets. In *First International Conference on Computational Creativity*.

[Birnbaum et al., 2005] Birnbaum, D., Fiebrink, R., Malloch, J., and Wanderley, M. M. (2005). Towards a Dimension Space for Musical Devices. In *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression (NIME-05)*, pages 192–195.

[Bischoff et al., 1978] Bischoff, J., Gold, R., and Horton, J. (1978). Music for an interactive network of microcomputers. *Computer Music Journal*, pages 24–29.

[Bishop, 2006] Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

[Bithell and Hill, 2014] Bithell, C. and Hill, J. (2014). *The Oxford Handbook of Music Revival*. Oxford University Press.

[Blaine and Fels, 2003] Blaine, T. and Fels, S. (2003). Contexts of Collaborative Musical Experiences. In *Proceedings of the 2003 conference on New interfaces for musical expression*, pages 129–134. National University of Singapore.

[Blaine and Forlines, 2002] Blaine, T. and Forlines, C. (2002). JAM-O-WORLD : Evolution of the Jam-O-Drum Multi-player Musical Controller into the Jam-O-Whirl Gaming Interface. pages 1–6.

[Boden, 2004a] Boden, M. A. (2004a). *The creative mind: Myths and mechanisms*. Psychology Press.

[Boden, 2004b] Boden, M. A. (2004b). *{T}he {C}reative {M}ind: {M}yths and {M}echanisms*. {R}outledge.

[Booth, 2010] Booth, G. (2010). *Inclusive Interconnections: Towards Open-Ended Parameter-Sharing for Laptop Ensemble*. Master thesis, University of Huddersfield.

[Booth and Gurevich, 2012] Booth, G. and Gurevich, M. (2012). Collaborative composition and socially constructed instruments : Ensemble laptop performance through the lens of ethnography. In *NIME 2012 Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 61–65.

[Boulanger, 2000] Boulanger, R. C. (2000). *The Csound book: perspectives in software synthesis, sound design, signal processing, and programming*. MIT press.

[Bowen, 2012] Bowen, N. (2012). 4Quarters: Real-time Collaborative Music Environment for Mobile Phones. In *Proceedings of the 1st Symposium on Laptop Ensembles & Orchestras (SLEO)*, pages 83–87.

[Brandt and Brandt, 2005] Brandt, L. and Brandt, P. A. (2005). Making sense of a blend: A cognitive-semiotic approach to metaphor. *Annual Review of Cognitive Linguistics*, 3(1):216–249.

[Bresin and Friberg, 2013] Bresin, R. and Friberg, A. (2013). Evaluation of computer systems for expressive music performance. In *Guide to Computing for Expressive Music Performance*, pages 181–203. Springer.

[Bresin and Umberto Battel, 2000] Bresin, R. and Umberto Battel, G. (2000). Articulation Strategies in Expressive Piano Performance Analysis of Legato, Staccato, and Repeated Notes in Performances of the Andante Movement of Mozart's Sonata in G Major (K 545). *Journal of New Music Research*, 29(3):211–224.

[Brown and Sorensen, 2007] Brown, A. R. and Sorensen, A. (2007). Dynamic media arts programming in impromptu. In *Proceedings of the 6th ACM SIGCHI conference on Creativity & Cognition*, pages 245–246. ACM.

[Brown, 2012] Brown, D. C. (2012). Creativity, surprise & design: An introduction and investigation. *The 2nd international conference on design creativity (* . . . , (September):1–8.

[Bryan-Kinns, 2004] Bryan-Kinns, N. (2004). *Computers in Support of Muscial Expression*. Queen Mary University, London.

[Bryan-Kinns, 2012] Bryan-Kinns, N. (2012). Mutual engagement in social music making. *Intelligent Technologies for Interactive Entertainment*, 44(0):1–6.

[Bryan-Kinns, 2013] Bryan-Kinns, N. (2013). Mutual engagement and collocation with shared representations. *International Journal of Human-Computer Studies*, 71(1):76–90.

[Bryan-Kinns and Hamilton, 2012] Bryan-Kinns, N. and Hamilton, F. (2012). Identifying mutual engagement. *Behaviour & Information Technology*, 31(2):101–125.

[Bryan-Kinns and Healey, 2006] Bryan-Kinns, N. and Healey, P. (2006). Decay in Collaborative Music Making. *Interfaces*, (August):114–117.

[Burns and Surges, 2008] Burns, C. and Surges, G. (2008). NRCI : SOFTWARE TOOLS FOR LAPTOP ENSEMBLE Christopher Burns and Greg Surges Department of Music. *International Computer Music Conference*.

[Caramiaux et al., 2010a] Caramiaux, B., Bevilacqua, F., and Schnell, N. (2010a). Analyzing Gesture and Sound Similarities With a Hmm-Based Divergence Measure. *Smc'10*, (section 3):59–66.

[Caramiaux et al., 2010b] Caramiaux, B., Bevilacqua, F., and Schnell, N. (2010b). Towards a gesture-sound cross-modal analysis. pages 158–170.

[Cardew, 1971] Cardew, C. (1971). *Treatise handbook: including Bun no. 2 [and] Volo solo*. Edition Peters.

[Caridakis et al., 2006] Caridakis, G., Raouzaiou, A., Karpouzis, K., and Kollias, S. (2006). Synthesizing gesture expressivity based on real sequences. In *Workshop "Multimodal Corpora. From Multimodal Behaviour Theories to Usable Models". In: 5th international conference on language resources and evaluation (LREC'2006), Genova, Italy*, pages 19–23.

[Cascone, 2002] Cascone, K. (2002). Laptop music-counterfeiting aura in the age of infinite reproduction. *Parachute*, pages 52–59.

[Cerqueira, 2010] Cerqueira, M. (2010). *Synchronization over Networks for Live Laptop Music Performance*. PhD thesis.

[Cerqueira et al., 2013] Cerqueira, M., Salazar, S., and Wang, G. (2013). SoundCraft: Transducing StarCraft 2. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 243–247, Daejeon, Republic of Korea. Graduate School of Culture Technology, KAIST.

[Chadabe, 1997] Chadabe, J. (1997). Electric Sound:{The} Past and Promise of Electronic Music.

[Chaffin et al., 2006] Chaffin, R., Lemieux, A. F., and Chen, C. (2006). 11 Spontaneity and creativity in highly practised performance. *Musical Creativity*, page 200.

[Clarke, 2004] Clarke, E. (2004). Empirical methods in the study of performance. *Empirical musicology: Aims, methods, prospects*, pages 77–102.

[Clarke, 2005a] Clarke, E. (2005a). Expression in performance: generativity, perception and semiosis. *The practice of performance: Studies in musical interpretation*, page 21.

[Clarke, 1988] Clarke, E. F. (1988). Generative principles in music performance.

[Clarke, 2005b] Clarke, E. F. (2005b). Creativity in performance. *Musicae Scientiae*, 9(1):157–182.

[Clarke, 2012] Clarke, E. F. (2012). Creativity in performance. *Musical Imaginations: Multidisciplinary perspectives on creativity, performance and perception*, pages 17–30.

[Clarke and Windsor, 2000] Clarke, E. F. and Windsor, W. L. (2000). Real and simulated expression: A listening study. *Music Perception*, pages 277–313.

[Cohen, 2012] Cohen, J. E. (2012). Decentering Creativity A Decentered Model of Creativity.

[Collins and D'Escriván, 2007] Collins, N. and D'Escriván, J. (2007). *The Cambridge companion to electronic music*. Cambridge University Press.

[Collins et al., 2003] Collins, N., Mclean, A., Rohrhuber, J., and Ward, A. (2003). Live coding in laptop performance. *Organised Sound*, 8(3):321–329.

[Comajuncosas and Guaus, 2014] Comajuncosas, J. and Guaus, E. (2014). Conducting Collective Instruments : A Case Study. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 513–516.

[Comajuncosas et al., 2011] Comajuncosas, J. M., Barrachina, A., and Connell, J. O. (2011). Nuvolet : 3D Gesture-driven Collaborative Audio Mosaicing. *Audio*, (June):2–5.

[Cope, 2005] Cope, D. (2005). *Computer models of musical creativity*. MIT Press Cambridge.

[Corbin and Strauss, 2014] Corbin, J. and Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage publications.

[Crab, 2004] Crab, S. (2004). 120 Years of Electronic Music.

[Cropley, 2005] Cropley, A. (2005). Engineering Creativity: A Systems Concept of Functional Creativity.

[Csikszentmihalyi, 1999] Csikszentmihalyi, M. (1999). 16 Implications of a Systems Perspective for the Study of Creativity. *Handbook of creativity*, page 313.

[Csikszentmihalyi, 2014] Csikszentmihalyi, M. (2014). *Society, culture, and person: A systems view of creativity*. Springer.

[Daniel, 2001] Daniel, R. (2001). Self-assessment in performance. *British Journal of Music Education*, 18(03):215–226.

[Dannenberg, 2012] Dannenberg, R. (2012). Laptop orchestra communication using publish-subscribe and peer-to-peer strategies. In *Proceedings of the 1st Symposium on Laptop Ensembles and Orchestras*, pages 88–93.

[Davidson and Coulam, 2006] Davidson, J. and Coulam, A. (2006). Exploring jazz and classical solo singing performance behaviours: A preliminary step towards understanding performer creativity. *I. Deli{è}ge & G. A. Wiggins, Musical creativity: Multidisciplinary research in theory and practice*, pages 181–199.

[Davidson and Coimbra, 2001] Davidson, J. W. and Coimbra, D. D. C. (2001). Investigating performance evaluation by assessors of singers in a music college setting. *Musicae Scientiae*, 5(1):33–53.

[Davis, 2009] Davis, S. (2009). Bring out the counterpoint: exploring the relationship between implied polyphony and rubato in Bach's solo violin music. *Psychology of Music*, 37(3):301–324.

[de Campo, 2013] de Campo, A. (2013). Republic: Collaborative Live Coding 2003-2013. *Collaboration and learning through live coding*, pages 152–153.

[Derbinsky and Essl, 2012] Derbinsky, N. and Essl, G. (2012). Exploring reinforcement learning for mobile percussive collaboration. *Ann Arbor*, 1001:48109–2121.

[Dixon et al., 2002] Dixon, S., Goebl, W., and Widmer, G. (2002). The performance worm: Real time visualisation of expression based on langner's tempo-loudness animation. In *Proceedings of the international computer music conference (icmc 2002)*.

[Dobrian, 1991] Dobrian, C. (1991). Thoughts on Composition and Improvisation.

[Dubnov and Assayag, 2005] Dubnov, S. and Assayag, G. (2005). Improvisation planning and jam session design using concepts of sequence variation and flow experience. In *Proc. International Conference on Sound and Music Computing*, pages 24–26.

[Duby, 2006] Duby, M. (2006). *Soundpainting as a system for the collaborative creation of music in performance*. PhD thesis, University of Pretoria.

[Duncan, 2010] Duncan, S. P. (2010). Re-Complexifying the Function (s) of Notation in the Music of Brian Ferneyhough and the New Complexity. *Perspectives of New Music*, pages 136–172.

[Essl et al., 2008] Essl, G., Wang, G., and Rohs, M. (2008). Developments and Challenges turning Mobile Phones into Generic Music Performance Platforms. *Signal Processing*, pages 13–16.

[Fadnes, 2015] Fadnes, P. F. (2015). Improvisational Conduct and Case Studies from the Margins. *The Cultural Politics of Jazz Collectives: This Is Our Music*, page 197.

[Fauconnier and Turner, 1998] Fauconnier, G. and Turner, M. (1998). Conceptual integration networks. *Cognitive science*, 22(2):133–187.

[Fels et al., 2002] Fels, S., Gadd, A., and Mulder, A. (2002). Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, 7(02):109–126.

[Fencott and Bryan-Kinns, 2010] Fencott, R. and Bryan-Kinns, N. (2010). Hey Man, you're invading my Personal Space! Privacy and Awareness in Collaborative Music. In *Proc. NIME 2010*, number June 2008, pages 198–203.

[Fencott and Bryan-Kinns, 2012] Fencott, R. and Bryan-Kinns, N. (2012). Audio delivery and territoriality in collaborative digital musical interaction. *BCS-HCI '12 Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers*, pages 69–78.

[Fiebrink et al., 2009a] Fiebrink, R., Cook, P., Smallwood, S., Trueman, D., and Wang, G. (2009a). Laptop Orchestras and Machine Learning in Real-time Music Performance. *Interfaces*, pages 1–4.

[Fiebrink et al., 2009b] Fiebrink, R., Cook, P. R., and Trueman, D. (2009b). Play-Along Mapping of Musical Controllers. *Icmc 2009*, pages 61–64.

[Fiebrink et al., 2007] Fiebrink, R., Wang, G., and Cook, P. (2007). Don't forget the laptop: using native input capabilities for expressive musical control. *Conference on New Interfaces for Musical Expression (NIME-07)*, page 4.

[Fiebrink, 2011] Fiebrink, R. A. (2011). Real-time Human Interaction with Supervised Learning Algorithms for Music Composition and Performance. (January).

[Finnendahl, 2012] Finnendahl, O. (2012). Technologie und performative Kompetenz. *Auditive Perspektiven*, 2.

[Fletcher, 2004] Fletcher, P. (2004). *World musics in context: a comprehensive survey of the world's major musical cultures*. Oxford University Press.

[Foote, 2000] Foote, J. (2000). Automatic audio segmentation using a measure of audio novelty. *2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No.00TH8532)*, 1:3–6.

[Fran et al., 2013] Fran, J., Caramiaux, B., and Fran, J. (2013). A Hierarchical Approach for the Design of Gesture-to-Sound Mappings To cite this version :.

[Fransella et al., 2004] Fransella, F., Bell, R., and Bannister, D. (2004). *A manual for repertory grid technique*. John Wiley & Sons.

[Freeman and Street, 2008] Freeman, J. and Street, M. (2008). Extreme Sight-Reading , Mediated Expression , and Audience Participation : Real-Time Music Notation in Live Performance. *Computer*, pages 25–42.

[Freeman et al., 2010] Freeman, J., Troyer, A. V., Colella, A., Lee, S. W., and Yao, S. (2010). Lolc (2010).

[Freeman et al., 2003] Freeman, J., Troyer, A. V., and Street, M. (2003). Collaborative Textual Improvisation in a Laptop Ensemble. pages 8–21.

[Friberg, 1991] Friberg, A. (1991). Generative rules for music performance: A formal description of a rule system. *Computer Music Journal*, pages 56–71.

[Friberg et al., 1991] Friberg, A., Fryden, L., Bodin, L.-G., and Sundberg, J. (1991). Performance rules for computer-controlled contemporary keyboard music. *Computer Music Journal*, pages 49–55.

[Friberg and Sundström, 2002] Friberg, A. and Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, 19(3):333–349.

[Fuller et al., 2010] Fuller, D., St, C., and Magerko, B. (2010). Shared Mental Models in Improvisational Performance. *Current*.

[Gabora and Aerts, 2009] Gabora, L. and Aerts, D. (2009). A model of the emergence and evolution of integrated worldviews. *Journal of Mathematical Psychology*, 53(5):434–451.

[Gabrielsson, 2003] Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of music*, 31(3):221–272.

[Gabrielsson and Juslin, 1996] Gabrielsson, A. and Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of music*, 24(1):68–91.

[Gaye et al., 2006] Gaye, L., Holmquist, L. E., Behrendt, F., and Tanaka, A. (2006). Mobile music technology: Report on an emerging community. In *Proceedings of the 2006 conference on New interfaces for musical expression*, pages 22–25. IRCAM—Centre Pompidou.

[Gelineck and Serafin, 2009] Gelineck, S. and Serafin, S. (2009). A quantitative evaluation of the differences between knobs and sliders. *Proceedings of NIME*.

[Gero, 2002] Gero, J. S. (2002). Computational models of creative designing based on situated cognition. *Proceedings of the fourth conference on Creativity & cognition - C&C '02*, pages 3–10.

[Gero and Maher, 2013] Gero, J. S. and Maher, M. L. (2013). *Modeling creativity and knowledge-based creative design*. Psychology Press.

[Ghamsari et al., 2013] Ghamsari, M., Pras, A., and Wanderley, M. M. (2013). Combining musical tasks and improvisation in evaluating novel digital musical instruments. In *Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research, Marseille*, pages 506–515.

[Gibbs, 2010] Gibbs, L. (2010). Evaluating creative (jazz) improvisation: Distinguishing invention and creativity. In *Leeds International Jazz Conference 2010:: Improvisation - jazz in the creative moment*.

[Gillian et al., 2011] Gillian, N., Knapp, R., and O'Modhrain, S. (2011). A machine learning toolbox for musician computer interaction. In *Proc. of NIME*, volume 11.

[Gillian and Paradiso, 2014] Gillian, N. and Paradiso, J. A. (2014). The gesture recognition toolkit. *The Journal of Machine Learning Research*, 15(1):3483–3487.

[Glowinski et al., 2012] Glowinski, D., Badino, L., Ausilio, A., Camurri, A., and Fadiga, L. (2012). Analysis of leadership in a string quartet. *Third International Workshop on Social Behaviour in Music at ACM ICMI 2012*.

[Godøy, 2006] Godøy, R. I. (2006). Gestural-Sonorous Objects: embodied extensions of Schaeffer's conceptual apparatus. *Organised Sound*, 11(02):149–157.

[Goebl and Palmer, 2009] Goebl, W. and Palmer, C. (2009). Synchronization of timing and motion among performing musicians.

[Goebl et al., 2004] Goebl, W., Pampalk, E., and Widmer, G. (2004). Exploring expressive performance trajectories: Six famous pianists play six Chopin pieces. In *Proceedings of the 8th international conference on music perception and cognition*, pages 505–509.

[Goehr, 1992] Goehr, L. (1992). *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music: An Essay in the Philosophy of Music*. Oxford University Press.

[Goldenberg and Mazursky, 2002] Goldenberg, J. and Mazursky, D. (2002). *Creativity in product innovation*. Cambridge University Press.

[Gómez de Silva Garza and Gero, 2010] Gómez de Silva Garza, A. and Gero, J. (2010). Elementary Social Interactions and Their Effects on Creativity: A Computational Simulation. *Proceedings of the International Conference on Computational Creativity*, pages 110–119.

[Goodman, 1968] Goodman, N. (1968). *Languages of art: An approach to a theory of symbols*. Hackett publishing.

[Grace and Maher, 2014] Grace, K. and Maher, M. L. (2014). Using Computational Creativity to Guide Data-Intensive Scientific Discovery. pages 35–38.

[Grace et al., 2014] Grace, K., Maher, M. L., Fisher, D., and Brady, K. (2014). Data-intensive evaluation of design creativity using novelty, value, and surprise. *International Journal of Design Creativity and Innovation*, (ahead-of-print):1–23.

[Grace et al., 2015] Grace, K., Maher, M. L., Fisher, D., and Brady, K. (2015). Modeling Expectation for Evaluating Surprise in Design Creativity. In Gero, J. S. and Hanna, S., editors, *Design Computing and Cognition '14*, pages 189–206. Springer International Publishing.

[Gresham-Lancaster, 1998] Gresham-Lancaster, S. (1998). The aesthetics and history of the hub: The effects of changing technology on network computer music. *Leonardo Music Journal*, 8:39–44.

[Gurevich et al., 2010] Gurevich, M., Stapleton, P., and Marquez-borbon, A. (2010). Style and Constraint in Electronic Musical Instruments. *Proceedings of the International Conference on New Interfaces for Musical Expression*, (Nime):106–111.

[Hajdu et al., 2011] Hajdu, G., Niggemann, K., and Szigetva, A. (2011). Notation in the Context of Quintet . net Projects. *Contemporary Music Review*, (789296666).

[Halkidi et al., 2001] Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2001). On Clustering Validation Techniques. *Journal of Intelligent Information Systems*, 17:107–145.

[Hamilton, 2009] Hamilton, R. (2009). Maps and Legends: Designing FPS-Based Interfaces for Multi-user Composition, Improvisation and Immersive Performance. *Computer Music Modeling and Retrieval. Sense of Sounds*, pages 478–486.

[Hansen et al., 2011] Hansen, A.-M. S., Andersen, H. J., and Raudaskoski, P. (2011). Play Fluency in Music Improvisation Games for Novices.

[Hansen et al., 2012] Hansen, A.-m. S., Andersen, H. J., and Raudaskoski, P. (2012). Two Shared Rapid Turn Taking Sound Interfaces for Novices. In *NIME 2012 Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 470–473.

[Hargreaves et al., 2012] Hargreaves, D., Miell, D., and MacDonald, R. (2012). *Musical imaginations: Multidisciplinary perspectives on creativity, performance and perception.* Oxford University Press.

[Hargreaves and MacDonald, 2012] Hargreaves, D. J. and MacDonald, R. (2012). Explaining musical imaginations: Creativity, performance, and perception. *Musical imaginations: Multidisciplinary perspectives on creativity, performance, and perception*, pages 1–14.

[Hartmann et al., 2005] Hartmann, B., Mancini, M., Buisine, S., and Pelachaud, C. (2005). Design and evaluation of expressive gesture synthesis for embodied conversational agents. *4th International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS'05)*, 2(35):1095–1096.

[Hattwick and Umezaki, 2012] Hattwick, I. and Umezaki, K. (2012). Approaches to Collaboration in a Digital Music Ensemble. In *NIME 2012 Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 466–469.

[Hattwick and Wanderley, 2012] Hattwick, I. and Wanderley, M. (2012). A Dimension Space for Evaluating Collaborative Musical Performance Systems. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, Ann Arbor, Michigan. University of Michigan.

[Hemment, 2006] Hemment, D. (2006). Locative Arts. *Leonardo*, 39(4):348—-355.

[Hennessey and Amabile, 1999] Hennessey, B. A. and Amabile, T. M. (1999). Consensual assessment. *Encyclopedia of creativity*, 1:347–359.

[Hennessey and Amabile, 2010] Hennessey, B. a. and Amabile, T. M. (2010). Creativity. *Annual review of psychology*, 61:569–598.

[Hewitt, 2014] Hewitt, S. (2014). Notational Approaches for Composing and Directing a Non-Homogeneous Laptop Orchestra.

[Hewitt and Tremblay, 2012] Hewitt, S. and Tremblay, P. (2012). Notational approaches for laptop ensembles. ... *of the Symposium on Laptop Ensembles ...*, pages 3–6.

[Hewitt et al., 2010] Hewitt, S., Tremblay, P. A., Freeman, S., and Booth, G. (2010). HELO: The Laptop Ensemble as an Incubator for Individual Laptop Performance Practices. *Proceedings of the International Computer Music Conference*, 2010:304–307.

[Hill, 2011] Hill, J. (2011). belief systems encourage or inhibit creative activities in music. *Musical Imaginations: Multidisciplinary perspectives on creativity, performance and perception: Multidisciplinary perspectives on creativity, performance and perception*, page 87.

[Hodson, 2007] Hodson, R. (2007). *Interaction, Improvisation, and Interplay in Jazz*. Routledge, 1st editio edition.

[Holmes, 2012] Holmes, T. (2012). *Electronic and experimental music: technology, music, and culture*. Routledge.

[Horn and Salvendy, 2006] Horn, D. and Salvendy, G. (2006). Product creativity: conceptual model, measurement and characteristics. *Theoretical Issues in Ergonomics Science*, 7(4):395–412.

[Horowitz, 1995] Horowitz, D. (1995). Representing Musical Knowledge in a Jazz Improvisation System. *Proc. of Artificial Intelligence and Music*, pages 16–23.

[Horvitz et al., 2005] Horvitz, E., Apacible, J., Sarin, R., and Liao, L. (2005). Prediction, expectation, and surprise: Methods, designs, and study of a deployed traffic forecasting service. *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence (UAI '05)*.

[Howard, 2012] Howard, S. A. (2012). The effect of selected nonmusical factors on adjudicators' ratings of high school solo vocal performances. *Journal of Research in Music Education*, page 0022429412444610.

[Humair and Genève., 1999] Humair, J.-D. and Genève., D. (1999). *Variations systématiques et types rythmiques: une approche musicologique du domaine microtemporel dans l'interpr{é}tation des styles musicaux*. Universit{é} de Gen{è}ve Facult{é} des lettres.

[Hunt and Wanderley, 2003] Hunt, A. and Wanderley, M. M. (2003). Mapping performer parameters to synthesis engines. *Organised Sound*, 7(02):97–108.

[Hunt et al., 2003] Hunt, A., Wanderley, M. M., and Paradis, M. (2003). The importance of parameter mapping in electronic instrument design. *Journal of New Music Research*, 32(4):429–440.

[Husserl et al., 1973] Husserl, E., AMERIKS, K., CHURCHILL, J. S., Eley, L., and Landgrebe, L. (1973). *Experience and Judgment: Investigations in a Genealogy of Logic. Revised and Edited by Ludwig Landgrebe. Translated by James S. Churchill and Karl Ameriks. Introduction by James S. Churchill. Afterword by Lothar Eley*. Northwestern University Press.

[Iazzetta, 2000] Iazzetta, F. (2000). Meaning in musical gesture.

[Isaacson, 1959] Isaacson, L. A. H. J. L. M. (1959). Experimental music. *New York*.

[Itti and Baldi, 2009] Itti, L. and Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, 49(10):1295–1306.

[Jacquemin et al., 2007] Jacquemin, C., Ajaj, R., Cahen, R., Olivier, Y., and Schwarz, D. (2007). Plumage: Design d'une interface 3D pour le parcours d'echantillons sonores granularises. In *Proceedings of the 19th International Conference of the Association Francophone d'Interaction Homme-Machine*, pages 71–74. ACM.

[Jain, 2010] Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern recognition letters*, 31(8):651–666.

[Jarvinen, 1995] Jarvinen, T. (1995). Tonal hierarchies in jazz improvisation. *Music Perception*, pages 415–437.

[Järvinen and Toiviainen, 2000] Järvinen, T. and Toiviainen, P. (2000). The effect of metre on the use of tones in jazz improvisation. *Musicae Scientiae*, 4(1):55–74.

[Jensenius, 2007] Jensenius, A. R. (2007). Action-sound: Developing methods and tools to study music-related body movement.

[Johnson, 1996] Johnson, C. M. (1996). The performance of Mozart: Study of rhythmic timing by skilled musicians. *Psychomusicology: A Journal of Research in Music Cognition*, 15(1-2):87.

[Johnson, 1999] Johnson, C. M. (1999). The performance of Bach: Study of rhythmic timing by skilled musicians. *Bulletin of the Council for Research in Music Education*, pages 66–71.

[Johnson, 2003] Johnson, C. M. (2003). Effect of rubato magnitude on the perception of musicianship in musical performance. *Journal of Research in Music Education*, 51(2):115–123.

[Johnson-Laird, 2002] Johnson-Laird, P. N. (2002). How Jazz Musicians Improvise. *Music Perception*, 19(3):415–442.

[Jordà, 2004a] Jordà, S. (2004a). Digital Instruments and Players : Part I – Efficiency and Apprenticeship. *Proceedings of the 2004 Conference on New Interfaces for Musical Expression*, (September):59–63.

[Jordà, 2004b] Jordà, S. (2004b). Digital Instruments and Players: Part II–Diversity, Freedom and Control. *Proceedings of the International Computer Music Conference*, pages 706–710.

[Jordà, 2004c] Jordà, S. (2004c). Instruments and Players: Some Thoughts on Digital Lutherie.

[Jorda, 2005] Jorda, S. (2005). *Digital Lutherie*. PhD thesis.

[Jordà, 2005a] Jordà, S. (2005a). Digital Lutherie Crafting musical computers for new musics ' performance and improvisation. *Departament de Tecnologia*, 26(3):531.

[Jordà, 2005b] Jordà, S. (2005b). Multi-user instruments: models, examples and promises. In *Proceedings of the 2005 conference on New interfaces for musical expression*, pages 23–26, National University of Singapore.

[Jordà, 2014] Jordà, S. (2014). A Methodological Framework for Teaching , Evaluating and Informing NIME Design with a Focus on Expressiveness and Mapping. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 233–238.

[Jordà et al., 2007] Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. (2007). The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 139–146. ACM.

[Jordanous, 2011] Jordanous, A. (2011). Evaluating Evaluation: Assessing Progress in Computational Creativity. pages 102–107.

[Jules Françoise, Baptiste Caramiaux, 2012] Jules Françoise, Baptiste Caramiaux, F. B. (2012). A hierarchical approach for the design of gesture-to-sound mappings. *Smc'12*, pages 1 – 8.

[Juslin, 2003] Juslin, P. N. (2003). Five facets of musical expression: A psychologist's perspective on music performance. *Psychology of Music*, 31(3):273–302.

[Kahol and Rikakis, 2003] Kahol, K. and Rikakis, T. (2003). Gesture segmentation in complex motion sequences. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 2, pages II—-105. IEEE.

[Kaiser and Peeters, 2013] Kaiser, F. and Peeters, G. (2013). A Simple Fusion Method of State and Sequence Segmentation for Music Structure Discovery. *Proc. of the 14th International Society for Music Information Retrieval Conference*.

[Kaltenbranner et al., 2006] Kaltenbranner, M., Jorda, S., Geiger, G., and Alonso, M. (2006). The reactable*: A collaborative musical instrument. In *Enabling Technologies: Infrastructure for Collaborative Enterprises, 2006. WETICE'06. 15th IEEE International Workshops on*, pages 406–411. IEEE.

[Kaltenbrunner et al., 2004] Kaltenbrunner, M., Geiger, G., and Jordà, S. (2004). Dynamic patches for live musical performance. *Proc. of NIME'04*, pages 19–22.

[Keil, 1987] Keil, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, 2(3):275–283.

[Keil, 1995] Keil, C. (1995). The Theory of Participatory Discrepancies: A Progress Report. *Ethnomusicology*, 39(1):pp. 1–19.

[Keller, 2007] Keller, P. E. (2007). Musical ensemble synchronisation. In *International Conference on Music Communication Science*, number December, pages 80–83.

[Keller, 2013] Keller, P. E. (2013). Musical ensemble performance : A theoretical framework and empirical findings on interpersonal coordination.

[Keller and Mills, 2008] Keller, P. E. and Mills, W. C. (2008). Joint Action in Music Performance. *Culture*, pages 205–221.

[Keller and Repp, 2008] Keller, P. E. and Repp, B. H. (2008). Multilevel coordination stability: integrated goal representations in simultaneous intra-personal and inter-agent coordination. *Acta psychologica*, 128(2):378–86.

[Kiefer et al., 2008] Kiefer, C., Collins, N., and Fitzpatrick, G. (2008). HCI Methodology For Evaluating Musical Controllers: A Case Study.

[Kirke and Miranda, 2009] Kirke, A. and Miranda, E. R. (2009). A survey of computer systems for expressive music performance. *ACM Computing Surveys*, 42(1):1–41.

[Kitani and Koike, 2010] Kitani, K. M. and Koike, H. (2010). Improvgenerator: Online grammatical induction for on-the-fly improvisation accompaniment. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010)*.

[Klügel et al., 2014] Klügel, N., Hagerer, G., and Groh, G. (2014). TreeQuencer: Collaborative Rhythm Sequencing - A Comparative Study. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 50–53.

[Knotts, 2013] Knotts, S. (2013). A portfolio of compositions: commentary.

[Knotts and Collins, 2014] Knotts, S. and Collins, N. (2014). The Politics of Laptop Ensembles: A Survey of 160 Laptop Ensembles and their Organisational Structures. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 191–194.

[Koestler, 1964] Koestler, A. (1964). The act of creation.

[Lartillot and Cereghetti, 2013] Lartillot, O. and Cereghetti, D. (2013). a Simple , High-Yield Method for Assessing Structural Novelty. *Proceedings of the 3rd International Conference of Music and Emotion*, (June).

[Lee et al., 2012] Lee, S. W., Freeman, J., Colella, A., Yao, S., and Van Troyer, A. (2012). Evaluating Collaborative Laptop Improvisation with LOLC. In *Proceedings of the Symposium on Laptop Ensembles and Orchestras*, pages 55–62.

[Lehmann and Kopiez, 2010] Lehmann, A. C. and Kopiez, R. (2010). The difficulty of discerning between composed and improvised music. *Style (DeKalb, IL)*, pages 113–129.

[Leman, 2010] Leman, M. (2010). An embodied approach to music semantics. *Musicae Scientiae*, 14(1 suppl):43–67.

[Lerdahl and Jackendoff, 1985] Lerdahl, F. and Jackendoff, R. (1985). *A generative theory of tonal music*. MIT press.

[Livingstone, 2003] Livingstone, D. (2003). Emergent Behavior in the context of Reactive Compositional Environments.

[Locke, 1980] Locke, D. (1980). Improvisation in West African musics. *Music Educators Journal*, 66(5):125–133.

[MA, 2005] MA, W. J. W. B. A. (2005). *Improving Music Performance Assessment*. PhD thesis, Faculty of Education, Griffith University.

[MacDonald, 2008] MacDonald, M. (2008). *Schoenberg*. Oxford University Press.

[Magdon-Ismail and Purnell, 2010] Magdon-Ismail, M. and Purnell, J. T. (2010). Approximating the Covariance Matrix with Low-rank Perturbations. *Intelligent Data Engineering and Automated Learning–IDEAL 2010*, pages 300—-307.

[Magnusson, 2011a] Magnusson, T. (2011a). Coding Live Music. 21:19–23.

[Magnusson, 2011b] Magnusson, T. (2011b). ixi lang: a SuperCollider parasite for live coding. In *Proceedings of International Computer Music Conference*, pages 503–506. University of Huddersfield.

[Maher, 2010] Maher, M. (2010). Evaluating creativity in humans, computers, and collectively intelligent systems. ... *Network Conference on Creativity and Innovation in ...*, pages 22–28.

[Maher and Fisher, 2012] Maher, M. and Fisher, D. (2012). Using AI to evaluate creative designs. *2nd International Conference on Design Creativity*.

[Maher, 2012] Maher, M. L. (2012). Computational and Collective Creativity : Who ' s Being Creative ? Creativity : Process and Product. *International Conference on Computational Creativity*, (2003):67–71.

[Maher et al., 2013] Maher, M. L., Brady, K., and Fisher, D. H. (2013). Computational Models of Surprise in Evaluating Creative Design. *Proceedings of the 4th International Conference on Computational Creativity*, pages 147–151.

[Maidhof, 2013] Maidhof, C. (2013). Error monitoring in musicians. *Frontiers in human neuroscience*, 7.

[Malloch et al., 2014] Malloch, J., Sinclair, S., and Wanderley, M. M. (2014). Distributed tools for interactive design of heterogeneous signal networks. *Multimedia Tools and Applications*, pages 1–25.

[Marchini, 2014] Marchini, M. (2014). *Analysis of Ensemble Expressive Performance in String Quartets: a Statistical and Machine Learning Approach*. phd, Univesitat Pompeu Fabra.

[Marquez-Borbon et al., 2011] Marquez-Borbon, A., Gurevich, M., Fyans, A., and Stapleton, P. (2011). Designing digital musical interactions in experimental contexts. *Contexts*, (June):373–376.

[Martin et al., 2015] Martin, C., Gardner, H., and Swift, B. (2015). Tracking Ensemble Performance on Touch-Screens with Gesture Classification and Transition Matrices. pages 359–364.

[Mathews, 1963] Mathews, M. V. (1963). The digital computer as a musical instrument. *Science*, 142(3592):553–557.

[Matthusen, 2009] Matthusen, P. (2009). Paula Matthusen: Lathyrus. *Leonardo Music Journal*, 19:108–109.

[McCartney, 2002] McCartney, J. (2002). Rethinking the computer music language: SuperCollider. *Computer Music Journal*, 26(4):61–68.

[Menezes, 2010] Menezes, J. (2010). *Creative Process in Free Improvisation*. PhD thesis, University of Sheffield.

[Minsky, 1980] Minsky, M. (1980). K-Lines: A theory of Memory. *Cognitive science*, 4(2):117—-133.

[Miranda and Wanderley, 2006] Miranda, E. R. and Wanderley, M. M. (2006). *New digital musical instruments: control and interaction beyond the keyboard*, volume 21. AR Editions, Inc.

[Mitra and Acharya, 2007] Mitra, S. and Acharya, T. (2007). Gesture recognition: A survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(3):311–324.

[Morreale et al., 2014] Morreale, F., Angeli, A. D., and O'Modhrain, S. (2014). Musical Interface Design: An Experience-oriented Framework. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 467–472.

[Murray-Browne and Mainstone, 2011] Murray-Browne, T. and Mainstone, D. (2011). The medium is the message: Composing instruments and performing mappings. *Proceedings of the* ....

[Narveson and Trueman, 2013] Narveson, J. and Trueman, D. (2013). LANdini: a networking utility for wireless LAN-based laptop ensembles. In *Proc. SMC Sound, Music and Computing Conference*.

[Neal, 2008] Neal, A. S. (2008). *The Continuum of Indeterminacy in Live Computer Music*. PhD thesis, Queen's University Belfast.

[Neeman, 2014] Neeman, E. (2014). *Free Improvisation as a Performance Technique: Group Creativity and Interpreting Graphic Scores*. PhD thesis, The Juilliard School.

[Noy et al., 2012] Noy, L., Hart, Y., Andrew, N., Ramote, O., Mayo, A., and Alon, U. (2012). A Quantitative Study of Creative Leaps. pages 72–76.

[Nymoen et al., 2012] Nymoen, K., Torresen, J., Godøy, R. I., and Jensenius, A. R. (2012). A Statistical Approach to Analyzing Sound Tracings. In *Speech, Sound and Music Processing: Embracing Research in India*, pages 120—-145. Springer.

[Ogborn, 2014] Ogborn, D. (2014). Live Coding in a Scalable, Participatory Laptop Orchestra. *Computer Music Journal*, 38(1):17–30.

[Oh et al., 2010] Oh, J., Herrera, J., Bryan, N. J., Dahl, L., and Wang, G. (2010). Evolving The Mobile Phone Orchestra. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 82–87.

[Oman et al., 2013] Oman, S. K., Tumer, I. Y., Wood, K., and Seepersad, C. (2013). A comparison of creativity and innovation metrics and sample validation through in-class design projects. *Research in Engineering Design*, 24(1):65–92.

[O'Modhrain, 2011] O'Modhrain, S. (2011). A framework for the evaluation of digital musical instruments. *Computer Music Journal*, 35(1):28–42.

[Paine, 2004] Paine, G. (2004). Gesture and Musical Interaction : Interactive Engagement Through Dynamic Morphology. *Gesture*, pages 80–86.

[Paine, 2009] Paine, G. (2009). Gesture and Morphology in Laptop Music Performance Gesture The Computer as Instrument. *The Oxford Handbook of Computer Music*, pages 214–232.

[Palmer, 1996a] Palmer, C. (1996a). Anatomy of a performance: Sources of musical expression. *Music Perception*, pages 433–453.

[Palmer, 1996b] Palmer, C. (1996b). On the assignment of structure in music performance. *Music Perception*, pages 23–56.

[Palmer and Deutsch, 2013] Palmer, C. and Deutsch, D. (2013). Music performance: Movement and coordination. *The psychology of music*, pages 405–422.

[Palmer and Drake, 1997] Palmer, C. and Drake, C. (1997). Monitoring and planning capacities in the acquisition of music performance skills. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie exp{é}rimentale*, 51(4):369.

[Papiotis et al., 2011] Papiotis, P., Maestre, M., Marchini, M., and Perez, A. (2011). Synchronization of intonation adjustments in violin duets; towards an objective evaluation of musical interaction. In *Proceedings of the 14th conference in Digital Audio Effects (DAFx-11)*.

[Papiotis et al., 2012] Papiotis, P., Marchini, M., Maestre, E., and Perez, A. (2012). Measuring ensemble synchrony through violin performance parameters: a preliminary progress report. In *Intelligent Technologies for Interactive Entertainment*, pages 267–272. Springer.

[Pecenka and Keller, 2009] Pecenka, N. and Keller, P. E. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences*, 1169(1):282–286.

[Pecenka and Keller, 2011] Pecenka, N. and Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 211(3-4):505–15.

[Peuquet et al., 2012] Peuquet, S., Neal, A. S., Brien, B. O., and Garrison, T. (2012). THE FLORIDA LEAGUE FOR INDETERMINATE PERFORMANCE : IDEALISM AND FAILURE IN IMPROVISATORY LAPTOP ENSEMBLE PERFORMANCE. In *Proceedings of the 1st Symposium on Laptop Ensembles and Orchestras*, pages 41–46.

[Pi and Larsen, 2015] Pi, N. and Larsen, A. (2015). A study of the effect of the concept of signs on musical creativity. pages 1–60.

[Piana et al., 2013] Piana, S., Staglianò, A., Camurri, A., and Odone, F. (2013). A set of full-body movement features for emotion recognition to help children affected by autism spectrum condition. In *IDGEI International Workshop*.

[Pimentel et al., 2014] Pimentel, M. a. F., Clifton, D. a., Clifton, L., and Tarassenko, L. (2014). A review of novelty detection. *Signal Processing*, 99:215–249.

[Poepel, 2005] Poepel, C. (2005). On interface expressivity: A player based study. *Nime*, pages 228–231.

[Pointeau, 2013] Pointeau, A. (2013). *Creativity Studies on the Reactable online community. Assessing the validity of distributed creativity models in an online community of creators*. Masther thesis, Universitat Pompeu Fabra.

[Prame, 1997] Prame, E. (1997). Vibrato extent and intonation in professional Western lyric singing. *The Journal of the Acoustical Society of America*, (102):616–621.

[Pressing, 1987] Pressing, J. (1987). Improvisations: methods and models. *Generative processes in music*, pages 129–178.

[Prögler, 1995] Prögler, J. A. (1995). Searching for swing: Participatory discrepancies in the jazz rhythm section. *Ethnomusicology*, pages 21–54.

[Puckette, 1988] Puckette, M. (1988). The Patcher. In *International Computer Music Conference*, pages 420–429.

[Puckette and Others, 1996] Puckette, M. and Others (1996). Pure Data: another integrated computer music environment. *Proceedings of the Second Intercollege Computer Music Concerts*, pages 37–41.

[Puckette et al., 1990] Puckette, M., Zicarelli, D., and Others (1990). Max/Msp. *Cycling*, 74:1990–2006.

[Quick and Hudak, 2013] Quick, D. and Hudak, P. (2013). Grammar-based automated music composition in Haskell. In *Proceedings of the first ACM SIGPLAN workshop on Functional art, music, modeling & design*, pages 59–70. ACM.

[Ramalho, 1999] Ramalho, G. (1999). An artificially intelligent jazz performer. *Journal of New Music ...*, 28(1):1–24.

[Ramalho and Ganascia, 1994] Ramalho, G. L. and Ganascia, J.-G. (1994). Simulating creativity in jazz performance. In *Proceedings of the National Conference on Artificial Intelligence*, pages 108–108.

[Ramsay and Paradiso, 2015] Ramsay, D. B. and Paradiso, J. a. (2015). GroupLoop : A Collaborative , Network-Enabled Audio Feedback Instrument. In *New Interfaces for Music Expression*.

[Rebelo, 2010] Rebelo, P. (2010). Notating the Unpredictable. *Contemporary Music Review*, 29(1):17–27.

[Reeves et al., 2005] Reeves, S., Benford, S., O'Malley, C., and Fraser, M. (2005). Designing the spectator experience. pages 741–750.

[Reinecke, 2010] Reinecke, D. (2010). How To Conduct A Laptop Orchestra: Issues Of Coordination In Computer Music. In *Proceedings of the 2010 International Computer Music Conference*, page 400.

[Reinholdsson, 1998] Reinholdsson, P. (1998). Making music together: An interactionist perspective on small-group performance in jazz.

[Repp, 1992] Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's "Tr{ä}umerei". *The Journal of the Acoustical Society of America*, 92(5):2546–2568.

[Repp, 1995] Repp, B. H. (1995). Expressive timing in Schumann's "Tr{ä}umerei:"An analysis of performances by graduate student pianists. *The Journal of the Acoustical Society of America*, 98(5):2413–2427.

[Repp, 1996a] Repp, B. H. (1996a). Pedal timing and tempo in expressive piano performance: A preliminary investigation. *Psychology of Music*, 24(2):199–221.

[Repp, 1996b] Repp, B. H. (1996b). The art of inaccuracy: Why pianists' errors are difficult to hear. *Music Perception*, pages 161–183.

[Repp, 1997a] Repp, B. H. (1997a). Expressive timing in a Debussy Prelude: A comparison of student and expert pianists. *Musicae Scientiae*, 1(2):257–268.

[Repp, 1997b] Repp, B. H. (1997b). The aesthetic quality of a quantitatively average music performance: Two preliminary experiments. *Music Perception*, pages 419–444.

[Repp, 1999a] Repp, B. H. (1999a). A microcosm of musical expression: II. Quantitative analysis of pianists' dynamics in the initial measures of Chopin's Etude in E major. *The Journal of the Acoustical Society of America*, 105(3):1972–1988.

[Repp, 1999b] Repp, B. H. (1999b). Individual differences in the expressive shaping of a musical phrase: The opening of Chopin's etude in E major. *Music, mind, and science Seoul, Korea: SeoulNational University Press1999p*, pages 239–270.

[Repp, 2005] Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic bulletin & review*, 12(6):969–992.

[Rhodes, 1961] Rhodes, M. (1961). An analysis of creativity. *Phi Delta Kappan*, pages 305–310.

[Rink, 2002] Rink, J. (2002). *Musical performance: a guide to understanding*. Cambridge University Press.

[Ritchie, 2007] Ritchie, G. (2007). Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, 17(1):67–99.

[Ritsch, 2014] Ritsch, W. (2014). ICE - towards distributed networked computermusic ensemble. (September):14–20.

[Ritter et al., 2012] Ritter, S. M., van Baaren, R. B., and Dijksterhuis, A. (2012). Creativity: The role of unconscious processes in idea generation and idea selection. *Thinking Skills and Creativity*, 7(1):21–27.

[Roads, 1996] Roads, C. (1996). *The computer music tutorial*. MIT press.

[Rogers, 1991] Rogers, G. L. (1991). Effect of color-coded notation on music achievement of elementary instrumental students. *Journal of Research in Music Education*, 39(1):64–73.

[Rohrhuber et al., 2007] Rohrhuber, J., de Campo, A., Wieser, R., van Kampen, J.-K., Ho, E., and Hölzl, H. (2007). Purloined letters and distributed persons. In *Music in the Global Village Conference (Budapest)*.

[Roma, 2015] Roma, G. (2015). *Algorithms and representations for supporting online music creation with large-scale audio databases*. PhD thesis, Universitat Pompeu Fabra, Barcelona.

[Rousseeuw and Van Zomeren, 1990] Rousseeuw, P. J. and Van Zomeren, B. C. (1990). Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association*, 85(411):633–639.

[Rovan et al., 1997] Rovan, J. B., Wanderley, M. M., Dubnov, S., and Depalle, P. (1997). Instrumental gestural mapping strategies as expressivity determinants in computer music performance. In *Kansei, The Technology of Emotion. Proceedings of the AIMI International Workshop*, pages 68–73. Citeseer.

[Rowe, 1992] Rowe, R. (1992). *Interactive music systems: machine listening and composing*. MIT press.

[Runco, 2008] Runco, M. A. (2008). Divergent thinking is not synonymous with creativity: Comments on "Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods.". *Psychology of Aesthetics, Creativity, and the Arts*, 2(2):0.

[Russell, 2015] Russell, B. E. (2015). An empirical study of a solo performance assessment model. *International Journal of Music Education*, page 0255761415581282.

[Ruviaro, 2011] Ruviaro, B. (2011). From Schaeffer to *LOrks: an expanded definition of musical instrument in the context of laptop orchestras. In *26th Annual Conference of the Society for Electro-Acoustic Music in the United States*, volume 31.

[Sarkar and Chakrabarti, 2011] Sarkar, P. and Chakrabarti, A. (2011). Assessing design creativity. *Design Studies*, pages 1–36.

[Sarlo, 2003] Sarlo, J. a. (2003). GrIPD: A Graphical Interface Editing Tool and Run-Time Environment for Pure Data. *Proc. of the International Computer Music Conference*, pages 305–307.

[Saunders and Gero, 2001] Saunders, R. and Gero, J. S. (2001). Designing for Interest and Novelty: Motivating Design Agents. In *Proceedings of CAAD Futures 2001*, number 1990, pages 725–738.

[Saunders and Holahan, 1997] Saunders, T. C. and Holahan, J. M. (1997). Criteria-specific rating scales in the evaluation of high school instrumental performance. *Journal of Research in Music Education*, 45(2):259–272.

[Sawyer, 2011] Sawyer, R. K. (2011). *Explaining creativity: The science of human innovation*. Oxford University Press.

[Sawyer and DeZutter, 2009] Sawyer, R. K. and DeZutter, S. (2009). Distributed creativity: How collective creations emerge from collaboration. *Psychology of Aesthetics, Creativity, and the Arts*, 3(2):81–92.

[Sawyer and Others, 2014] Sawyer, R. K. and Others (2014). *Group creativity: Music, theater, collaboration*. Psychology Press.

[Schmeder et al., 2010] Schmeder, A., Freed, A., and Wessel, D. (2010). Best practices for open sound control. In *Linux Audio Conference*.

[Schoenberg and Stein, 1975] Schoenberg, A. and Stein, L. (1975). *Style and Idea: selected writings of Arnold Schoenberg*. Univ of California Press.

[Scholar and Zanter, 2013] Scholar, M. D. and Zanter, M. (2013). Lawrence D . " Butch " Morris , Conduction and the Network of Composition.

[Schwarz, 2003] Schwarz, D. (2003). The caterpillar system for data-driven concatenative sound synthesis. In *Proceedings of the COST-G6 Conference on Digital Audio Effects (DAFx)*, pages 135–140.

[Schwarz and Others, 2004] Schwarz, D. and Others (2004). Data-driven concatenative sound synthesis.

[Scott, 2014] Scott, R. (2014). The Molecular Imagination: John Stevens, The Spontaneous Music Ensemble and Free Group Improvisation. In HAodha, M. Ó. and Schroeder, F., editors, *Soundweaving: Writings on Improvisation*, pages 95–110. Cambridge Scholard Publishing.

[Seashore, 1938] Seashore, C. E. (1938). *Psychology of music*. Courier Corporation.

[Settel and Lippe, 2003] Settel, Z. and Lippe, C. (2003). Convolution Brother ' s Instrument Design. In *Proceedings of the 2003 International Conference on New Interfaces for Musical Expression (NIME-03)*, number August, pages 197–200.

[Shah et al., 2000] Shah, J. J., Kulkarni, S. V., and Vargas-Hernandez, N. (2000). Evaluation of Idea Generation Methods for Conceptual Design: Effectiveness Metrics and Design of Experiments.

[Simpson, 1982] Simpson, R. M. (1982). Social psychology of creativity: A consensual assessment technique. *Pers. & Soc. Psychol*, 43:997–1013.

[Sloboda, 1992] Sloboda, J. A. (1992). Empirical studies of emotional response to music.

[Smalley, 1997] Smalley, D. (1997). Spectromorphology: explaining sound-shapes. *Organised sound*, 2(02):107–126.

[Smallwood et al., 2008] Smallwood, S., Trueman, D., Cook, P. R., and Wang, G. (2008). Composing for Laptop Orchestra. *Computer Music Journal*, 32(1):9–25.

[Smith and Garnett, 2012] Smith, B. D. and Garnett, G. E. (2012). Unsupervised play: Machine learning toolkit for Max. In *Proceedings of NIME*.

[Sosa and Gero, 2003] Sosa, R. and Gero, J. S. (2003). Design and change: a model of situated creativity. page 25. IJCAI 2003 Workshop 3rd. Workshop on Creative Systems.

[Stockhausen and Maconie, 1989] Stockhausen, K. and Maconie, R. (1989). Stockhausen on Music Lectures and Interviews.

[Stowell et al., 2008] Stowell, D., Plumbley, M. D., and Bryan-Kinns, N. (2008). Discourse analysis evaluation method for expressive musical interfaces. In *Proc. of NIME'08*.

[Stowell et al., 2009] Stowell, D., Robertson, a., Bryan-Kinns, N., and Plumbley, M. (2009). Evaluation of live human–computer music-making: Quantitative and qualitative approaches. *International Journal of Human-Computer Studies*, 67(11):960–975.

[Stravinsky, 1970] Stravinsky, I. (1970). *Poetics of music in the form of six lessons*, volume 66. Harvard University Press.

[Sundberg et al., 1983] Sundberg, J., Askenfelt, A., and Frydén, L. (1983). Musical performance: A synthesis-by-rule approach. *Computer Music Journal*, pages 37–43.

[Susilo, 2001] Susilo, H. (2001). Toward an Appreciation of Javanese Gamelan.

[Sutton, 2001] Sutton, J. (2001). *'The Invisible Handshake': An Investigation of Free Musical Improvisation as a Form of Conversation*.

[Sutton, 1998] Sutton, R. A. (1998). Do Javanese gamelan musicians really improvise? *In the course of performance: Studies in the world of musical improvisation*, pages 69–92.

[Swain, 1997] Swain, J. P. (1997). *Musical languages*. WW Norton & Company.

[Swift, 2012] Swift, B. (2012). *The Design of a Smartphone-based Digital Musical Instrument for Jamming*. PhD thesis, Australian National University.

[Swift, 2013] Swift, B. (2013). Chasing a feeling: Experience in computer supported jamming. In *Music and human-computer interaction*, pages 85–89. Springer London.

[Tanaka et al., 2011] Tanaka, A., Altavilla, A., and Spowage, N. (2011). Gestural Musical Affordances.

[Taylor, 1988] Taylor, C. W. (1988). Various approaches to and definitions of creativity. *The nature of creativity*, pages 99–121.

[Tech et al., 2010] Tech, V., Martin, T., Standley, E., and Matthews, M. (2010). Introducing L 2 Ork : Linux Laptop Orchestra. *Interfaces*, (Nime):170–173.

[Thompson and Williamon, 2003] Thompson, S. and Williamon, A. (2003). Evaluating evaluation: Musical performance assessment as a research tool.

[Thompson et al., 1998] Thompson, W. F., Diamond, C. T. P., and Balkwill, L.-L. (1998). The adjudication of six performances of a Chopin Etude: A study of expert knowledge. *Psychology of Music*, 26(2):154–174.

[Todd, 1985] Todd, N. (1985). A model of expressive timing in tonal music. *Music perception*, pages 33–57.

[Torres et al., 2008] Torres, G., Basnet, R., and Sung, A. (2008). A similarity measure for clustering and its applications. *Proceedings of World Academy of Science, Engineering and Technology*, 31(1307-6884):490–496.

[Troyer et al., 2012] Troyer, A. V., Textual, C., Link, C., and Freeman, J. (2012). Collaborative Textual Improvisation in a Laptop Ensemble.

[Trueman, 2007] Trueman, D. (2007). Why a laptop orchestra? *Organised Sound*, 12(2):171–179.

[Trueman, 2010] Trueman, D. (2010). Digital Instrument Building and the Laptop Orchestra. In *US Frontiers of Engineering Symposium. Armonk.*

[Trueman and Cook, 2000a] Trueman, D. and Cook, P. (2000a). BoSSA: The deconstructed violin reconstructed. *Journal of New Music Research*, 29(2):121–130.

[Trueman and Cook, 2000b] Trueman, D. and Cook, P. (2000b). BoSSA: The Deconstructed Violin Reconstructed. *Journal of New Music Research*, 29(2):121–130.

[Trueman et al., 2006] Trueman, D., Cook, P. R., and Smallwood, S. (2006). PLOrk: Princeton laptop orchestra, year 1. *Proceedings of the ...*.

[Turner and Fauconnier, 2002] Turner, M. and Fauconnier, G. (2002). The way we think: conceptual blending and the mind's hidden complexities. *EUA: Basic Books*.

[Vallis et al., 2012] Vallis, O., Diakopoulos, D., Hochenbaum, J., and Kapur, A. (2012). Building on the Foundations of Network Music : Exploring interaction contexts and shared robotic instruments *. *Organised Sound*, 17(1):62–72.

[Van Nort and Wanderley, 2006] Van Nort, D. and Wanderley, M. (2006). The LoM Mapping Toolbox for Max/MSP/Jitter. In *Proc. of the 2006 International Computer Music Conference (ICMC)*, pages 397–400.

[Wakefield et al., 2010] Wakefield, G., Smith, W., and Roberts, C. (2010). LuaAV: Extensibility and Heterogeneity for Audiovisual Computing. *Proceedings of the Linux Audio Conference*.

[Wallas, 1926] Wallas, G. (1926). The art of thought.

[Wallbott, 1998] Wallbott, H. G. (1998). Bodily expression of emotion. *European journal of social psychology*, 28(6):879–896.

[Wanderley and Orio, 2002] Wanderley, M. M. and Orio, N. (2002). Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI.

[Wanderley et al., 1998] Wanderley, M. M., Schnell, N., and Rovan, J. (1998). Escher-modeling and performing composed instruments in real-time. In *Systems, Man, and Cybernetics, 1998. 1998 IEEE International Conference on*, volume 2, pages 1080–1084. IEEE.

[Wang, 2014] Wang, G. (2014). Principles of Visual Design for Computer Music. (September):391–396.

[Wang et al., 2009] Wang, G., Bryan, N., Oh, J., and Hamilton, R. (2009). Stanford laptop orchestra (slork). In *Proceedings of the International Computer Music Conference*, number Icmc, page 508.

[Wang and Cook, 2004] Wang, G. and Cook, P. (2004). ChucK : A Programming Language for On-the-fly , Real-time Audio Synthesis and Multimedia. *12th annual ACM international conference on Multimedia*, pages 812–815.

[Wang and Cook, 2006] Wang, G. and Cook, P. R. (2006). Building Collaborative Graphical interFaces in the Audicle. *Proc. NIME 2006*, pages 49–52.

[Wang et al., 2008] Wang, G., Essl, G., and Penttinen, H. (2008). Do mobile phones dream of electric orchestras. *Proceedings of the International ...*, 16(10):1252–61.

[Wang et al., 2005] Wang, G., Misra, A., Davidson, P., and Cook, P. R. (2005). CoAudicle: A Collaborative Audio Programming Space. *International Computer Music Conference*.

[Weaver, 1948] Weaver, W. (1948). Probability, rarity, interest, and surprise. *The Scientific Monthly*, 67(6):390.

[Weinberg, 2002] Weinberg, G. (2002). Playpens, Fireflies and Squeezables: New Musical Instruments for Bridging the Thoughtful and the Joyful. *Leonardo Music Journal*, (12):43–51.

[Weinberg, 2003] Weinberg, G. (2003). Interconnected Musical Networks - Bringing Expression and Thoughtfulness to Collaborative Group Playing. *Sciences-New York*, (1994).

[Weinberg, 2005] Weinberg, G. (2005). Interconnected Musical Networks : Toward a Theoretical Framework Interconnected Musical Networks : Toward a Theoretical Framework. *Computer*, 29(2):23–39.

[Weinberg and Gan, 2000] Weinberg, G. and Gan, S.-l. (2000). The Squeezables : Toward an Expressive and Interdependent Multi-player Musical Instrument. pages 37–45.

[Weinberg et al., 2008] Weinberg, G., Godfrey, M., Rae, A., and Rhoads, J. (2008). A real-time genetic algorithm in human-robot musical improvisation. In *Computer music modeling and retrieval. Sense of sounds*, pages 351–359. Springer.

[Weisberg et al., 2004] Weisberg, R. W., Brinkman, a. R., Folio, C. J., Dick, a. S., Fleck, J. I., Niederberger, B., and Bassett, F. (2004). Toward a cognitive analysis of creativity: Improvisation in jazz. *Proceedings of the Conference on Interdisciplinary Musicology Annekatrin Kessler and Frank Zimmer Graz Austria Conference on Interdisciplinary Musicology*, pages 1–9.

[Weitzner et al., 2012] Weitzner, N., Freeman, J., Garrett, S., and Chen, Y.-l. (2012). massMobile – an Audience Participation Framework. *NIME 2012 Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 92–95.

[Wessel, 1979] Wessel, D. L. (1979). Timbre space as a musical control structure. *Computer music journal*, pages 45–52.

[Widmer and Goebl, 2004] Widmer, G. and Goebl, W. (2004). Computational models of expressive music performance: The state of the art. *Journal of New Music Research*, 33(3):203–216.

[Wiggins, 2006] Wiggins, G. a. (2006). A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, 19(7):449–458.

[Williamon et al., 2006] Williamon, A., Thompson, S., Lisboa, T., and Wiffen, C. (2006). Creativity, originality and value in music performance. *Musical creativity: Multidisciplinary research in theory and practice*, page 26.

[Wilson et al., 2010] Wilson, S., Lorway, N., Vasilakos, K., and Moyers, T. (2010). Free as in BEER : Some Explorations into Structured Improvisation Using Networked Live-Coding Systems. (Fei 2001):54–64.

[Wing et al., 2014] Wing, A. M., Endo, S., Bradbury, A., and Vorberg, D. (2014). Optimal feedback correction in string quartet synchronization. *Journal of The Royal Society Interface*, 11(93):20131125.

[Winkler and Czigler, 2012] Winkler, I. and Czigler, I. (2012). Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology*, 83(2):132–143.

[Winkler et al., 2012] Winkler, I., Denham, S., Mill, R., B\Hohm, T. M., and Bendixen, A. (2012). Multistability in auditory stream segregation: a predictive coding view. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1591):1001–1012.

[Winkler et al., 2009] Winkler, I., Denham, S. L., and Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in cognitive sciences*, 13(12):532–540.

[Winkler, 1998] Winkler, T. (1998). *Composing interactive music: techniques and ideas using Max*.

[Wolterstorff, 1980] Wolterstorff, N. (1980). Works and worlds of art.

[Wright, 1997] Wright, M. (1997). OpenSound Control: A New Protocol for Communicationg with Sound Synthesizers. *ICMC 97*.

[Xambó, 2015] Xambó, A. (2015). *Tabletop Tangible Interfaces for Music Performance: Design and Evaluation*. PhD thesis, The Open University.

[Xambó et al., 2013] Xambó, A., Hornecker, E., Marshall, P., Jorda, S., Dobbyn, C., and Laney, R. (2013). Let's jam the reactable: Peer learning during musical improvisation with a tabletop tangible interface. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(6):36.

[Xambó et al., 2011] Xambó, A., Laney, R., Dobbyn, C., and Jordà, S. (2011). Multi-touch interaction principles for collaborative real-time music activities: towards a pattern language. In *Proc. of ICMC'11*, pages 403–406.

[Zils and Pachet, 2001] Zils, A. and Pachet, F. (2001). Musical mosaicing. In *Digital Audio Effects (DAFx)*, volume 2.

[Zmölnig and Eckel, 2007] Zmölnig, I. and Eckel, G. (2007). Live coding: An overview. In *Proceedings of the International Computer Music Conference*.