

© 2011 by Benjamin D. Smith. All rights reserved.

TELEMATIC COMPOSITION

BY

BENJAMIN D. SMITH

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Musical Arts in Music
with a concentration in Music Composition
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Associate Professor Guy E. Garnett, Chair, Director of Research
Assistant Professor Kyratso George Karahalios
Assistant Professor Heinrich Taube
Professor Thomas Turino

Abstract

The practice of distributed, networked music performance, involving video and audio representations, is rapidly growing and presenting new possibilities and opportunities for aesthetic expression. This thesis examines the expanding body of works in telematic music and formulates a critical analysis of the primary aesthetic and technical concerns for music composition and performance. Presence, agency, and interactivity inform telematic music performance, stemming from the transparency or mediation of the network, directly impacting the expressive capabilities of the medium. Practical issues in the problem domain of networked ensemble performance were examined through a series of experimental research sessions, showing that latency affects the perception of interactivity and presence and strongly coerces musical results. Visual representations can be critical for communication at key ensemble moments, regardless of synchrony with other media, yet audio streams remain the primary channel for musical connection and interaction. The multiple composite musical results, different at each location due to network transmission times, leads to a unique reformulation of the notion of “piece,” creating many valid aesthetic experiences from a single, distributed production.

To Meagan and Rowan.

Acknowledgments

I would like to thank my advisor, Guy E. Garnett, for his enduring support; my thesis committee, Karrie Karahalios, Rick Taube, and Tom Turino, for encouraging and enabling this work; my family, Meagan and Rowan, and my parents, Barbara and Steve, for setting me on this path and inspiring my perseverance. Financial support for this research came from the Illinois Informatics Institute, eDream and the Cultural Computing Lab, the Center for Technology and Public Policy, and the Illinois Japan Performing Arts Network.

Table of Contents

List of Abbreviations	vi
Chapter 1 Introduction	1
Chapter 2 Telematics	4
2.1 Composition	4
2.2 Performance	5
2.3 Aesthetic Concerns	7
2.4 Latency	16
2.5 Musical Dichotomies	21
Chapter 3 Experimental Research	27
3.1 Methodology	27
3.2 Technology	27
3.3 Findings	30
3.4 Style	36
3.5 Conclusion	37
Appendix A Survey and Analysis	39
A.1 Earliest Examples	39
A.2 Pieces	40
Appendix B Research Experiment Report	50
B.1 Methodology	50
B.2 Research Log	51
Appendix C Appendix 3	70
C.1 Telematic Etudes	70
References	72

List of Abbreviations

CD	Compact Disc
db	decibels
HD	High Definition
ms	milliseconds
sec	second/seconds

Chapter 1

Introduction

How might new technologies and order of magnitude improvements in existing technologies lead to new art, new art forms, new modes of artistic expression or new forms of artistic engagement? Artists have historically embraced the advent of new technologies, often pioneering their development: oil paint, photography, film, and the computer all had significant impact on existing art forms and led the field, and the technology in new directions. How might the technologies of high speed networks and multi-core, high performance desktop computing impact and transform the arts, and music, today?

This thesis approaches these questions through the examination of telematic music and networked music performance. “Telematic” connotes a focus on telecommunications and informatics (Nora and Minc, 1980), and in music typically refers to ensemble performance where the performers are distributed between disparate physical locations that are connected via the internet (Dresser, 2008). While the first examples of telematic art were seen in the 1920s (Ascott and Shanken, 2007), the technical expertise and hardware requirements have resulted in very little exploration until the last decade.

The first musical collaborations over network connections were seen in the 1980s (Cooperstock, 2011), employing satellite connections to bring artists across the United States into communication. Around this time the League of Automatic Music Composers and its offshoot The Hub formed with the express intent of exploring the potentials of network-based music and art (Gresham-Lancaster, 1998). Since the turn of the twenty-first century, with widespread institutional access to high-speed networks, artists have begun exploring telematics with much more frequency. This has resulted in a plethora of examples, bridging all the visual and performance arts as they intersect or extend into distributed, networked spaces. Musical works set in the networked domain come from two different approaches: one focusing on the computers and network topography, seeking to employ them as instruments for artistic creation (Barbosa, 2003; Gresham-Lancaster, 1998), and the other focusing on the communicative aspect of networks and their ability to bring people together across large physical (and temporal) distances. Of course, as with all oppositions many cases encompass both aspects, such as the work of Weinberg (2002) and The Hub. However, the application of computers *as* instruments and algorithmic music generators as opposed to enabling communication is the

key distinction.

I take this latter set as my focus, approaching digital networks as a way to expand acoustic ensemble practice, and confronting a further philosophical divide at many turns: between artists modeling co-present musical practices (i.e. the video-conferencing model) and artists looking for new musical practices in reaction to the unique potentials of the networked medium. As we will see, this contrast of approach is found at many levels, from the approach to the technology, to the music composed and chosen, to the staging of the performances. While both are certainly valid endeavors, and the former promises to facilitate the development of higher grade video conferencing technology, my overall argument is for the latter, pursuing advances for the field of composition and its practitioners.

Concerns exposed through the examination of the extent telematic music literature and documented performances are covered first. These start with the premises of musical performance as practiced in concert halls today, and considers the primary problems these practices face when transposed to distributed settings. Access to the presence and agency (i.e. knowledge of the actions and musical effects) of remote performers becomes a focal concern, which is exacerbated by the mediated nature of the environment (the projections and sound systems) and the time delay introduced by the network transmission (*latency*). However, the issues of presence, agency, and latency also inform the search for the unique potentials of the medium and the ability to create new music and new musical expressions.

To date, the bulk of studies of music in distributed settings have implicitly held face-to-face interactions as their measure of truth. However, their primary contribution has been to show the difficulties of co-present music making over a network connection, and to expose the unspoken assumptions about the ideals of music making as practiced in concert hall performance settings today.

In order to uncover the potentials and problems of distributed music making, and begin mapping the domain, I set up and conducted a series of constrained musical experiments. Taking the model of *Research through Art*, as described by Dallow (2003) and Schiphorst (2009), these experiments privileged the creative intuitions of the musicians within a frame of intellectual and methodological rigor that allowed the adoption, creation, and testing of new technologies towards discovering new musical expressions. These experiments informed the composition of a series of small ensemble pieces (for three to seven musicians), which were performed in conjunction with the lecture presentation of this thesis.

The findings of this research are interwoven in the following text, informing and responding to the philosophical and analytical discussion. Details of setup and technology choices confirm the approach to presence and agency and layout their use towards aesthetic ends. The examination of latency produced a spectrum of modalities (based on the length of the time delay), each of which strongly coerces and impacts

the music being played. These modalities present a set of affordances, ultimately bringing the focus to the creation of a unique piece of music at each physical location, as all of the musical parts are combined differently at each point based on the latencies and network topography.

The appendices comprise detailed analysis and critiques of historical telematic music performances (Appendix A), the detailed reports from my research experiments (Appendix B), and a series of etudes written to facilitate the exploration of ensemble playing into distributed environments (Appendix C).

Chapter 2

Telematics

2.1 Composition

Across the spectrum of contemporary compositional domains, from acoustic concert works to electro-acoustic fixed media recordings to interactive computer-facilitated performance, composers are primarily concerned (as creative artists) with the structuring and ordering of sonic events and materials. Telematic situations present numerous challenges to this pursuit and may fundamentally change and redefine long-held assumptions about compositional and musical practice. Issues of perceptibility, musical agency, identity, synchronization, alignment, and the location of the “work” are raised for examination.

While telematic musical performance embraces all forms of musical expression, a solid understanding must be built from the foundations of human musical practice, that is, from acoustic sound. Thus this thesis takes acoustic music performance as its locus of origin, allowing that additional elements of electronic instruments, artificially intelligent performing agents, and pre-recorded material will be reconciled when their concerns and challenges are understood and addressed.

Even at a basic conceptual level, there is an emerging divide in the approach to telematic music as evidenced by the goals and works of its practitioners, between the intentional replication of co-present, face-to-face musical practices and the contrasting pursuit of new aesthetic expressions and musical possibilities. This difference of objectives is seen again and again, especially in regards to the conceptualization of latency and the resulting musical choices.

This chapter first presents a look at telematic philosophy as it relates to the field of music, and begins to formulate methods for the evaluation of telematic musical works. Discussion of the potentials and problems revealed by practice in this burgeoning field follows, including some examination of exemplary cases. Consideration is then given to existing research in distributed performance, and evidence from human perception and cognition is presented to guide the discourse. Finally, the identity of the musical work, its location, and the ultimate impact of the network time delay are presented towards the reformulation of the nature of composition and the ‘piece.’

2.2 Performance

The origins of telematics are rooted in human interaction through telecommunications, and its potential for societal, or even global, transformation (Ascott and Shanken, 2007). Emphasis is placed on engagement over physical distance, or with the virtually real; and on human agency, even amid a vastly growing web of data and artificial systems. Telematic art requires at a minimum a duality, communication between two individuals, either concurrently or distanced in some measure over time; both as instigators, participants, and viewers, taking an active role in the consumption and continual creation of the art.

The term “*télématique*” was first coined by the French *Inspecteur des Finances*, Simon Nora and his assistant Alain Minc, in their 1978 treatise “The Computerization of Society” (Nora and Minc, 1980). The word was born out of a need to frame the “increasing interconnection[s] between computers and telecommunications.” The emphasis in their treatise is on telecommunications, as the French government anticipated that the increasing proliferation of digital networks would have a wide reaching and dramatically transformational impact on society.

Roy Ascott, a vocal proponent of telematic arts, lauded practitioner (Ascott and Shanken, 2007), and prolific writer on the topic, extends the definition and application of telematics, quoting Nora and Minc (here in 1984): “telematics is the ‘springing to life born of the marriage between computers and communications networks’ ” (Ascott and Shanken, 2007). Six years later, in 1990, Ascott set forth his clearest definition of telematics as: “computer-mediated communications networks between geographically dispersed individuals and institutions . . . and between the human mind and artificial systems of intelligence and perception.”

Throughout his work Ascott is focally concerned with the connection of human beings, with communication, and the creation of a transformational network of systems resulting in a new space for “planetary conviviality and creativity” (Ascott and Shanken, 2007). Intentionally discarding terms such as “information technology” and “computer-mediated interaction,” he preferred “telematics” as the denominator for the developing network he hoped would lead to a “quantum leap in human consciousness.” For Ascott, telematic art fundamentally challenges the traditional distinction between observing audience and passive art object, instead creating interactive contexts that draw viewers in through “negotiation, participation, [and] collaboration.” Finally, he saw that the information network (literally, “the web”) that begins to encompass the globe would lead to the next evolution in collective human consciousness (an idea picked up, or arrived at independently by many artists (Bartlem, 2005)).

While the accuracy or desirability of his utopian vision is easily questioned, his enthusiasm is inspiring and his focus on the personal, human, transformative aspects and possibilities is a healthy reminder that technology and its development is ultimately in and for our service (despite frequent appearances to the

contrary). His challenge to passive, physical art and conventional modes of consumption in the twenty-first century supports the ideal of the interactive, participatory engagement of every observer (also espoused by Tofts (2005) and Munster and Lovink (2005)), allowing every individual to shape his/her own experience of the art work and participate in its production and evolution. Due to technological constraints telematic music performance today must limit the number of simultaneous connections and performers, and while the eventual goal is to involve every viewer concurrently as a performer this is not currently feasible. In a way, the performers in the telematic concert hall today act as emissary and representative for their co-located audience, interacting, participating, and collaborating through the network on the audience's behalf, facilitating their engagement with the performance.

More recent definitions by musicians tend to be less particular and more encompassing, such as Dresser's posit: "telematics generally refers to the interface of computers, communication and performance" (Dresser, 2008). Yet, since all telecommunications today are mediated by computers on some level this application could simply be called *distributed performance*. This latter term on the surface claims much less, aesthetically and philosophically, simply requiring that physically disparate performers be connected through a network. In one sense this direction seeks to minimize the impact of the distributed nature, simply extending face-to-face performance over longer distances (in effect enhancing and augmenting the eyes and ears of the performers without requiring any other changes, or acknowledging the ramifications).

This calls to the fore a fundamental divide in the pursuit of telematic music, between modeling current interactions versus seeking new models of musical practice. The former holds conventional face-to-face music making as the gold-standard, measuring every aspect against the common practices of physically proximal ensemble playing (for examples see Olmos et al. (2009); Chew et al. (2004a, 2005c); Dresser (2008)). This seems a natural goal, yet one that may never be obtainable. Given current technological directions and their ultimate advancements, telematic music will still face issues of time synchrony and latency that fundamentally belie the face-to-face model (even transmitting data at the speed of light results in noticeable delays, especially when traversing large portions of the globe).

On the other end of this spectrum are artists and works that use existing practice as a spring board, attempting to discover new ways of interacting and new expressions that uniquely respond to the challenges of networked performance. Dessen (2010) extolls this search, and Braasch (2009) presents a few possibilities that move in this direction, such as exploiting the clarity or obscurity of the connection between locations as a compositional device. These works still employ musicians, privileging the dynamics of human ensemble playing, but rather than mimic the positions and patterns of conventional acoustic practice, new practices are sought. To date there are few examples that convincingly move in this direction, yet the motivational

differences are apparent, guiding the artists to distinctly different choices.

The Stanford Laptop Orchestra (SLOrk) presents one such example (see A.2.8), where the performance actions of all of the musicians, both local and remote, exist primarily in the digital domain. In one notable case the ensemble employed EEG headsets and used brainwave signals to drive digital sound synthesis. The result is a performance that joins the human organism with the computer, enabling connections between performers from multiple locations. Another group pursuing a similar connection is the Avatar Orchestra Metaverse (A.2.9), which uses graphically simulated virtual spaces as platforms for musical expression. In this way they bring people together from many different physical locations into a space that only exists with technological facilitation.

On the other extreme, is it possible to ignore the philosophical and aesthetic implications of networked technology in a distributed performance, and simply play conventional chamber music? Current technology denies the performance of classical era works, primarily due to latency and performer-to-performer awareness (tested extensively by Chew et al. (2004a,b, 2005b,c); Sawchuk et al. (2003)), but many recent compositions that allow for different timings and alignments of parts (such as many works of John Cage) could be adapted for a distributed performance. Yet is a distributed realization of any given work aesthetically (even phenomenologically) equivalent to a conventionally staged performance of the same?¹

Intentional confrontation of the technology and its concerns is the extreme alternative, leading to works in which this impact is seen on all levels of the performance, from conception and presentation, to the musical materials and practical techniques of the musicians. In any case, just as composers of concert music reconcile each piece with the unique affordances presented by particular instruments and ensembles, so must the telematic composer (or performing improviser) consider the unique affordances of the networked medium and place their musical expression in relation. Doing otherwise results in an incomplete and weakly based artistic statement.

2.3 Aesthetic Concerns

In order to identify the fundamental concerns of the composer of telematic music we must first consider the goals of current musical performance practice. While music performance is often tied up with entertainment, it is also built on action, intentionality and evaluation. Musical performance first involves the act of doing, of making the music, live. This performance aspect contrasts with participatory musical practices (Turino, 2008) in that it is staged for an audience, crafted by the performers for the appreciation of the observers.

¹While it may someday be possible to make the two apparently identical this in itself becomes a convoluted objective for the composer, to ignore the aesthetic advances of the technology and replicate a historicized practice.

Involving live musicians in the act of presenting their art also brings the focus to their skill and expression in the moment. Thus, we can understand music performance as *music presented for the intentional evaluation and appreciation of others* (i.e. an audience, at a minimum the musicians themselves).

A further distinction can be drawn between live performance, and the presentation of recorded media for evaluative purposes (such as the practice of electroacoustic concerts, which telematic music closely resembles on a surface level). Starting with the human element of performance is key, as the inventiveness, spontaneity, and variability of the performer's engagement with the artistic work informs and defines an audience's experience of the event. This openness to interpretation is a large part of what makes performance meaningful and culturally valuable, as it is capable of adapting and reflecting the concerns of the moment. While performance is about representation (of ideas, narratives, expressions), according to Phelan (1993) it "cannot be saved, recorded, documented, or otherwise participate in the circulation of representations of representations" without losing its identity as *performance*. The ephemerality, the lack of "left-overs," and the sense of immediacy are defining characteristics. Popular forms of entertainment concerned with reproduction, such as cinema, are not performance, and live simulcasts of opera and other staged events are equivalent, in so far as they are perceptually identical to film. Concerts of pre-recorded media can be valuable and entertaining (and also presented for evaluation), but they are not a performance art, as they are not enacted in the moment.

Telematic music must negotiate this divide as well, as performances frequently involve large projections of video and electronic audio from remote musicians, not physically present in the hall with the audience. Showing the "liveness" of the video and the audio of the remote participants is one of the biggest problems in telematic performance. Without it, the responsibility of engaging the audience, defining the performance, and creating a compelling experience falls solely to the musicians that are co-located with the audience.

Distributed ensemble performances require a sense of "liveness" between the musicians as well. Performers working in collaboration and ensemble need awareness of one another and the ability to interact. Without the sense of being "live," the event becomes equivalent to a pre-recorded video, or a concert for musician and fixed media. Building the live performance aspect is directly informed by a sense of presence and by the agency and responsiveness of the remote performers. While this does not necessarily require physical proximity, as musicians both on- and off-stage are capable of playing together in a symphonic concert hall, yet it does require access to some form of communicative channel, typically video and/or audio.

If such a link can be created for the performers it can be extended to the audience, allowing the musicians to convey their intentions, interpretation, and artistry to the observers for their evaluation. Although the performer's actions may be variously concealed or transformed before reaching an audience, the performer's

presence, agency, and artistic contribution must be perceptible. Without this sense of presence they are not a part of the stage performance (although their efforts may still have a significant impact on the production, just as lighting and sound engineers enable concert performances).

These considerations in turn facilitate the goals of the composer, who seeks to create, control, and direct an artistic work towards a particular aesthetic aim. This requires distinguishable, discernible elements that can be arranged over the course of a performance to construct a relational language for expressive ends (as typified by Lerdahl and Jackendoff (1996)). In music these elements are primarily sonic events, with every work developing or implying its own context for interpretation through the progressive statement of these sonic components. This again requires clarity of presentation, although in telematic performance additional components can become compositional elements, such as: the visibility or obscurity of remote performers (as shown by Braasch (2009), for example), the temporal relationship between performers, and even the sense of impact or existence of the networked locations.

2.3.1 Presence

All human interaction, involving sound and sight, is mediated on some level (by both physical and phenomenological factors). In literature dealing with technologically facilitated communication (i.e. through telephones or video-conferencing systems) a distinction between “mediated” and “non-mediated” interactions is drawn (Lombard and Ditton, 1997) based on the nature of the technology. This conceit arises from the apparent contrast between the “natural, immediate, direct, and real” conception of co-present interaction and the contrastingly unnatural, indirect, and unreal sense of technologically assisted, remote communication.

In a distributed, networked performance the technology is often glaringly apparent and disruptive to the intended activity. However, when such an experience seems “natural, immediate, direct, and real” (Lombard and Ditton, 1997), it is said to create for the user a strong sense of *presence*. This feeling produces heightened emotional responses and is apparently what informs the intense popular interest in mediated entertainment experiences, such as cinema, television, video games, and through the internet. Lombard and Ditton (1997) also note that presence engenders implicit trust, enabling the influence of news media presentations that change people’s judgements of the real world based on information incorporated from highly technologically mediated news sources.

Presence is “the perceptual illusion of non-mediation,” which occurs when a person “fails to perceive or acknowledge the existence of a medium in his/her communication environment and responds as he/she would if the medium were not there” (Lombard and Ditton, 1997). A simple analogy is to eye-glasses that

enable the wearer to see the world. The glasses mediate vision but are phenomenally invisible to the wearer in most situations, allowing the wearer to act and behave with complete disregard to the additional sensory mediation.

The sense of presence is primarily an issue of awareness. As long as the individual in question is focally aware of the technology, of the enabling medium in his/her environment, then they are not experiencing presence. However, when focal awareness of the technology disappears and the viewer approaches and accepts the “mediated” content directly, we understand that presence is in effect.

It is significant that not only is a sense of presence individualistic and relative (i.e. one person may experience a form of presence while another in the same moment or situation does not), but it also does not have a universal standard of truth. Mis-informed or unfounded presence is still presence. Artificially created presence, such as the feeling that an artificial intelligence animating a graphical avatar is “real,” is still presence (this can be seen in the emotional responses and attachments formed with animated characters in film). As an artist, and as a technician, one attempts to encourage the desired sense and aspect of presence in a given setting, yet there is no guarantee of general, comprehensive success.

In telematic music presence is the foundation of musical interactions between performers and a necessary prerequisite to audience reception and understanding of the performance.² Absence, the opposite of presence, is not only a real condition (when the remote participants are not connected or not in their performance space) but also a perceptual one (when a remote musician is visible and audible but appears to act heedless of the musical input of anyone else). In this latter state there is no responsiveness or perceived interaction between performers, the ‘absent’ ones disregarding the musical progression as they are effectively removed and disconnected from the performance.

Without a sense of a performer’s agency and presence the conception of the performance is endangered. Performers conventionally place high value on their spontaneity, expression, and communication with the audience (Blum and Quartet, 1987), and an audience’s sense of engagement is informed by the same. While classical musicians typically refrain from spoken dialog with an audience during a performance the concept of communication is applicable, as ideas and expressions are formed by the performers, conveyed through a musical medium (as variations of intensity, tempo, dynamics, etc.), and heard by an audience. The return loop of communication, from audience to performer, can be very subtle, such as in concert hall traditions, or much more apparent, as in congregational gospel singing traditions where the distinction between performer and audience is intentionally obscured. Yet even the quietest audiences respond and their presence is

²The term “immersion” is often employed to denote the same concept, as is the misappropriated “telepresence.” The former implies a narrowing of perceptual focus, where presence connotes a greater comprehension, and the latter specifically refers to the feeling of “being there,” and is used primarily in discussion and analysis of remote robotic control (Held, 1992).

important to the performers. When this connection between performer and audience is denied, due to the mediation of telematic settings, the communicative relationship is broken and the notion of *performance* is denied.

2.3.2 Agency

In many examples of telematic music performance everything from visible performers clearly making sounds (A.2.1), to performers clearly visible but making no sound (A.2.6), to performers that only seem to appear in the program notes (A.2.8, A.2.14), can be found. This relationship, between a performer’s visible actions and his/her musical effects, is not unique to telematic performance, but is strongly apparent, resulting from intentional choices on the director or producer’s part. While it is largely dependent on technology and implementation, the music that is performed facilitates the relationship as well. Considering the visibility or occlusion of a performer’s manipulations in combination with the perceptibility of his/her musical effects gives us four model cases (based on Reeves et al. (2005)).

The primary operative mode in acoustic concert hall performance is the “expressive” model, wherein the actions—the physical manipulations of the performer—are perceptibly connected to artistic effects. Concealing the manipulations of the performer leads to a “magical” sense, wherein the effects are clearly perceived but the generative gestures are not (note that this is still dependent on a sense of the agency of the performer, especially if they are fully concealed from view). On the other side of the spectrum is the “suspenseful” modality (wherein the manipulations are visible but the effects are not, such as a musician visible through a sound-proof window), and the “secretive,” in which the audience neither sees nor hears the performance. “Secretive” requires the knowledge that the actions and effects are taking place, such as standing outside a radio studio with the “on air” light engaged, even if they are fully hidden from perceptive access. A truly secret performance, of which one is unaware, effectively does not exist to the audience. “Suspenseful” and “secretive” typically rely on a later period of revelation, when the effects become audible or the manipulations become apparent, enabling the observer to reinterpret and understand what has already happened.

The threshold between the display and occlusion of the artistic effects on one level is discrete: either the audience is aware of the effects or not. However, between minimal awareness and full clarity exists a continuum of gradations. If this variation of access is carefully controlled (such as by sonically blurring or diffusing the musical effects) the perceptive movement between “expressive” and “suspenseful” or “magical” and “secretive” may be employed to great compositional effect.

Similarly, movement between showing and hiding the actions and manipulations of the performer is con-

tinuous, capable of a wide range of variation between “expressive” and “magical.” Despite these possibilities telematic performance typically fixes the operative model in the “suspenseful” or “secretive” realm, with performers variously visible, their actions variously concealed, and with no discernible relationship between cause and effect. In many cases the agency of remote performers disappears completely, becoming entirely absent. Simply putting the performer’s name in the program is not enough.

Telematic music inherently and consciously confronts these perceptive modes. The agency of remote performers can be completely absent—as in cases where we are told various remote performers are involved but have no access to their appearance or effects (A.2.6, A.2.8, A.2.9, A.2.14). While they may be contributing in some significant manner, the audience has no access to their impact. Due to the frequent contemporary practice of involving pre-recorded material in performances it is a natural assumption that any sound without a visible source is simply digital playback. Employing a laptop as an instrument reinforces this assumption since the nature and function of computer-based instruments are typically opaque to the observer. If the participation of the remote musicians is intended to be in some way significant to the work, then their tele-presence³ (as opposed to co-presence) must be a focal concern of the presentation.

From the perspective of other performers the same elements of presence and interactivity are equally important. In order to perform with another musician each must be aware of the other, receptive to the musical dialog, and capable of responding. If any of these elements are missing then the two performers are, in a sense, no longer engaged in playing together, but merely playing at the same time.

Once the characteristics of the audience-performer conceptual models are understood it becomes possible to compose with this relationship: one could conceal the performer’s actions then gradually reveal them as if to bring the performer “on stage” for a key moment, before obscuring his/her actions once again (A.2.12). This can serve to draw attention to the music, to the performer’s actions, or to the mediation between audience and performer, bringing the telematic aspects of the performance to cognitive focus.

Awareness of the remote performers’ actions in a telematic performance is enabled by the reproduction of their movements in the audience’s space. This is often accomplished with video, but can also be served by other media, such as very sensitive microphones that relay the rustling of fabric and the breathing of the performer. Similarly tele-present robots or abstract graphical displays could be created to convey the movements of the performers, producing a convincing sense of the agency of the remote musicians. What then is at the core of defining and creating this sense of presence?

³This phrase has been generally appropriated by the field of robotics where it refers to the sense of “being there” experienced while remotely operating a robotic system (Held, 1992). It is unfortunately also used by telematic artists as a generic combination of “telematic” and “present.”

2.3.3 Creating Presence

While many elements have been identified that encourage and feed into the sense of presence (Lombard and Ditton, 1997), I conclude that the primary informer in telematic performance is the perception of interactivity. Mantovani and Riva (1999) similarly finds that more than any technical elements, presence is most strongly based on the ability to “negotiate” with an environment and effectively interact with it. Given the distributed nature of telematic performance the reality and agency of remote participants can always be questioned (i.e. they could easily be prerecorded playback), and thus demonstrating interactivity is largely the responsibility of the co-present performers (who are typically visible to the audience and thus their reality is not in question), through the enactment of the musical performance. This is possible because music follows an internal order, the particularities of which inform our ideas of *style*, involving both specific musical conventions and also performance practices.

On a basic level, *playing music together* involves synchronizing musical changes, such as dynamics and rubato, and in improvised situations, textural and sectional changes. Too many deviations from the set of appropriate interactions casts doubt on the remote performers’ engagement, and hence their *liveness*. Whether the failure to respond is due to automation, oblivion (the remote performer’s audio may be disconnected), or too much delay in the network, the result is a destruction of the observer’s sense of interactivity and presence. While a co-performer may have a more accurate conception of the immediacy of interaction than an observer (i.e. the latter could be fooled into a false sense of “live”), the authenticity of the perception, in any objective sense, is irrelevant. Once the sense of interactivity is instilled (correctly or not) it must remain consistent, or again come into question, further challenging the trust of the audience.

A fundamental problem then is the perception of the “liveness” of the remote video and audio streams, the authenticity of their origin and hence performative relevance. Certainly one could craft a video recording, have a musician practice a live component, and stage a performance of the two, telling the audience that they are witnessing a live duet. Another example might be between a movie, which we know is prerecorded, versus a live TV event, which we believe is presented in real time. Yet both are viewed remotely and we have no definitive access to the immediacy of the presentation.

In fact, the concept of presence is still at work in all of these media, informing the interpretation and reception of the expressive content. Cinema and TV rely on creating a sense of presence, of false “non-mediation,” to enable connection with their audience (Lombard and Ditton, 1997). In the best circumstances the screen becomes invisible, phenomenologically absent, to the viewer, allowing them to become immersed in the artificial world presented by the video. This is the same situation telematic performance seeks to contrive, bridging the mediation of the network and bringing the audience and the remote performer into

communicative contact.

The distinction then is not with presence, but with differences of time, mutability, and expectations of immediacy. If a viewer believes that what they are witnessing is truly “live”, i.e. without the mediation of technology or time, then it is effectively that, and the subsequent expressive ability of the media cannot be denied. However, this is extremely difficult to create, especially in a musical setting where the inter-musician connection is audibly scrutable (and the smallest glitch or missed note is potentially audible). In actual practice the field is generally far from approaching this extreme, and simply fostering a semblance of presence of the remote participants is a challenge. Additionally, as previously noted, performance requires communication between audience and performer and this is impossible in the presentation of media recorded prior to the concert.

In a networked performance visual media reinforces the sense of presence of a remote participant and enables interaction and communication. This is typically accomplished with a video-conference type projection of the remote performers, in a sense bring them onto the same stage as the local performers. Elements that contribute to the believability of the video and presence of the portrayed individuals are (from Lombard and Ditton (1997)): quality, fluidity, and consistency (which break down into: clarity of image, larger size, brighter, close or well framed camera position, fluidity and smoothness of playback, all of which should be consistent over time). This is further augmented by high quality, well amplified audio, projected appropriately for the setting, which emanates from the same source shown on the video (Petty et al., 2008). The particularities of these elements are further considered later in this paper.

Once a sense of presence is created the audience has access to the actions, effects, and connections between musicians during the performance. According to Turino (2008), elements of performer virtuosity, inter-performer coordination, and interpretive spontaneity are hallmarks of the presentational art form. In telematic settings these only become appreciable when the presence of the remote participants is understood by the observer or co-performer.

Yet, simply playing the audio and video of remote participants in a telematic performance is rarely sufficient to engender this sense of presence and connection, and may simply confuse things by highlighting the actions of specific performers without giving them any artistic significance. Theory dictates that large video projections (with optimal size dependent on angle and field of view for the observer Lombard and Ditton (1997)) and high-quality audio systems will foster presence and connection, but the music that is performed must facilitate the relationship as well. In a concert hall the soloist in a concerto is placed at the front, visibly emphasizing their musical role. In the same way the telematic director must consider what the setup is communicating about the priorities of the music, as well as the music’s relationship to the setup

(for example, placing the non-soloist, 2nd bassoon player at the front and making the soloist sit in the wind section causes extreme friction between the musical and situational messages—also seen in A.2.6, A.2.7).

The notion of presence allows for the definition of a continuum covering the range of existent modes, from no presence or absence to hyper-presence and hyper-awareness. Once an observer trusts the medium, and implicitly accepts the agency of the remote participants, the ambiguity (Biocca, 1997) of presence can become an element of artistic and compositional expression. By playing with the apparent responsiveness, presence, and perceptive existence of the remote performers a piece can be continually asking “are they live, or not?” However, this is a difficult balance to achieve. As in cinematic narrative, characters can only have impact if the audience is aware of their existence and has accepted their creation and definition as creative agents. Yet, once instantiated, the removal of a character, even their ambiguous or false termination (to later have the agent rediscovered or resurrected), is often used as a dramatic device.

The notion of presence described thus far is implicitly based on face-to-face models of interaction. However the telematic space allows a new standard to be used in the measure of presence. Through the employ of surround sound speakers, larger than life displays, contact microphones and HD video it is possible to explore notions of hyper-presence and hyper-awareness (Biocca, 1997). These are effectively extensions of presence beyond face-to-face practicalities, achieving super-human levels of projection into remote places. Eventually, real-time biometric data from remote performers driving musical sonifications would begin to explore Ascott’s aim of convolving human, cybernetic, and artificial intelligences.

Through hyper-presence musicians can interact, respond, and connect with one another at a new level of detail. Acoustic musicians reportedly find the sound of the breath of the other players to be a primary indicator of intention, and use their breath to synchronize playing (Blum and Quartet, 1987). As an example, a hyper-present setup can amplify the breath, allowing fine variations in the motion to be heard consciously. It also becomes possible to amplify other changes of the body (such as skin temperature and surface conductivity, meant to be an indicator of emotional response) which may lead to the discovery of new modes of musical communication.

Musicians are typically much more willing to tolerate the close proximity of microphones and cameras, where another person standing as close would be intrusive. The result is a more intimate projection into the remote location than is typically available in co-present settings. The sound captured by a microphone clipped onto a violin is markedly different than the acoustic sound heard in the audience, and yet this is the only access a telematic audience has to a remote violinist, hearing the sound as if their ears were four inches away from the instrument. The same is true of the video, depicting a musician from a vantage of only a few feet away, while the typical concert audience sits many yards away from the acoustic performer. Thus a

sensation of closeness, used so effectively in cinema, is possible in live performance in stark contrast to the actual distances of the distributed performance.

Relaying a single audio and video stream to a remote location is the norm in telematic performance, as this typically encourages the sense of interactivity and sound-to-source identification. However through the use of multiple cameras it is possible to project a performer as seen from many angles and distances, and at many sizes and positions at the receiving end (Ahmad, 2007). The result is a real-time representation of a performer in a way that is not possible in physical, acoustic performance. The audience sees the remote musician in a multiplicity of images, showing a 360 degree view, revealing every gesture in several dimensions. Additionally, these images could be selected and directed individually by each audience member, allowing each observer to create their own unique view into the performance.

2.4 Latency

The issues described thus far are primarily of a perceptive nature and deal with the aesthetic and expressive challenges and potentials of telematic music performance. However technical complications have wide reaching ramifications that further inform work in the telematic domain.

The foremost of these challenges stems from the physical distances involved in a distributed ensemble and the resulting delays in transmission and reception of video and audio data. Latency is a measure of the time it takes for sound (or image) to be captured, transmitted, and reproduced at a remote destination. The amount of time for a given connection is comprised of the following elements:

- Microphone transduction and transmission as electrical signal through copper cable,
- Analog-to-digital conversion in a computer audio interface device,
- Buffering, packetizing, and sending of the digital audio out of the computer,
- Transmission of the audio packets over the internet, typically as UDP datagrams,
- Reception, buffering, and unpacking of the network packets on the destination computer,
- Digital-to-analog conversion in a computer audio interface device,
- Transmission of the electrical audio signal through copper cable to a speaker where it is reproduced at the destination.

Each of these steps adds additional time (on the order of milliseconds) to the path, and serve as the focus for continuing development and research in music and technology communities. The largest contributor

to telematic latency is the transmission of audio packets over the internet and the requisite buffering that enables continuous playback. The most popular audio solution, JackTrip (Cáceres and Chafe, 2010), employs eleven millisecond buffers on both ends of the connection and sends each packet three times, in order to best ensure complete transmission and accurate, CD quality playback. However, audio interfaces, required at both sides of a connection, can also add significant amounts of time (anywhere from two milliseconds for professional quality hardware to twenty milliseconds and more for interfaces built in to consumer laptops). In the case of short network distances these interface delays become the biggest hurdle to low-latency ensemble interactions.

For reference, sound travels at nearly one millisecond per foot in our atmosphere. Thus, it can take 20 to 40 milliseconds for the sound from the principal french horn player to reach the conductor in a typical concert stage. The delay is even longer between the timpanist (located at back, stage left) and the back of the first violin section. These delays are compensated for in an orchestra and the sound, ideally, arrives at the conductor in a desirable temporal alignment. The same issues inform the anti-phonical choral music written for large cathedral settings (such as the works of Jacob Handl), where the acoustic latency between choirs and organ could be as much as a second or more.⁴

Typical audio network latencies today are (using JackTrip and 1 gigabit-per-second high-bandwidth networks):

- under 20 ms for local networks (i.e. with one or two switches, in one building),
- 30-60 ms for connections across parts of the US,
- 240 ms between Urbana, IL and Melbourne, Australia.

As previously noted, video transmission is typically even more delayed, resulting from longer capture times and higher bandwidth requirements. Due to the significant limitations of current solutions (in terms of quality, flexibility, and speed), I developed and released a video compression package, *uiuc.jit.codec*, for telematic performance based in Max 5 (a commercially available software environment for audio and video programming and performance). This new software exposes a number of popular compression algorithms and standards (including *mpeg2*, *mpeg4*, and *h263*) for real-time control, and facilitates point-to-point streaming of live video. Additionally, *uiuc.jit.codec* makes available several lossless codecs (compression-decompression algorithms, namely *huffyuv*, *ffvhuff*, and *ffv1* (FFMpeg)), which have the distinct advantage of transmitting and reproducing the original image exactly, without any degradation or compression artifacts. While the later

⁴This practice of antiphony was a primary inspiration for the early telematic musicians under the tutelage of Pauline Oliveros (Oliveros et al., 2009), which can still be seen in a vast number of recent concert works.

do necessitate larger bandwidth allowances, the fidelity of the lossless compression is currently unavailable through nearly any other solution (and certainly not as accessible).

Clearly music performance, as conventionally practiced, is challenged by significant time delays between the members of the ensemble. However, telematic researchers have yet to ascertain its impact or devise a metric for measuring its effect. Some projects (such as Sawchuk et al. (2003); Chew et al. (2004a,b, 2005b,c,a); Olmos et al. (2009)), conclude with ranges of numbers based on the opinions and experiences of a handful of study participants. However, they also note that the typical coping strategy on the part of the performers is for one member to ignore all the others and serve as a fixed, musical reference point for all the others. While this not only challenges the notion of *playing together*, it also calls in to question the validity of the realization of the musical piece. If the amount of asynchrony causes that much distress then where is the actual “piece” heard, “correctly” aligned?

Schuett (2002) and Chafe and Gurevich (2004) tried to quantify latency’s effect on ensemble accuracy by observing pairs of musicians trying to clap together in rhythm with various network delays. These studies, which included sizable test populations, concluded that delays of 10 to 20 ms were ideal, allowing the players to accurately maintain a tempo, while longer delays caused the performers to slow down, and shorter delays to speed up. While this work is often cited, it has some significant problems. First, the test population consisted of anyone who could clap, and yet the ability to keep a steady beat is not a given in our society. Correlating the results with the participants capacity for maintaining a steady rhythm on their own might have provided significantly different results. Second, the conclusion that the musicians speed up when the delay is less than 10 ms is bizarre. Musicians are perfectly able to maintain steady tempos in co-present situations, and yet this study indicates that they should ideally stand ten to twenty feet apart or else they will tend to accelerate (and conversely standing farther away they will slow down). Why these effects were observed in these studies remains to be described or explained.

However, human perception and cognition has been the focus of much research over the last century and a half, resulting in a high level of understanding of the workings of the human organism. From this work it becomes possible to formulate metrics for the measuring of telematic latencies, and the prediction of its impacts and resulting aesthetic possibilities.

Typically, studies of human aural temporal discrimination examine the ability of a participant to differentiate two sounds occurring in rapid succession. This differentiation is measured at two points, first as the ability to accurately say that there are two sounds versus one, and second the order of the two sounds (by playing one sound on the left and one on the right). The threshold at which the participant can identify not only that the sounds are separate but which one occurs first is considered the threshold of perceived

simultaneity.

The common finding is that 20 ms is the point at which most humans can identify the sequence of strongly attacked sounds (Pastor and Artieda, 1996).⁵ This is dependent not only on the individual, but also on the sounds being examined—two legato string notes are much harder to separate than two glockenspiel strikes (studies typically employ electronic clicks). Below 20 ms, and down to about 3 ms, individuals can still identify the asynchrony of the sounds, but not the order. Going even shorter, down to a few hundred microseconds, the brain processes the sounds as one, but uses the time difference as an interpretation of source direction.

Processing of image has been similarly studied but has more variables. Visual stimuli presented sequentially is processed as movement, so long as the distance between images, brightness, and display rate are tolerable. Again, 20 ms has been found as the ideal presentation rate (given suitable brightness and distance of change) for the brain to perceive movement. The lower limit is largely dependent on brightness—the light from an image or object must be strong enough and last long enough to register.

The transduction of light in the eye and sound in the ear each operate at different speeds, producing a striking discrepancy. While the eye takes 50 ms to convert light stimuli to neurological stimuli, the ear converts sound in less than 5ms (van Eijk, 2008). Thus sound and light arriving at the ear and eye, respectively, reaches the brain roughly 45 ms apart. Apparently, as a result of physical stimuli (the speed of light being much faster than the speed of sound in our atmosphere), our body perceives simultaneity even when sound arrives significantly later than image or light. This in turn has very interesting implications for telematic art where the transmission of the two media occurs at different rates.

Studies testing the limits of the perception of audio and video simultaneity have found the range to be +50 ms (with audio arriving before video) to -150 ms (with video arriving before audio)(van Eijk, 2008). This also parallels situations in the physical world, where images typically arrive at our eye before the corresponding sounds arrive at the ear (based on the slower speed of sound compared to that of light). Studies independently found that -50 ms (video arriving at the eye 50 ms before the corresponding audio) is the most natural, easiest perceptible alignment for visual and aural stimuli, which fits with the transduction times just discussed.

While many studies have found surprising fidelity and repeatability over large test populations, all of these numbers cited are somewhat approximate and vary from individual to individual. Apparently, no study has attempted to quantify the effects of extensive musical training, or other practice that requires intense listening or observation on a regular basis, and thus we can only hypothesize that musicians may

⁵Pöppel (1996, 1997) presents a convincing argument generalizing this threshold at 30 ms, however herein we retain the more stringent 20 ms limit.

perform on the lower, more precise end of these discrimination thresholds.

Based on these numbers we can make the following presumptions for networked music performance:

- Audio network latency of 20 ms or less ideally mimics face-to-face delays.
- Audio delays up to 50 ms match those observed in orchestral concert settings and can be handled in similar fashions.
- Video frame-rates of 50 times a second are ideal for portraying the movements of remote participants. Faster than this is apparently imperceptible while much slower causes discomfort in the viewer, who has to consciously reassure themselves of the fluidity and accuracy of the image.
- Video that arrives after the audio by up to 150 ms can still be considered *in sync*. Synchronized media makes interaction and presence easier and also enables conducting type gestures.

Methods of artistically treating and responding to latency diverge in two directions, one that seeks to downplay, minimize or even ignore it, and another that attempts to emphasize the unique implications of the distributed performance setting. These mentalities can be seen in all aspects of the telematic works, but are most dramatically evident around issues of latency. The former desires, ultimately, to reproduce face-to-face practices, removing the mediation from perceptive access, enabling a continuation of contemporary concert hall performances. This typically leads to musical results that are full of amorphous, aleatoric, improvised textures, or rhythmic material that allows for the looseness of ensemble that results from the small network delays (see Dresser (2008)). Another approach uses the period of the latency as a beat in the music, allowing the musicians to play rhythmically together but offset metrically. This was employed for a trans-Pacific concert by Cáceres et al. (2008), however, the choice of music in this case (Terry Riley's In C) effectively erases the latency, as once the proto-minimalist ensemble texture is achieved it remains consistent throughout the work (the performance was deemed successful and reportedly repeated between CA and Italy a year later).

My own works of telematic music attempt to exploit latency to create new aesthetic experiences. The single biggest implication of latency is the unique alignment of parts at each physical location involved in the telematic performance. This change is typically ignored in the literature, either allowing the music to happen randomly (in effect), or by designating one location the 'main stage' and cheapening the experience of the remote participants. However, it is possible to compose works where each location witnesses a rich and satisfying musical piece, each uniquely synchronized and balanced but employing the same component musical elements.

This impact of latency emphasizes a fundamental aspect of telematic performance: that the musical piece is dramatically unique at each location. This is perceptual relativism to the extreme, as in telematic music there is no one true version of the piece. Not only do the presence of the acoustic instruments and the acoustical properties of each place make the sound different, but the alignment of the parts will be offset differently at each location based on the topography of the network connections. While this is also applicable in an acoustic concert hall (for example, an audience member sitting in front of the bass players on the audience's right hears a different sonic alignment from the person in front of the percussion, on the opposite side of the hall), it is rarely a concern and is otherwise unavoidable. In telematic music the offset of the parts can often be felt as an eighth-note or longer, and serves to highlight the fact that the reception of any piece of music is an individual experience, every perception is valid, and no single true, or correct version exists.

Musically, latency turns syncopation into unison, synchrony into asynchrony, and anticipations into arrivals. When two telematically distributed players appear to be playing in rhythmic synchrony at one location they will be syncopated, or offset at another. When a player intuits the time of the transmission they can anticipate musical arrivals, causing their remote partners to hear a unison attack (just as orchestral brass players play ahead of the 'beat' to ensure audible ensemble alignment at the conductor's position). When the latency is long enough it can be felt as a regular period in the music, such as a beat or a bar, allowing the players to play complex rhythms in alignment, but effectively phased by one unit of the period. This can lead to very satisfying canonic musical passages (still different at each location), where the tempo is enforced by the distance of the connection.

2.5 Musical Dichotomies

The pursuit of face-to-face telematics is apparently such an obvious choice that its reign goes largely unquestioned. However, this pursuit may be based on an even deeper assumption about playing music. Playing together, in nearly every musical tradition, involves a shared temporality, a sense of pulse or passing of time that is felt by all the members of the group. This can be seen on a wide spectrum, from hard-bop jazz that requires millisecond accuracy, to loose participatory styles that allow much more rhythmic variation yet still consider everyone to be playing together. The tradition of antiphonal choirs, as noted previously, takes this further, requiring each individual choir to be tightly synchronized but allowing the opposing choirs more leeway in their alignment. Thus it must have seemed self evident that research looking at telematic music making would attempt to locate the point at which the delay made playing together possible or impossible

(and many have proposed numbers that define this threshold—see Chafe and Gurevich (2004); Chew et al. (2005a); Olmos et al. (2009); Schuett (2002)).

However, another option is to reconsider the assumptions that inform this pursuit. Sharing of time and rhythm informs the ideology of the *true piece*, the idea that a single correct version of the work, and of the performance of the work, exists. For example, in an orchestral performance this “true” version is typically heard by the conductor who stands at a focal point for the sound of the ensemble. However, the distances between the instruments in addition to the performer’s compensation for acoustic latency causes the sound to align differently at every point in the hall (with offsets from tens to hundreds of milliseconds between the parts). In extreme cases notes sounded ostensibly simultaneously by orchestral musicians on opposite sides of the stage will be heard a sixteenth-note or eighth-note apart in different places in the hall.

This is precisely the same challenge faced by telematic ensembles, where the delays are equally long and much more direct (i.e. the sound is not softened by the acoustic dispersion and reflections present in a concert hall). Thus each location in a distributed ensemble hears a unique alignment of the parts, creating a unique version of the piece being performed. If the delays are small enough the participants may agree they are hearing the “same thing,” yet if the delays are large it becomes futile to argue that the piece is equivalent at both ends.

Based on psychoacoustic data presented earlier we can argue that the low threshold for piece equality is around twenty milliseconds, with greater delays causing perceptibility different alignments of the music. However, depending on the nature of the music being played it may or may not be relevant to locate the musical differences and attempt to define the identity of each version (for example A.2.6 or Cáceres et al. (2008)). Each version may still be considered equivalent based on aesthetic considerations even with large delays.

At a certain threshold (pending stylistic and aesthetic applicability), the musical events will perceptually align differently, for any piece (although the threshold will be different in each case). Music based on long slow un-foldings over long time scales (such as A.2.6) will tolerate much longer delays than tightly synchronous, pulsed, rhythmic music (such as A.2.2). While different rhythmic composites may be the first revealing characteristic, differences in harmonic alignment make clear distinctions as well. This is especially apparent in situations where one location witnesses a unison line and another, due to the time offsets induced by the network, hears vertical harmonic simultaneities.

The following figure depicts this principle at work. The middle staff shows a melodic line played at the first location in a two part telematic performance. The lower staff is the part played by the second musician at a second location in response to the first part (downward arrows, note that the pitches and rhythms of

the lower staff are a unison copy of the middle staff). The sound from the second player is then transmitted back (arrows up from bottom staff) and heard by the first player offset by one quarter note. The composite heard at the first location is the combination of the top and middle staves. The second location hears a unison line of the bottom staff played by both musicians.



Figure 2.1: Middle staff played at location A is heard and matched with bottom staff at location B. Top staff shows alignment of part B as heard at location A.

Works may be composed that exploit this temporal schism intentionally and dramatically, creating different aesthetic expressions at each location that are unique sums of the component parts. When the discussion revolves around a few milliseconds it may seem superfluous⁶, yet if the delay is undeniably perceptible (such as a half-second or more, where it is felt as an temporally organizing pulse in the music) it can become a defining element. In these latter cases a line that sounds in unison at one location (say *Twinkle, Twinkle*) will be heard with the two parts offset by a beat (or two beats) at the other. Thus the simple diatonic motion of the melody becomes major and minor seconds, throwing the two contrasting (and still simultaneous) versions into stark relief.

The aim here is a new way of conceptualizing “the work.” If each location in the distributed ensemble hears a different piece (i.e. different temporal alignment of the same parts) they are no longer sharing the musical togetherness of concert hall ensembles. However, they are still sharing the same musical work, as each of the different pieces depends on the same parts, the same musical utterances from each of the performers. Additionally, musicians playing in a telematic performance often are fully convinced of the interaction and engagement of the whole ensemble and embrace the “togetherness” of the group (see Appendix B). Thus “work” and “piece” must be expanded, or wrapped into a new term that embraces the multiplicity of musical experiences stemming from a single set of parts aligned differently in space and time.

Time is the biggest factor in defining the identity of each of the resulting ‘pieces’ in a telematic performance, but many other elements may be used to inform their distinction. Moving on or off a microphone,

⁶However, even at this granularity one location may be able to execute a tightly synchronous rhythmic passage with matched articulations, while another location hears a smeared, dispersed version resulting from the phasing of the ensemble attacks.

keeping the acoustic sound present in the local space but disconnecting its transmission to remote places, reveals and highlights the multiplicity. In this case one musician becomes a listener only, receiving the audio of the remote performers but not sending any sound in return. Yet, the local audience and other performers continue to hear the full ensemble. A striking example of this is a distributed duet where the accompanying musician stays on the microphone but the soloing musician steps away, leaving the accompanist in the musical ‘dark’ as to what is transpiring. Involving technicians actively mixing the streams of audio differently at each location would serve this same effect and allow even more degrees of variation (and play heavily with the notions of agency).

A useful analogy of this new multiple-piece construct is to windows into a chamber, each providing a view into a musical space from a different angle. Each orientation portrays the same space, the same objects and boundary walls (possibly the other windows as well), but each from a divergent direction. These views (pieces) can also be likened to facets on a classically cut diamond, each of which allows the entry and exit of light at different angles, creating continually sparkling refractions and colors. This analogy calls to the fore the influence of each face/window/telematic-connection on the performance, as each functions as both an inlet and an outlet for musical content. Just as a diamond with a single face appears dull, a telematic performance necessitates multiple locations and connections to retain its aesthetic distinction. Finally, the proverbial blind men and the elephant presents a strong analogy to telematic performance: just as each man describes a unique and contrasting aspect of the same animal based on their limited perception, so each location in a telematic performance experiences a different (and potentially contrasting) musical construct reflecting the same central work.

From the vantage of logical comparison we can argue the import and validity of the multiple-piece aspect of telematic music performance, yet, for an observer in a concert hall this aspect may be entirely opaque, as each location only bears witness to one arrangement and has no access to the others for comparative purposes. Knowing that each is different, on an intellectual level, does little to change or augment the expressive impact of the performance at a given location. In this way the notion of a “telematic musical performance” sets up expectations for the viewer of network visibility and impact, while the actual experience is typically perceptively equivalent to a non-telematic, mixed-media performance.

The solution to this problem may be two-fold: first, each location’s version of the work ought to be complete in its own right, leaving the notions of multiplicity as intellectual rigor for the creators alone. Towards this end the telematic aspect becomes purely functional, enabling a performance by employing networks but focusing the aesthetic expression on other elements. Connecting specific individual musicians into a performance for their unique musical contribution exemplifies this case, and keeps the presentational focus on the

musical content separate from technical considerations. Telematic connections also allow creators to bring together equipment, instruments, and acoustical spaces in a single performance that would be otherwise prohibited. For example, a cathedral organ, which has extreme space and construction requirements, employed in a telematic performance with a chamber concert hall brings new musical possibilities to the smaller venue. However, this solution requires an intentional change to the advertised expectations, down-playing the aspect of “telematic performance” and emphasizing the content of the performance.

The other solution requires the uniqueness of each local version of the performed work be shown in the enactment of the concert—by having players gradually recede off camera, disconnecting and reconnecting to create “suspenseful” and “secretive” sections, or be audibly playing through the network but as if at a great distance—bringing questions as to the identity of the piece at other locations to the fore. This is certainly a challenge to the composer, who must consider and create each individual piece in addition to the perception of the remote “pieces” at each location. Musically this is a matter of playing with presence and agency, presenting and emphasizing the contrast between local and remote perceptions of the same performance elements.

While an individual observer cannot witness more than one version of a telematic piece at a time, the idea that something is happening differently at another location (such as a soloist that can be seen on a video projection playing an emotional melody but is not heard over the network) can be instigated and transformed into a compositional tool. Leaving the participant with questions, wondering what transpired behind the ‘closed doors’ of the network, is a challenging and promising objective for the composer.

Employing verbal dialogue or narrative presents perhaps the clearest space for depicting the different work-views within a telematic performance, as spoken language typically relies on more explicit and accessible forms and structures (compared to those of contemporary art music). Setting portions of the text solely in each location, modeling the telematic connections as windows or doorways between rooms of a house, plays into the audience’s embodied knowledge and experience of conversations that continue even while the viewer is absent. Thus characters can engage in dialogue between locations and present monologues that are only accessible to the local audience. This portrays vibrantly different angles of the same work, that would ultimately be appreciated best by attending each location of the performance in sequence, revealing all of the content over the full course of viewing.

This approach to telematic music performance becomes as much about disconnection and restriction of access as it is about connection and sharing. The creation of each unique piece here is strengthened by convincing the viewer of their disconnection, first of their inability to perceive the work as a whole and that their view is in some way incomplete, and then of their inability to perceive the work as seen by other remote

audiences. Denying each audience access to certain elements of the musical work, and making this exclusion explicit, can further reinforce the special identity of each location in the telematic network. Each view of the work thus becomes an exclusive event, relying on and reflecting the other views, yet existing ephemerally in each physical location as a unique experiential performance.

While the primary thrust in telematic music to date has been focused on connecting performers as seamlessly and transparently as possible (ideally reproducing the sense of face-to-face interaction), this next step in the art form may be about subtly disconnecting the ensemble and making the technological mediation explicitly visible. Recreating co-present performances poses myriad technical challenges but does little for the development and advancement of contemporary art practices. Embracing the ‘difficulties,’ on the other hand, and bringing the disconnection and lack of access to the aesthetic foreground presents dramatically new elements to the compositional and performative domain.

Chapter 3

Experimental Research

3.1 Methodology

In order to better understand the problems of telematic music and formulate the operative concerns, challenges, and potentials I conducted a seven month long series of experiments in distributed ensemble performance. The method of inquiry for this research was posed as *Research through Art*, as described by Dallow (2003) and Schiphorst (2009), extending the principles of *interaction design research* as described by Zimmerman et al. (2007) (all of which are inspired by (Frayling, 1993), “Research in Art and Design”). This methodology holds that the art researcher is uniquely positioned to tackle under-constrained problems, and contribute to the field through the production of artifacts that push the world they exist in from its current state toward a preferred state. These artifacts are intended as carefully crafted questions, exposing assumptions, facilitating the mapping and constraining of the problem space. Rather than begin from a known problem, the art researcher constructs “experiential starting points from which practice follows,” diving into new, unmapped areas of development (Haseman, 2006). See Appendix B for further methodological details and documentation.

3.2 Technology

The continuity of the experiments allowed for the direct comparison of a number of different supporting technologies and setups. Before proceeding to the findings of the research I will first describe the technical setup and considerations in selecting the best solutions.

Since the goal of this work is to address the potential for distributed ensemble performance I began with the current popular setup of musicians connected with CD quality audio streams and video-conferencing software. Before the experiments began I was able to test and discard Skype and iChatAV (both of which are used in many performances, see Appendix A) due to the poor quality and lack of configuration capabilities. The Scalable Adaptive Graphics Environment (SAGE, Jeong et al. (2006); Renambot et al. (2004)) was

similarly evaluated and discarded based on unreliability and extreme configuration costs (in terms of setup time). Thus the experiments commenced using Max 5's Jitter video and audio capabilities, which allowed for real-time streaming of uncompressed audio and video.

The setup for each session employed two musicians in two different rooms in the National Center for Supercomputing Applications, connected by a 10 gigabit-per-second network. One room had a large (approximately 10-by-30 foot) projection screen and the other a 30" Apple Cinema LCD display. Both had professional quality audio amplification systems and microphones. The setup was arranged to minimize the obtrusiveness of the computers by having the musicians face the screens, placing a small video camera in front of and aimed at them, and setting a single microphone on a stand directly in front of each musician. Musicians stood if possible, or sat on a chair with plenty of room around them (to mimic the feeling of being on a stage). They were strongly lit from above (to increase their visibility as well as the projection screen or monitor display).

The initial audio solution proved problematic over many sessions, and while it had a very low latency (on the order of 20 ms), it introduced many audio artifacts (heard as clicks and beeps). Thus it was replaced with Jack Trip, which performed comparably in terms of delay but with far fewer distortions.

Streaming uncompressed HD video was functional but proved challenging, producing unexpected delays that were unresolvable. The best guess is that additional traffic on the local network caused backups in the TCP data, which resulted in buffers filling up and delaying the video frames in transmission. Thus the experimental log (in Appendix B) notes that video delay amounts changed dramatically over the course of a two hour session (from roughly one-quarter to two seconds). Several additional attempts to use this software to connect to distant partners (namely Louisiana State University in Baton Rouge, LA) failed due to packet loss causing extreme delays (on the order of sixty seconds or more). This eventually led to the examination of several other solutions.

ConferenceXP was used successfully for several performances between UIUC, NY, IN, and Japan, but the software's dependency on Windows OS was deemed prohibitive. Additionally, it does not provide any extra features for artistic use, and camera support was found to be unpredictable (especially for the HD capture cards and cameras that were available).

Unreal Media Server was employed for one session, and observed in use at several sessions between UIUC and the Sonic Arts Research Centre (SARC) in Belfast, Northern Ireland, but discarded due to long and detailed configuration requirements. Also, the end result was equivalent to ConferenceXP in quality terms (i.e. low resolution support and limited compression options) and is also dependent on Windows OS.

The final solution was to implement a new codec package for Jitter, providing real-time control of the

compression parameters during live streaming. This new library, *uiuc.jit.codec*, was built on the open-source package *ffmpeg*, implementing a large number of popular and commercial video compression algorithms. The performance of the software remains to be tested empirically, but initial results show compression times of approximately 10 ms for HD 720p frames using mpeg4 (resulting in a 90 to 98% reduction in data size), and much less time for decompression (5 ms and less), on consumer grade hardware (a Mac Pro tower).

The resulting amount of data space that is saved indicates not only greatly reduced transmission times (since the receiving computer has to receive a whole frame before displaying, it can commence much sooner since the data is 70-90% smaller) but also the ability to use smaller internet connections (such as those used by consumers). This opens the potential to transmit multiple streams simultaneously, where the uncompressed video required most of the available bandwidth for a single bi-directional stream. The connections employed across the Illinois campus during the final presentations of the musical component of this thesis leveraged this capability to advantage.

In addition to the popular lossy codecs (compression-decompression algorithms, such as mpeg2, mpeg4, and h264) *uiuc.jit.codec* implements a number of lossless algorithms (namely *huffyuv* and variations). This means that the image that is captured at the source is reproduced on the receiving side exactly, without any transformation as a result of compressing and decompressing. The tradeoff is a slightly more costly computation load (requiring more CPU power and time) and a larger bandwidth requirement (only compressing by roughly 60%). However if these expenses are acceptable the higher quality appears to be noticeable on many levels.

Another benefit that I was able to provide through *uiuc.jit.codec* is the real-time control of many of the codec's configuration parameters. These allow not only for dynamic optimization, based on knowledge of the video's content (such as compressing more for relatively static scenes and less for moving dancers or musicians), but also aesthetic possibilities resulting from intentionally over compressing or blurring the images. The exposed control parameters consist of the codec choice, the target bit-rate, the frequency of key-frames, and the option to interlace images in addition to the frame rate and frame resolution (controlled through standard Jitter objects). The result is a highly configurable system that is easy to use right away.

Already this software has been employed during performances on the UIUC campus, between UIUC and Melbourne, Australia, a dance performance between UIUC and the Sonic Arts Research Centre (SARC) at Queen's University in Belfast, Northern Ireland, and for test sessions between SARC and Brazil.

The final sessions continued to extend the practices of face-to-face musical ensemble playing, but by the end the cameras were being positioned to capture more unusual and interesting angles (such as focusing on the keys of the piano or the fingers of the flutist) and the audio software had been enhanced to allow

dynamic control of the amount of latency. Microphones were placed very close to the sound sources, both to reduce system echo and feedback, but also to relay a very intimate and personal sound perspective to the remote locations. The results were still strongly rooted in acoustic ensemble practice, but sought to move into new realms of aesthetic possibility, as described in the following sections.

3.3 Findings

3.3.1 Latency

Through my telematic experiments I discovered two foundational effects of latency and how it coerces a musician's playing as a result. First, latency lengths can be categorized into five areas that encourage different musical actions, and second, a musician's conceptualization of the length of the latency is the primary guide as to what they play. Let us consider the categories first (note that these numbers indicate round-trip latency times).

1. Under 20 ms forms the first category. Here the delay is imperceptible, as previously noted, and musicians are able to play anything they would play in a co-present situation. This is very, very difficult to create in a distributed setting.
2. The next category is roughly between 20 and 120 ms. This range is audibly asynchronous, but too short to feel as a pulse or beat. The result is that musicians are drawn to slow, loose textures that do not require tight rhythmic alignment. Playing pulsed music is very challenging and problematic in this range.
3. Between 120 and 400ms the delay is heard as a part of a beat, typically an eighth- or sixteenth-note. Slow, loose textures are still possible, but rhythmic music that plays at the period of the round-trip latency also becomes possible and can be very rewarding.
4. 400 to 600ms comprises a range that aligns with most musician's "spontaneous tempo" or "foot-tapping tempo," which is the speed that people start tapping or clapping when asked (Krumhansl, 2000). Thus it is very easy to perceive the delay as a beat and play rhythmically at this rate. Slower, looser music becomes even more out of sync, showing the interactions to be evidently delayed.
5. Over 600ms and the delay is difficult to track as a beat, and tapping at a subdivision of the time is very hard to keep synchronized. Thus pulsed, rhythmic music becomes very difficult again without some sort of mechanical assistance, such as a metronome or click track.

These five categories of latency encourage possibilities but do not enforce them—they were formulated as the result of many musicians attempting musical devices, failing, and arriving at these conclusions. I have observed that the two most comfortable ranges are under 20 ms and between 400 and 600 ms, which both allow rhythmically synchronous and free tempo playing. The “uncanny valley” hypothesis (Mori, 1970) observed in human-robotic interaction is potentially a useful analogy here, as delays that are either radically like face-to-face interaction, or radically different are the most easily approached, musically. Delays in the middle, in the “uncanny valley” of latency, prove problematic, denying many conventional musical forms and learned musical practices.

Over many sessions I was able to observe that once a musician forms a mental conception of the length of the latency they retain this idea and it is very difficult to change. Even when conditions change (i.e. the delay increases or decreases significantly) the musician typically does not reevaluate the latency unless the situation is so dramatic that it is severely obvious. However, musically showing the latency is very hard to manage even intentionally, thus most players form an idea about the delay at the start of a performance and maintain it through the end. Actual jitter in a network connection is typically very low, meaning that the amount of change in the ‘natural’ latency only varies by a few milliseconds during a ten-minute piece. Thus, latency preconceptions only become problematic when a composition artificially alters the delay (such as Appendix B.2.4–6), typically attempting to setup certain musical effects, as the musicians are disinclined to notice the shift and hence change their musical direction.

3.3.2 Video

The impact of the video comprises another significant area of inquiry in telematic performance research. Video displays and projections are frequently employed in telematic performance, ostensibly to encourage interaction, presence, and inform the sense of agency of the remote musicians. However, its adoption should not be taken for granted and the particularities of its use must be examined.

The first important question is: what are the implications of removing the video stream (thereby getting at what adding video signifies)? My experiments, as well as evidence from the literature, show that musicians are still capable of playing together without video (see B.2.5, B.2.10, B.2.12, and Oliveros et al. (2009)), however certain situations become more difficult. If the video is gone, the responsibility for showing the liveness, interactivity, responsiveness, and agency of a remote musician falls solely to the sounds coming through the audio network streams. This is less of an issue when only two musicians are playing, but very problematic when there are many musicians at each location. Image is very helpful in informing a sense of presence when the sonic field is very dense. Thus removing the video may necessitate musical changes that

can be restrictive.

Without video the lack of postural cues and communication can have a significant musical impact in many situations. While the asynchrony of video and audio and the measurable delay times of both often preclude precise gestural cues, the way a musician moves and how they stand communicates a lot about their intentions and their attitude in the moment. With the video this is transparent, and musicians employ a lot of the same subconscious movements that are seen in face-to-face interactions. These motions typically facilitate the intensifying or relaxing of dynamics, tempos, articulations, textural densities, and the like, as an ensemble. Postural cues are most evident at moments of starting and stopping, where players conventionally look at one another to coordinate clean changes. When the video element is removed these points are thrown into stark focus, forcing evident constraints on the music being played.

For example, stopping an improvisation takes much longer without video, as the only way to communicate a clean ensemble ending is to gradually fade out, which may yet be ignored by others who wish to play further. Other musical alternatives are easily lost on distributed musicians, who can interpret them as a desire to start something new. Starting together can only be accomplished with an aural cue (such as counting “1, 2, 3,” seen in A.2.1), and changes in dynamic and texture require an intense focus on listening. A video representation provides focus for these issues, allowing the distributed ensemble to interpret the sounds of their remote counterparts without forcing the music to bear the full burden of coordination.

Video components in a telematic performance have unavoidable aesthetic consequences, in addition to the functional considerations presented thus far. My experiments found that close camera angles and positions that show the movements of the performer but also leave significant elements off screen are generally preferable (i.e. artistically richer) to wide angle shots that show a performer against a motionless backdrop. Capturing faces is also a consideration, as humans typically find a sense of connection through facial expressions. Depicting local actions as projected in a remote space and fed back through the network can also be evocative, as this serves to show the looped nature of the interactions and the liveness of the display. This can be further enhanced by displaying video of the audience—allowing spontaneous “wave at myself” tests to prove that the whole system is functioning as promised.

Due to the size of the displays employed and the seductive quality of video images, the audience’s attention can be easily consumed by the portrayal of the remote performers, minimizing the impact of the local musicians. As the remote musicians are projected larger than life, their hyper-present nature overwhelms the merely present nature of the physically co-present performers. Displaying the video of the local, co-present performers (which is being sent to the remote locations) next to the video of the remote performers can balance this issue, and simultaneously reinforce an observer’s trust in the system and further

emphasize the ensemble interactions of the musicians. While this still subsumes the co-presence of the local performers it succeeds in bringing all of the musicians onto the same virtual stage, in the form of the projected screen, encouraging the perception of group unity, despite their physical separation.

Another consideration is the visible, physical interaction of the co-present performers with the projected video and the co-located audience. Placing the performers in front of the projection, and having them face out towards the audience, creates a perceptive disconnect as the co-present musicians apparently ignore their remote counterparts. On the other hand having the physically present musicians turn to face the screen conceals their faces from view and reduces the audience-performer connection. Possible solutions may be the incorporation of projectable surfaces set amongst the co-present ensemble, allowing the remote video to be shown in the middle of the stage, in the same position that the remote players would assume if they were physically present. Another possibility is the use of a semi-transparent scrim or screen hanging between the stage and the audience, which accepts the projection of the remote video. This would allow the local musicians to see the video from the stage side, facing out towards the audience, while the audience sees the remote video and the co-present musicians through the screen.

Finally, despite the logical arguments for the inclusion of video streams in telematic performance, my experimental and concert experience has shown that performing musicians almost exclusively ignore the video. After the occasional glance to start or end a piece the performers rely primarily on their aural faculties for ensemble interaction and synchrony, using the video mostly during non-playing conversation and setup times. However, its impact for the audience is significant, as it is the primary method of demonstrating the liveness and showing the agency of the remote ensemble. Although the video element may seem superfluous during a concert, its absence is immediately apparent and a cause for confusion and consternation in the attentive observer.

3.3.3 Intimacy

One of the most intense experiences that arose during the research experiments, both for myself and others, was the incongruous sensation of closeness with remote participants. Throughout the experiments I repeatedly observed, and felt, a strong musical connection, interaction, and dialogue on a level rarely achieved in face-to-face situations. This notion was reinforced by feedback from other participants and leads to the consideration of intimacy in a fully mediated environment.¹

A possible way to account for this in telematic performance, in which the performers are physically isolated from one another, is based in the intense listening focus required on the part of the musicians.

¹Not unrelated to Ascott's most famous notion of "love in the telematic embrace," in which he argues the potential for deep human connection through technological mediation (Ascott and Shanken, 2007).

When musicians are physically co-located this listening awareness should also be of primary concern, yet it is easy to relax and believe that playing in the same room is the same as playing together, musically. Not so in distributed settings, where failing to listen destroys all sense of connection.

This is additionally compounded on a technical level by the typically close microphone positions and the previously discussed preference for narrow camera angles. Due to the necessities of audio amplification systems, microphones are typically placed close to the sound source, to reduce unwanted noise, echos and feedback. Visually, close-up images of remote performers appear to be much more informationally rich and satisfying. The result is that the remote projection, aurally and visually, of a musician is from a very close position, atypical of face-to-face interactions. This also presents a unique perspective when many musicians are involved, as it is physically impossible to keep one's ear within six inches of more than a couple acoustic instruments, and yet a whole telematic ensemble may be projected so.

Musically this means that very subtle sounds are transmitted at a disproportionately greater degree than compared co-present, acoustic sound. Key-clicks, on woodwind and brass instruments, finger impacts and bow-changes on strings, and the breathing of all musicians are heard more distinctly at the remote end. While the awareness of these sounds may not be at a focal level their presence is palpable and transformative. In turn, these differences may reward and further stimulate the intense listening required of the telematic performer. Of course this situation is not unique to telematics, as recording studios and live events that employ amplification can also take advantage of these sonic artifacts of instrumental performance. However, removing the physical body and instrument to a remote location and restricting aural access to the microphone alone is not a choice in telematics, but a necessity. This, in combination with the ephemerality of live performance, takes the experience beyond that of simply heightened awareness and into a loop of intense, focused interaction that is understood as intimacy after the fact.²

This heightened focus on listening leads to the hypothesis that other cognitive distractions, such as self-criticism and social evaluations, are displaced, resulting in a freer mental state for musical practice and creativity. Although this is not yet proven quantitatively, I strongly believe, based on my observations and discussions, that the telematic space can and does instill a sense of freedom in the participant, leading them ultimately to a higher degree of musical and technical achievement, as mental inhibitions are ignored in the face of artistic engagement.

The notion of intimacy in telematics can be explicitly employed for aesthetic purposes as well. Just as cinema makes effective use of soft filters, close camera angles, and warm audio filters so can live telematic performance, bringing this affective transformation to musical performance.

²This notion of intimacy is closely related to “flow theory” (Csikszentmihalyi, 1991), and can be understood as a specific experience of flow in the context of a close, musical dialogue.

In a bizarre twist of mediated presence, it is possible for the remote participant to seem more present, more intimately portrayed to the audience than the physically present musicians (as previously noted). Due to the large, close-up video images, and focused microphone placement the physically absent musicians appear close, in stark contrast to the actual distances involved in the distributed ensemble! On the other hand, the local musicians can appear diminished, in their merely large-as-life, physical presence on the stage, playing acoustically in the hall. Yet, the same effect is taking place at the other end of the connection as well, with *our* local musicians appearing large and close just as the remote musicians are merely human in their own space.

The musical differences between each physical location in a distributed performance highlight a myriad of issues, and challenge many commonly held assumptions about music. As previously discussed, there is no single right version of a telematic piece. Even if the theoreticians are willing to ignore the small sound differences between the front, back, and sides of a concert hall, in telematics the musical parts are aligned differently (or phased uniquely) at each location. Attempting to record the parts and create the correct alignment after the fact, speaks to an outdated philosophical conceit that misses this fundamental shift that telematic performance brings to music. On one level even the small changes in sound noticeable between the front seats in a concert hall and the back of the balcony effectively change the piece, and yet each is unique and valid in its own way.

The latencies inherent in telematic performance result in dramatic experiential differences between each performance location. By employing carefully crafted rhythmic patterns, and the right delay amounts, it is possible to setup situations where one location hears stylistically different results at the same time. Harmonically it is a simple matter to have a consonant alignment at one stage, and a fully dissonant, chromatic alignment at another. Creating these situations is possible by writing parts that add up to different wholes when offset by a beat in any direction (as caused by the network latency).

The multiplicity of aesthetic results is a unique potential and inspires interesting philosophical problems. Telematic, distributed performance denies searches to locate the ‘real piece,’ or rather, attempts at fixing one location as the ‘authentic’ version ignore the real implications. Each location here *is* the piece, and unlike a concert performance the different readings are undeniably, phenomenologically unique. For better or worse there is no way to produce a single recording of the performance, or of the piece, for later consumption. To appreciate the performance it would be necessary to watch video from each location in sequence, understanding that all are windows looking in to the work of art from different directions, but none ever fully capture the creation.

3.4 Style

The distributed musical environment employed in telematic performance does not inherently state stylistic constraints. However, the challenges and issues discussed thus far clearly coerce the type of music that can be played. I approached this research from the stance of a contemporary art music composer, interested in abstract, atonal sounds and musical expressions, which guided the bulk of the musical experiments. Yet, the participants that volunteered and were selected for the sessions were from a variety of backgrounds and musical interests. Thus a range of musical types were briefly exposed and tried.

Due to the time offsets inherent in distributed performances music that is dependent on ensemble aligned harmonic progressions does not transpose well. While some have accomplished this with a modicum of success (such as Dresser (2008), A.2.1, A.2.2, A.210), this necessitates exceedingly low latency connections (on the order of 50ms or less) or a high degree of tolerance for rhythmic inaccuracy. Thus musical styles that depend on rapid, synchronized chord changes do not fare well in a telematic situation (this includes many types of Jazz, scored classical music, many ethnic musical traditions, and most contemporary pop music).

Similarly, styles that rely on intricate melodic or contrapuntal lines are challenged by the latencies of networked music performance. If the amount of delay fits into a regular portion of the metric pulse (such as a beat or a bar), and the different locations in the setup are willing to tolerate different melodic alignments then this can be overcome. However, this requires a re-conception of the style on a fundamental level.

On the other hand, music that is based on harmonic drones or loose chordal alignment can work very well. Strongly pulsed music that accepts different composite rhythms at each location also fits the setting well. Typically, most of the pieces performed over networked connections in the last few decades have been forced to forego rhythmic cohesion and aim for more continuous, amorphous textures (for example *Telocello Concerto*, A.2.4).

As a performance medium telematics denies most physical, energetic stage interactions. Musicians cannot move around one another, easily change their setup on stage, or approach another performer to emphasize their connection and interaction. Stage acrobatics, such as running through the audience or crowd surfing are clearly prohibited in any comprehensive sense (each performer only being able to physically interact with their co-located audience). While contemporary art concerts rarely venture into this realm, popular music venues thrive on such audience-performer dynamics.

3.5 Conclusion

Telematic musical practice is still in its infancy and as with all new artistic developments a lot of varied experiments and research questions are yet to be asked. The goal of replicating face-to-face practices has lead a lot of researchers towards simple video-conferencing applications and results of dubious value. Yet the possibilities of aesthetic transformation in light of telematic implications remain and are gradually being explored and mapped.

Remote musical collaboration will continue to gain adherents and credence, in both academic communities and beyond. While the technological hurdles at the moment are extreme these requirements will continue to become more accessible. Broadband proliferation and more powerful consumer personal computers will allow the dissemination of telematic practice, facilitating communal and cultural developments in directions that are yet unknown. Artists need to be at the front of technological adoption, here and elsewhere, asking questions to uncover implications that may guide our advance. Both for the sake of art and its impact on our growth as a society and a people.

This thesis has focused on the aesthetic implications of telematic and networking technologies towards the creation of new modes of expression. However, telematic music performance also enables the creation of ensembles of musicians that are otherwise denied due to cultural or logistical difficulties. Bridging physical distances enables musical interactions between individuals that would otherwise have to invest significant resources to travel and meet in person. This is not simply a matter of convenience, as often telematic setups are anything but convenient, but a matter of possibility. In the mid 1980s, the pop icon Paul Simon undertook a transformative musical endeavor, bringing the music of black South Africa to the radios of western listeners, with his album *Graceland*. However the actual production process involved recording the African musicians separately, bringing the tracks to a studio in America, and overdubbing the western musicians. While there is an undeniable meeting of musical aesthetics, the meeting of musicians was fleeting at best. Telematic connections promise that similar multi-cultural bridges can be much stronger today, affording the musicians time to interact and co-create, recording in real-time, joining studios and stages around the world. Eventually it seeks to level the playing field of musical collaboration, discarding the currently elite requirement of access to global transportation.

While the problems of telematic music performance abound, and often dominate its conception, the promises are unique and exciting. Setting up and operating a telematic network connection requires training and access to the technology. Even with a connection, problems due to delays and latencies may seem insurmountable, and the issues surrounding actual performance events are many. Yet once these are engaged and dealt with the affordances speak strongly to the concerns of today. Technological mediation becomes

more and more a part of our lives as we begin to move more and more of our experiences out of physical places and into digitally facilitated spaces (Giddens, 1990). Uniting unique, creative performers in concert, presenting a multiplicity of musical creations uniquely at each venue, speaks directly to our modern, western ideals of individuality in a global community.

Appendix A

Survey and Analysis

The following comprises an examination of a number of telematic musical works and performances that exemplify the state of the art as it is practiced today. The list is not intended to be comprehensive, but captures the breadth of work being undertaken (for example, A.2.9 exemplifies all of the work of Avatar Orchestra Metaverse, and A.2.3, A.2.5-7, and A.2.13 trace the aesthetics of Oliveros' adherents).

For many, the goal of reproducing physically co-present practices in a distributed environment is the primary aim (see A.2.1, A.2.2, A.2.10, A.2.11). In contrast, others retain the notion of performance but look for new aesthetics that are afforded by the network connections (evidenced most strongly by A.2.12). Finally, some strongly embrace telematic ideals, but have a more tenuous relationship with the performance setting (see A.2.8, A.2.9).

The agency of the participants is a useful tool for analysis, but typically appears incidentally in a performance (A.2.6, A.2.7, A.2.11) and is only rarely employed explicitly (A.2.8). The intimacy of the ensemble can be heard in some examples (A.2.3), but also appears to be an artifact of the setup more than an intentional aesthetic creation. Intersections with other computer systems, such as artificially intelligent algorithms or data collecting sensors, while certainly a unique possibility in telematic performance only appears rarely (A.2.8, A.2.12, A.2.13).

A.1 Earliest Examples

The first use of telecommunications as an artistic medium occurred in 1922 when the Hungarian constructivist artist László Moholy-Nagy made the work *Telephone Pictures*. In this work the artist dictated a set of paintings to a factory foreman over the phone, who then had the images printed, all without any direct connection to the hand or person of Moholy-Nagy. Thus the technologically assisted communication channel became a tool employed for artistic creation.

It was not until 1970 that it became feasible to connect performing artists in real-time. At this point two projects employing earth orbiting communications satellites connected artists on the east and west coast of

the United States. NASA facilitated this event, in which artists produced composite images of participants, enabling an interactive dance concert amongst geographically disparate performers. An estimated 25,000 audience members witnessed these improvised, interactive dance and music performances that were displayed in real time (Ascott and Shanken, 2007).

A.2 Pieces

A.2.1 Telematic Concert, 2007

Available at: <http://www.youtube.com/watch?v=wLJcDoJQ2ck>

This performance, instigated by Mark Dresser, brought seven musicians at three institutions together in a full length concert of telematic music. Ostensibly in a Jazz-derived improvisational context, the performance comprised a series of compositions that explored a number of different musical textures and affects. Dresser wrote about his experience producing this concert in Dresser (2008). The technology for this performance was JackTrip for audio streaming and Skype or iChatAV for the video representation (with connections over Internet2).

Clearly, this performance takes conventional concert hall practice as the start and end objective. The remote performers are displayed on large screen above the heads of the musicians on stage, who face out towards the audience. All of the remote musicians are visible, with the exception of the laptop player, who only shows up to accept applause at the end of pieces. With that notable caveat, the agency and presence of all the other musicians is quite strong and expressive. Only one audience was assembled for this performance and thus the main stage heard the ‘definitive’ version of the pieces and issues of alignment and multiplicity were not raised.

The latency for audio in this concert was probably around 60 ms each way (based on knowledge of the software and hardware employed and careful audition of rhythmically synchronous parts of the performance). It was thus possible for the ensemble to play a metric piece together (“Parawaltz”, around 23’ into the video), although the performance is still very loose and ensemble attacks are relatively diffuse.

A.2.2 Telematic Concert, 2011

Available at: <http://www.youtube.com/watch?v=qVW93o1huA>

This performance is effectively an updated version of Dresser’s 2007 concert, with the same goals. In this case the video technology is significantly advanced (and required a team of several engineers to setup and operate for the performance) and the staging is very convincing. Now the remote musician (only three

performers appear in the concert) is projected on the back of the stage and the lighting is set to give the appearance of co-location. Despite the distributed aspect of the performance the surface appearance is of a face-to-face concert. Again, only one audience was assembled, and in this case the physical distance was so short as to render the network latency almost imperceptible (Jack Trip was again employed for audio streaming).

A.2.3 Free Telematic Trio Improvisation

Available at: http://www.mitpressjournals.org/doi/suppl/10.1162/lmj.2009.19.95/suppl_file/

LMJ19-TelematicMusic-esupp4.mp3

This piece was played by Chris Chafe at Stanford University and Jonas Braasch and Doug Van Nort at Rensselaer Polytechnic Institute on February 10, 2009. Video streaming was not employed for this work and Jack Trip was used for the audio. The latency was probably around 80ms in each direction, but the exact number is unknown.

In the recording available at the link above, the initial florid outbursts from the soprano saxophone and celletto (a custom built electric cello with extend range) give way to rough textures in the celletto and brief outbursts from the sax. Both instruments change textures together shortly thereafter, demonstrating the responsiveness and communication enabled by the network connection. A pizzicato passage from the cello precedes the apparent entrance of the laptop, who joins with sustained tones which are shortly taken up by the whole trio. The timbre gradually becomes brighter, and a glissando with an echo effect indicates that the laptop has begun processing the sound of the other instruments. Gradually, the acoustic instruments are subsumed by the sonorities of the laptop at the end of the audio excerpt.

This recording shows the dynamics of the trio and the interaction and responsiveness of the ensemble. My reading is that Chafe exhibits the dominant musical personality in this case, with all of the players eventually following him despite attempts at contrasting material (especially on Braasch's part). This may be due simply to personality, or some particulars of the setup that was employed in this case.

The agency of the three performers presents an interesting contrast. On the one hand we have Chafe and Braasch, playing instruments that are sonically separable, and playing sounds that are perceptually connected to the physical operation of their instruments. Van Nort on the other hand is using a laptop, the workings of which are completely opaque to the listener (other than the moments of effects processing of the other instruments). Thus the agency of the first two is very expressive while Van Nort is operating in the magical or potentially secretive realm.

Finally, it is worth noting that the two physical instruments are closely mic'd, conveying a very intimate

sound.

A.2.4 Telocello Concerto

Audio excerpt available at: http://www.mitpressjournals.org/doi/suppl/10.1162/lmj.2009.19.95/suppl_file/LMJ19-TelematicMusic-esupp5.mp3

This piece features Chris Chafe on celletto and Sara Weaver, soundpainting (a set of physical conducting type gestures that are used to give directions for improvisation to musicians). The performance featured 36 other musicians including the SoundWIRE ensemble at Stanford University, Tintinnabulate at Rensselaer Polytechnic Institute, and VistaMuse at University of California San Diego. Recorded on, November 16, 2007, although the published recording excerpt features a new solo recorded over the original.

The piece itself consists of a dense orchestral aliatoric texture that gradually builds in energy to the mid-point where a gestalt rhythmic alternation is affected. The chaotic state thins before building again, revealing a lone, steady guitar ostinato. The second build features more rhythmic playing generally and concludes with a low drum filling in the bass of the chaos.

We have to wonder why a new solo was recorded over the original before it was released to the public. Clearly it was not deemed satisfactory, but why? It may have simply not been up to the musical standards of the performers, but given the extreme technical setting it may have been due to other problems related to coordinating 36 musicians over network connections. This would not have been surprising, but the accompanying write-up makes no mention of this fact in either case, extolling the success of the event (Oliveros et al., 2009).

A.2.5 SLM

Video available at: http://www.mitpressjournals.org/doi/suppl/10.1162/lmj.2009.19.95/suppl_file/LMJ19-TelematicMusic-esupp3.mov

This piece was staged between New York and Jerusalem and is scored for an ensemble of 10 (or so) bass players. The work is largely improvisatory and features “sound-painting” conducting, a formalized set of gestures that loosely guide group improvisation. Sarah Weaver and Mark Dresser conduct the ensembles.

The audio latency, using JackTrip, was likely between a tenth and a fifth of a second each way, while the video delay is a half second or more (using Skype or iChatAV).

Although billed as a concert, one of the venues appears to be a lab or class room, with a large ‘M’ poster on the wall and the musicians dressed less formally. Thus, there appears to be a discrepancy of approach immediately, between the two locations. The audience was located at the non-‘M’ space, where

the performance ostensibly “took place,” in this case.

This work is based on the idea of antiphonal ensembles, as promoted by Oliveros et al. (2009), yet the musical textures are thick and opaque, making perception of the call and response aspects very difficult. The latency is effectively ignored in this performance, as rhythmic material is avoided and ensemble unison is never attempted. This then suggests that the work may have been more effectively staged in an acoustic antiphonal setting, with each ensemble on opposite sides (or front and back) of a concert hall.

A.2.6 Droniphonia (2010)

Available at: <http://www.youtube.com/watch?v=eqGo7qRaDZ0>

This work was staged at the Sonic Arts Research Centre (SARC) at Queen’s University, Belfast, Northern Ireland, and with musicians connecting from Alberta, CA, and New York in 2010. From the authors: “Droniphonia has polytonal drones continually morphing timbres, volumes and fundamentals moving in space. Networked musicians (Belfast, Banff and Troy NY) listen to the drones and develop gradually overlapping improvised sounds and phrases—first solo and then between two or three players at a time in a slowly growing density and texture. Different musical characters appear unexpectedly from time to time out of no where on a one time basis. Finally all players weave their way in and out of the droneophony sharing the unknown finale.”

Droniphonia features drones throughout (mostly electronic), as suggested by the title, shifting over long time scales (several minutes), with various chaotic elements woven in, especially from the piano, accordion, violin playing short outbursts and pizzicati that slowly vary in density. Sudden shifts in the drones (such as at 6’20”) appear to be arbitrary, as do the other shorter outbursts. Around 4’ the violin repeats a single motive many times, and a similar idea is hinted at by the accordion, in what may be the most approachable moment of melodic dialog. The highest point of density is achieved at roughly six minutes. At 7’30” the piano is finally played with the keys (leaving us with the question “what does it signify?”). The clip ends short of the end of the piece, although it is probable that it continues in the same vein, fading out to the end.

Overall, this is a musically functional concept to accommodate the latency of networked music. The continuity of sound removes the need for identifiable sound-source connections, although the on-stage musicians (especially the pianist, and violinist to a lesser degree) are very focal and enabling of the performance (i.e. of the audience’s connection to the music). Chafe and Braasch’s invisibility is not really an issue since their sounds are obscured by the density and complexity of the music. Shown on the projection, Van Nort’s huge size is interesting and his DJ-esque movements with the head phones and tablet stylus provide a source and

reason for the abstract (studio-audio art) sounds and disconnection of the music.

Several elements remain unexplained, such as the significance of the co-located but offstage musicians as well as the apparent absence (i.e. lack of presence) of the remote musicians. This was probably a result of purely functional and technical issues, but as an artistic expression this is not satisfactory. Since the absence of the remote musicians is not highlighted, and their *magical* appearances are effectively obscured, we can argue that this work would have been as effectively staged with a fully co-present ensemble. The work also skirts all the other promises of the medium (latency, the creation of unique pieces), placing the performance clearly in the hall at SARC.

A.2.7 100 Meeting Places (2007)

Available at: <http://www.youtube.com/watch?v=46v4MWsMUxw>

Staged on March 27, 2007, “This was a 4-way real time concert directed by Pauline Oliveros at RPI in new york, Chris Chafe at CCRMA, Sarah Weaver at Loyola, Chicago, and Synthia Payne at UC Santa Cruz in partial completion of her MFA. Audio mix via JackTrip software developed by SoundWIRE Group at Stanford’s center for computer research in music and acoustics. Visual here is iChat with the audio disabled.”

Again, this work utilizes “sound-painting” to coordinate and direct the musicians. The effects can be clearly heard at 0’32”, where Weaver makes a series of strong gesture that result in an analogous, and coordinated ensemble change. However, at around 1’08” another set of similar gestures produce no audible result (while the sounds do change they do not seem to coordinate with the sound-painting or with one another).

This performance presents a very interesting juxtaposition of spaces: a concert hall (in NY, on the right side), Weaver in a class room (at Loyola, Chicago), Chafe (at CCRMA, on the left side) in a conference room (with a very shiny table), and a lab, office, or bed room at UC Santa Cruz (at lower center). The contrast between spaces traditionally used for performance, and where the music is given great reverence, with spaces that are used much more generically is striking. This is further shown by Weaver wearing all black while the rest are wearing solid colors and khaki-type pants, indicating various measures of formality.

However, this is one of the first works that is not “in” one place alone. Since there is no prominent audience, each location has an equally valid claim on the identity of the work, creating effectively four different versions of the piece. This may also be considered an essentially participatory event, since it was produced primarily for the benefit of the players. This critique is based on the presentational and performance aspects of the work, however if this is considered primarily a participatory event than much of

this would need to be reformulated.

In either case, this piece puts most of the musicians in a disconnected or *suspenseful* condition. Although they are all basically visible, they are all so small (and were in concert too, based on the limitations of iChatAV) that the details of their actions are hidden. The conference room view presents another challenge as it visually cuts off the lower halves of the musicians and their instruments, further obscuring their movements. The final result is a disconnection of sounds that makes expressive access very difficult, given the highly mediated nature of the ensemble.

A.2.8 SLOrk—Electrode (2010)

Available at: <http://www.youtube.com/watch?v=BBe2Hb04Ehw> and
<http://www.youtube.com/watch?v=j2KodWLnVr4>

From the authors: “Stanford Laptop Orchestra (SLOrk) with special guests: Synthia Payne, voice, Zach Herchen, saxophone, Sean-David Cunningham, violin, Chris Chafe, conductor. Electrode combines music from outside the hall and inside the brain. Soloists piped in from Colorado and New Jersey are performing with the laptop orchestra, which itself is making music derived from about 150” of brain activity. Brain recording data and the idea of ‘musifying’ it are thanks to neurologist Dr. Josef Parvizi of the Stanford University School of Medicine. . . . Performed as part of the Stanford Laptop Orchestra Spring Concert in the Dinkelspiel Auditorium, Stanford University. June 2, 2010.”

The music and the staging of this work present a severe schism of aesthetic. The performers are arranged on stage like a percussion battery, each kneeling at a black laptop, spread out evenly around the stage and facing out towards the audience. Chafe directs from a chair at center stage, strongly down lit, periodically waving a cello bow, dramatically. There is a large projection screen behind the stage that remains dark for the duration, and the remote musicians are never seen in any fashion.

The music itself is basic digital synthesis, reminiscent of 1960’s sci-fi sound scores. The actions of the performers are almost entirely invisible, either hidden behind laptops or in their heads (in the form of brain activity). The remote musicians are completely undetectable. Although the saxophone and voice become audible at around 1’50” in the second video, they could easily be pre-recorded samples. In fact, with fifteen laptops on stage this is the easiest explanation. In either case, their presence or absence appears to be incidental and provides no expressive or performative content.

The primary critique of this work is not with its telematic nature but with its relationship to the implications of performance. In a strange twist of agency, the performers are simultaneously *suspenseful* in that we see them on stage but have no access to what they are doing, and *magical* because we hear sounds

but cannot relate them to any sources. In the end the observer has nothing to grasp. Creating music from brain waves is certainly a valid premise, but inherently denies a performance. Involving remote musicians and then removing all access to their contribution only aggravates the problems.

A.2.9 Avatar Orchestra Metaverse—Rotating Brains, Beating Heart (2010)

Available at: <http://vimeo.com/15426324>

The Avatar Orchestra Metaverse is an ongoing project that seeks to exploit the virtual environment of Second Life for artistic and musical performance. *Rotating Brains, Beating Heart* was a work created for the Digital Resources for the Humanities and Arts Conference 2010 at Brunel University, UK. This event involved two performers on stage and 18 people in Second Life.

The Second Life participants each are equipped with a sampling synthesizer of sorts, allowing them to trigger the playback of sound files at any point. On stage a saxophonist and a laptop musician play along with the virtual world, although it's not clear if the virtual participants were able to hear the concert performers. Also worthy of note, several of the Second Life performers are setup on stage with their computers behind the musicians.

The primary aesthetic challenge for this, and many similar pieces, is the observational access to the agency of the performers. Without any expression possible there is no performance, and this piece confronts this at every turn. It is impossible to know which sounds originate from Second Life, which from the laptop performer, and often even the saxophonist's contribution. A moment of laptop alone, or Second Life alone would have allowed the observer to form an idea as to the operation and sonic identities of the participants. Since everyone effectively plays on top of one another all the time there is no way to tie effects to actions (with a few momentary exceptions).

As it stands, the saxophonist provides the strongest, indexical sound source, becoming the center of performer expression for the piece. For the electronic sounds there is no way to differentiate the contributions of any participants, and the Second Life performers are often completely absent from the process. We believe that they are triggering sounds, but since there is no apparent visual to aural connection for all practical purposes the video could have been pre-recorded, as the audience has no access to its liveness or connection with its ephemerality or immediacy.

A.2.10 Network Music Performance

Available at: <http://www.youtube.com/watch?v=Jjpi0jURIGA>

This short demo ostensibly shows a jazz trio using a network setup to perform a piece. This work vocally

promotes the co-present gold standard, using video-conferencing models to reproduce face-to-face practices. Despite the apparent advances of the technology the performance in the video is rhythmically uneven, which is surprising considering that the two “rhythm” players were co-located. The actual latency involved is unknown, as is the manner in which the system attempts to synchronize the performers. Approaches of intentional offsets would not work, given the strongly rhythmic and harmonic nature of their chosen music.

All in all, this performance reproduces the results of Dresser, 2007, above.

A.2.11 Disparate Bodies 2.0 (2007)

Available at: <http://www.youtube.com/watch?v=k7qqCPh0HEw&feature=related>

A performance staged between three locations, Belfast, Hamburg, and Graz, which was additionally broadcast into Second Life (to enable remote audience access). Music by John Cage, Pedro Rebelo and others was played.

This work shares the “piece” in an admirable way, providing audiences at all locations (and in virtual space), each with a unique figuration of the music. However, the music chosen makes comprehending the differences (between each location) difficult.

The presence of the remote musicians, which we see in the video, is often confused. The only instruments that make a strong appearance are the ones that are sonically contrasting to the local musicians (such as the trombone) and which are clearly not present on stage. However the interactions between local and remote players is difficult to grasp and appreciate. In fact three pianists, two saxophonists, one clarinetist, a flutist, and a trombonist were involved in this performance.

A.2.12 Auksalaq (2010)

Available at: <http://auksalaq.org/>

This work is a telematic opera of sorts, premiered in NY on October 31, 2010.

One of the more ambitious telematic performances to date, foregoes the typical aliatoric, improvisatory textures for composed ensemble parts. The piece itself is comprised of a number of short movements for small ensembles that can be played in any order, overlapping as the performers choose.

Due to the divided nature of the composition, latency does not play a role in defining the performance, approaching the issues with an antiphonal mindset. Video streams are employed to show the remote musicians at certain times, while at others the projectors display data visualizations, photos and pre-recorded video material. However, the presence of the remote players is largely preserved, due to the focused nature of the music and the controlled aspect of the video changes.

Although the performance is clearly staged in NY, making the remote participants auxiliary to the main event, the extra musical aspects (of video and real time data systems) place this work solidly in the telematic domain. While the physical distance of the remote musicians is largely incidental to the musical result, their appearance as images amongst the digital projections arguably serves to bridge the human and electronic components of the narrative.

A.2.13 Distributed Composition 1 (2011)

Soon to be available through the NIME 2011 web-archive.

Performed at the international New Interfaces for Musical Expression (NIME) conference, 2011 on June 1st in Oslo, Norway. Doug Van Nort played laptop on the main stage, Jonas Braasch connected from Troy, NY, and Pauline Oliveros from Stanford, CA. The piece lasted for approximately twenty minutes. The two remote musicians were projected on screens (roughly six by eight feet) set side-by-side on the stage next to Van Nort. Jack Trip was used for the audio (with a latency on the order of 150 ms) and Skype for the video streams.

The premise of the work relied heavily on algorithmic processes running on the laptop which directed the improvisatory textures played by the musicians. This effectively skirted the issues of latency, as synchronous textures or gestures were never employed. The presence of all three musicians was strong, analogous to the Free Telematic Trio Improvisation, above, with the same caveats about the laptop's involvement.

Despite the title, this work removed the distributed nature from aesthetic relevance, from the audience's perspective. The musicians could have been on the same stage, or in any number of locations, and the end aesthetic result would have been effectively equivalent. Given the automated nature of the interaction, based on computer algorithms rather than musician intuitions, the mediation of the network became largely insignificant.

A.2.14 The Loop (2011)

Soon to be available through the NIME 2011 web-archive.

This short performance, presented on May 31 at NIME 2011, is only notable in its complete inability to incorporate any aspect of telematic aesthetics. The work was ostensibly for four laptop performers, three on stage in Oslo, Norway, and one connecting remotely from Norwich, UK. The audience was told that they were hearing the fourth musician, although no video stream was shown.

However, due to the completely opaque nature of the laptop as an instrument I, as an audience member, had no access to impact of the remote musician. For that matter, the whole work could have been a fixed-

media piece with minimally staged performance-art choreography, and it would have appeared identically. At the very least, promoting the participation of a remote musician who is completely absent during the performance frustrates expectations.

Appendix B

Research Experiment Report

The following are a series of reports on experimental research sessions staged between locations on the University of Illinois at Urbana-Champaign campus and various remote locations. The initial objective of these sessions was to explore and identify the challenges and affordances of available network software and setups for musical ensemble playing. Over time this evolved by changing hardware and software, developing new software, and trying new musical problems to continue the exploration.

The first tests, starting in October, 2010, employed a new 10 gigabit network connection within the National Center for Supercomputing Applications (NCSA) and attempted to stream HD audio and video between two rooms on different floors of the building, aiming for the lowest possible latency. Eventually, after writing new video compression software for the Max 5 programming environment, I found that lossless compressed HD video streamed just as fast and more reliably than fully uncompressed video. I also switched audio software, arriving at Jack Trip (from CCRMA at Stanford, CA) as the best, most reliable and flexible option.

B.1 Methodology

To distinguish art *research* from conventional *practice*, Zimmerman et al. (2007) state the following four lenses through which the work should be evaluated.

Process: the researcher must document the process employed with sufficient detail that it may be reproduced. Unlike the sciences the goal is not to ensure a reproduction of results, but to ensure a rigor, validity, and relevance of method. No two artists presented with the same problem will solve them identically, but the artifacts produced may still be equally significant to the field. Also, this is the distinction between a practitioner and a researcher. The latter is obliged to document the path that led to the final product, providing evidence of originality, creative skills, significance of contribution, and signaling a high degree of conceptual rigor.

The following report is intended as a significant contribution to the documentation of this process in the

context of artistic research.

Invention: the work must be highly innovative, demonstrating a novel integration of relevant elements to address a specific situation. To show this innovation, the work must be strongly grounded in the body of academic work, detailing the advancements from the current state of the art and technology. The goal of the artifacts produced is not to satisfy any market or economic conditions through slowly iterative improvements, but to push the state of the art in a significant jump of innovation.

Relevance: the researcher must communicate the intentionality, motivation, current situation, and preferred state the work attempts to achieve. The argument for the preferred state must consider why the larger community should desire this change in order to critically justify the work. The relevance of the work is significant because the measure of the success of the research is not in the replicability of the findings, but the strength of the pertinence and impact of those findings.

Extensibility: as with all academic research, the art researcher's work must be made available to other researchers to build upon, leveraging the knowledge resulting from the products of this research for further academic work. This requires a description and documentation in a usable form for the larger community.

B.2 Research Log

Each entry in the following log states first the date and time, locations, and participants. A brief abstract is followed by the initial research questions for the session, and technical notes. The rest of each entry is comprised of observations and suppositions recorded immediately after the session. The language is primarily in the first person, as this report is the author's record of the events that transpired.

B.2.1 October 13, 2010. Wednesday. 3 p.m.

Locations: NCSA 2103, NCSA 1005

Participants: Participant A (flute), Ben Smith (Violin)

Abstract: Conducted the first test of the 10 gig audio/video connection in NCSA between rooms 2103 and 1005. Participant A Carrizo played flute in 2103, I played violin in 1005. The primary objective was to see if two improvisers could communicate through this medium, and at what level problems and challenges would appear. The result was very affirmative, with issues surrounding synchrony and unison playing coming to the fore.

Research Questions: Can two musicians play over a video-conferencing type connection, and if so what sort of musical interactions are afforded? How does it feel to play in this setting, and what technical

details strongly influence this?

Technical Notes: in NCSA 2103: 8 core mac tower, blackmagic Intensity HDMI video capture card, Myrinet 10gig network card, canon mini HD camera, crown 700 mic, mackie mixer, cabled to the mac's builtin audio jacks. Audio playback through 2 JBL PA speakers. The image was projected on the big screen on the east wall, the camera (and computer) were at the south, and Participant A stood in the middle of the room, lit by 3 LED cans on the ceiling. In NCSA 1005: new mac tower w/ 29" display, same cards, mic, rack-mount mixer, same cabling, camera, internal mac audio. Audio played through 1 small Genelec speaker. Incandescent lights in the ceiling.

Both computers were sending video at 1280x720 (the network delayed anything larger) at ~20 fps. Audio was stereo, uncompressed, 16/44khz. Both audio and video were sent through Max 5 using the `jit.net.send/receive` objects, and `catch/release` for audio.

Ben saw Participant A from the side as she faced the projection but Participant A saw Ben from the front.

Report: We spoke first in person, toured both spaces, then split up. We first gauged the audio/video synchronization delay by clapping for one another. Then we tried to tune for about 3 minutes, then played a 10 minute improvisation. This was followed by a discussion and a few musical "tests."

Delay: Audio delay appeared to be consistent at about 10ms each way, maybe faster. Video delay from 2103 to 1005 was around 30 ms (roughly one frame of video) while the reverse, 1005 to 2103 started very high, ~500 ms. This was due to trying to push video at too high a frame rate, and as we progressed the delay diminished to roughly 50 ms. The cameras also introduce about a 30ms delay so in the end we both saw the other ~100 ms later.

During the improvisation we incidentally explored call & response textures, sustain versus active (or melody and accompaniment), rhythmic synchrony, and gradual decelerando. Since we had never played together prior to this event it was very encouraging to observe our ability to wordlessly find these modalities through the technology.

Observations from Participant A: She stared at the video intently, but mostly for the sake of developing a sense of connection, not for musical communication. During the rhythmically synchronous sections this became a focal point as she remembered that there was a significant video delay and her instinct to watch the violinists movements would not help. The video delay did not become a serious cognitive issue for her.

Tuning was very difficult, and maintaining a sense of blending or matching was hard. It was hard to hear the violin at times, when she was playing. The resonance of being co-present was removed here and thus playing in tune together was a significant cognitive element.

She was very excited post improvisation, and impressed with the ease of interaction. She felt as if I was very present in her space, and as if she was not alone, although physically she was alone in the room.

Observations from Ben: I hardly looked at the video at all, only watching during moments when I wanted to coordinate attacks or releases. This was relatively unsuccessful (due to the video delay), and I otherwise spent all my energy listening. The audio felt very present and responsive, as if we were standing next to each other playing. This was very encouraging and enabled the improvisation greatly.

Tuning was very hard, and there were some audio artifacts on certain pitches of the flute. This slight distortion made it even harder to match pitch, but otherwise our musical communication felt very natural and fluid.

About two thirds of the way through the improv I started to hear rhythmic hints in our playing and intentionally brought it into a regular eighth-note pattern (I played a 5 eighth-note ascending 4ths and whole steps pattern as an ostinato for a minute or so). Participant A picked up on this immediately and we played very closely together, as close as I would expect most co-present musicians to achieve (our attacks were barely audibly off, as in an orchestra, not as tight as a string quartet can achieve). We then slowed it down gradually over a minute, remaining in synchrony. This was very impressive, given the expectations for the medium.

We ended the session with 2 tests, one where I visually queued a series (~10) long notes and we attempted to attack them audibly together, and then the reverse, where Participant A visibly queued the notes. As the follower we both independently realized that the only way to line the notes up was to intuit/learn the delay and anticipate the video, synchronizing the audio. We were able to get a couple out of each series “together” (probably still 10-20 ms off).

B.2.2 October 14, 2010, Thursday. 11:18 am

Locations: NCSA 2103, NCSA 1005

Participants: Participant B (alto sax), Ben Smith (violin)

Abstract: This session, with Participant B, alto sax, and Ben at NCSA between 2103 and 1005, examined the same basic questions as before. Again, the participants felt very connected and present, and tuning was not the same issue as previously. Also, intentionally feeding the audio back through the system presented a strong setup for rhythmically synchronous playing.

Research Questions: See previous session, now with different participants. Also: can rhythmically synchronous music be played in this environment?

Report: This time, inexplicably, there was a significant audio delay, on the order of 160ms round-trip.

As we began we could hear lots of clicks and dropped packets so I set a 40ms buffer on the audio receiver on each side. It turned out that part of the problem was a bad cable on the 2103 side.

We also had a ~ 1.5 second video delay from 1005 to 2103 but a ~ 40 ms delay the other way. I had started the session with both sides pushing over 200 MB/sec of data, which previous tests had indicated was too much. This caused significant buffering in the network and dropped frames of data. Before we actually started playing I set both to ~ 100 MB/s of data, but the delay from downstairs to up remained huge.

The result was a seemingly perfect synchrony of audio & video arriving in 1005, but a hugely delayed video in 2103.

We played for ~ 6 minutes, talked, and followed it with a few experiments. We tried several rhythmic experiments, setting a beat on one side then having the other follow. Due to the noticeable delay this was uncomfortable for the ‘leader,’ who was very tempted to join with what they heard. We played with this, switching the leader ad hoc and each trying to change so that what they heard was ‘lined up.’ This broke down after a minute or so and we gave up. We didn’t bother trying visual cueing since we knew the video was terribly off.

Comments from Participant B: He did not watch the video other than to confirm my presence at the start. The audio was very clear and felt present, in his opinion, and that was all he needed (I think the video gave him confidence and then the audio was all he really needed to play). The improv was very fluid and responsive.

During the rhythmic tests Participant B noticed that he could hear himself echoing back slightly. I then put my mic closer to the speaker and he played a short rhythmic groove based solo at the period of the delay. I joined in, playing in synchrony with what I heard. He felt this was very satisfying, as I then came back to him in sync with his delayed sound.

Ben’s comments: The audio feedback, in order to assess the round-trip delay was very interesting. I want to build that in to the patch so it can be intentionally mixed in.

The video delay was unanticipated and I don’t know why it persisted. I set the video to 1280x720 on each side and it was still slow going the one way. I will have to test it further to try and figure it out. It turned out not to be a problem as we did not try to do any visual/aural synchronizing. Oddly enough, it didn’t bother either of us, perhaps because we couldn’t see the other’s face well enough to know if their mouth and voice were arriving together.

Compared to the previous day, tuning was not an issue. We seemed to play right together, but now the rhythmic challenges abounded.

B.2.3 October 27, 2010. Wednesday. 1:30 p.m.

Location: NCSA 2103

Participants: Participant A (flute), Ben Smith (violin)

Abstract: This session failed, due to technical difficulties.

Technical notes: Something happened with the tcp settings on one of the macs and I was unable to make any audio/video connections! It was devastating. By the beginning of November I discovered that the tcp send/receive buffers were set impossibly small and correcting this (setting them to 1MB each) made it possible.

To make some use of our time we played a co-present duet in 2103 and I setup the video animation software for a new piece of mine, *Sharpen*, for Participant A to play with. We briefly talked about multi-media art. It was interesting to note that we looked at each other maybe even less (while playing) than when we were telematically displaced. Our playing was satisfactory, and very hard to compare to telematic situations.

B.2.4 November 11, 2010. 1-2:30 p.m.

Locations: NCSA 2103, NCSA 1005

Participants: Participant A (flute), Ben Smith (violin)

Abstract: This session explored different delay amounts, attempting to discover the ramifications. The conclusion is that the round-trip time is important, but the symmetry of the delays (i.e. where the delay is) has no effect. Also, this session exposed the contrast of musical alignment at each location (i.e. when one hears a unison melody the other hears an offset canon).

Research Questions: How do different delay amounts effect what music can be played? What happens if the delay is not symmetrical (i.e. transmission in one direction is longer than the other)?

Technical notes: Setup was same as previously.

Report: We ran the following settings: 2 sec & 0, 200 ms & 300 ms, 3 sec & 100 ms, 50 ms & 40 ms, 100 ms & 200 ms. We ended playing a simple melody in unison with 5 sec & 0 and then 0 & 5 sec. During the later (but not the 5 sec tests) cases we set each end to relay the received audio back to the source, producing a loop (decaying 8 db over the course).

From Ben: Delays over 100ms caused unexpected and “uncomfortable” pauses in the improvisation. Pauses that were a result of “miscommunication,” when each waits for the other and neither anticipates the break. Such pauses don’t occur between practiced musicians—all pauses are intentional and shared. These are the type where each expects the other to continue, but the delay causes breaks.

The a-synchrony with the video seemed to have no discernible impact. We are both playing instruments where it is quite hard to detect the audio/video synchrony in the first place, so a fluid image was enough to instill trust. Further experiments could specifically target this—and might be noticeable with a percussionist or other dramatically visual means of sound production.

The final test, playing *Twinkle Twinkle* at a rhythmic unison at one end, resulting in a 5 second canon at the other was interesting. Mostly because for one player it was banal and for the other it was bizarre. Playing at a rhythmic unison was no challenge, but the canon was intense, demanding extreme focus from the leading player and producing a rich texture. This perceptive dis-synchrony is conceptually challenging, resulting in disparate levels of satisfaction with the musical experience. (Participant A’s comment was that it may be like the comparison of Flautist and Tubist in an orchestra. The one plays the melody, supported by an orchestra of harmony, while the other is the underpinning, unable to hear anything else clearly and having to play early to the conductor.)

From Participant A: Hearing her own sound echoing in the background (acoustically between speaker and mic in the remote room, delayed by the artificial delay amount) was initially very distracting. She was eventually able to dissociate her playing from the echo and accept it as a background sound, becoming more comfortable with the “environment.”

Once the audio was looping back she observed that we tended to play very pan-diatonically. This was presumed to be a result of the long decay of any input to the system (at least 10-20 seconds), encouraging us to play open intervals (5ths, 4ths) and repeat consonant intervals.

Future tests: Construct specific musical examples to test textural combinations with different delays. For example, unison, complex rhythmic and/or melodic patterns, intentional canons, ostinato patterns, matching metric patterns, matching quick pulse with different meters. . . and changing delay amounts programmatically.

B.2.5 November 16, 2010. 10:00-10:40 a.m.

Locations: NCSA 2103, NCSA 1005

Participants: Participant B (alto sax), Ben Smith (5 string violin).

Abstract: This session accidentally exposed issues surrounding the use or removal of video, in addition to varying the amounts of delay and feedback in the system (now exposed in software, rather than by moving mics and speakers as previously).

Research Questions: Same as previous session, plus: how does feeding back a small amount of audio to loop through the network encourage or discourage musical expressions? How does removing the video connection impact the musician’s ability to coordinate ensemble playing?

Technical Notes: as before, each room has a crown 700 mic, mac tower with 10 gig connection, HD camera, speakers. 2103 has a life size projection and PA speakers, 1005 has a 30” monitor and small Genelec speakers.

Changes: in this instance the video capture was not working on the mac in 2103 so 1005 received no image. However for all appearances 2103 looked fine, displaying video from 1005.

JackTrip was used for the audio relay, rather than the TCP based Max5 code previously employed. This apparently removed random distortions and packet loss, but may have introduced slightly more latency.

Report:

Lack of video: Although accidental it proved very significant. Previously we had surmised that the video, and especially the video synchrony, was not important, other than to give a sense of “presence.” It seems that sense of presence is very significant. I, in 1005, saw no video, operating by audio only. This made conversing more difficult (i.e. synchronizing our turn taking patterns), where we actually started speaking on top of each other (was never a problem with bi-directional video). Starting musical pieces was even more difficult since I couldn’t see when Participant B was moving to play. Similarly stopping a piece was very hard because I could not see his postural queues. I can easily extrapolate this to say that while musicians may not focus on the specific movements of other players they are very highly attuned to postural communication, seeing intention such as retreating/advancing, strength, engagement, etc. Without the video our interactions were much more difficult than I had anticipated.

Audio Delay: We played one 5 minute improvisation with a 1000 ms/0 ms setup, no feedback. Our primary observation was that we felt musically inept, unable to employ learned patterns of response and unable to detect responses in the other player. We attempted to impose a rhythmic/metric feel on the second half of the improv, but couldn’t find any multiples of 1 second, necessitating continual adjustment on both parts to try and align. At various points we each would try to remain steady, keeping an arbitrary groove going, but it would be impossible for the other to match (i.e. it would sound correct at one end but the other end would hear the parts misaligned, rhythmically).

We next played a longer improv with feedback set at -8 db at each end, resulting in 2-3 audible repeats of each player. The result was a 2 part piece in which we first played strongly with the period of the delay (now set to 500ms/800ms) and then attempted to get away from the period, resulting in a sound scape, free section. Our perception of the first half was that we were forced, or strongly coerced into playing at the period of the delays (effectively 1200ms, although we each heard the combination of sounds differently). We played short rhythmic patterns and then played a harmonization or other variation with the return on our own sound. We maintained a 1/2 bar (600ms) antiphonal texture for several minutes. Getting away from

the period of the delay was very difficult, although we tried again and again.

The third test involved me manually changing the delay amounts periodically throughout the improv, starting at 0/100, going to 100/200, then 2000/200, and finally 2000/2000. We guessed that the system was imposing another 40ms or so on each of these. The shortest delays (100+40) felt more like an environmental effect and did not lead us to rhythmic playing. The next step up (100-200+40) felt like a fast eighth-note reflection after playing and encouraged some quick rhythmic feels, although we did not have the same sense of enforcement, it was easy to break away and play against the delay here. The longer periods got back to the original feeling of musical ineptitude, as we were now unable to play rhythmically together and also with the return of our own audio. When we disregarded our own audio returning, and did not care about rhythmic alignment, it felt fine, but incidental. Every attempt at lining up resulted in a “bad” sound, or cacophony.

Finally, I attempted to dial in a reverb type effect setting the additional delays to 0 (just using the unavoidable system delay), and turning up the feedback amounts to the point of explosion. This was fun, as the engineer, slowly turning up the dials and listening intently, it was very engaging and requiring of human skill (although this could potentially be done algorithmically). The result was very surprising. As I turned up the feedback amounts we began to hear ringing tones in the audio, as Participant B played across the registers of the sax. At first it was strongly middle C plus several higher tones, and later settled on the D above middle C. At the end I got the D to sustain at a mezzo-forte dynamic, making a strong drone/pad tone. This is perfectly explicable as a waveguide function, but very interesting and not anticipated.

It would be very interesting to setup a self-adjusting multi-band filter that attempts to turn up the feedback and squelch any tones that ring too strongly.

B.2.6 December 7th, 2010 and December 14th, 2010

Locations: NCSA 2103, NCSA 1005

Participants: Participant C (keyboard), Ben Smith (violin)

Abstract: Introducing a new person required a quick survey of all previous questions, such as the viability of the setup and the sense of presence and connection. This was again affirmed, the participants commenting that it was even better than most co-present experiences. Using foot pedals to vary the delay in real-time proved insignificant in this case, although playing rhythmically together at the period of the round-trip delay was very successful and strong, aesthetically. These sessions achieved a strong metric alignment for the first time.

Research Questions: Same as previous, especially focusing on the impact of variable delay changes

during a piece. Also, how effectively can musicians lock into a rhythmic alignment given a fixed delay amount?

Technical notes: same setup as previous session, Participant C was playing a mini 3-octave keyboard. Set the Jack audio buffer to 32 on both systems which seemed to reduce the latency, although we did not measure it scientifically (other than when feedback was set to max it actually fed-back at a piercing pitch). There was some delay in the video from 1005 to 2103, as we have seen before. Set at 12080x720 it appeared to have a 1/2 second delay. At 640x480 it was much less, but still not ideal (and the resolution in this case was very much not ideal).

Report:

Observations from Participant C: it was easy. Playing together, between floors of the building, was easy and responsive, very comparable to being co-present. At first this seemed like an insignificant fact, but is actually quite important. Compared to past experiences (with IUPUI and Arizona) this is a huge break through. Participant C described it as “fluid” (sliding), due to the changing delays we were able to impose.

We attempted a simple ‘game,’ with the latency set at various ‘high’ amounts (1 second+): to play single hits all in a verticality. The desired result, with everyone playing a note every chance they get, is a single pulse followed by silence at the rate of the round-trip delay.

Ben’s observation: it proved difficult at first until the rate of the delay, and the re-audition of the fed-back pulse, was internalized and then it was just a matter of counting a sub-pulse. Another trick: once a mistake note was played (i.e. a note that did not line up with the big attack) it could be heard as an anticipation, giving a time marker that was close to the ‘correct’ time. As a minimalist piece it was very interesting, especially since it meant everyone heard the “same thing” at each end, but offset in GMT by a second or more.

New ideas: smooth delay changes, we’ve only been able to try discrete changes so far. Also, 3-way telematic music opens up a lot more possibilities. Finally: instead of thinking of each end of a connection as an antiphonal ensemble (playing tightly together but only loosely responding to the other), split the “ensembles” between locations. I.e. have 2 string quartets playing with all the violins in one location and the violas/cellos in another. Have them play an octet but behave as 2 separate quartets, paying closest attention to their own “quartet” and then next most attention to members of the other quartet.

B.2.7 January 30th, 2011

Locations: NCSA 2103, NCSA 1005

Participants: Participant D (didjeridu), Ben Smith (violin)

Abstract: Again, the introduction of a new musician involved reviewing previous issues, making progress especially on understand the rhythmic possibilities and challenges. This session also confirmed some stylistic freedom in telematics, although the delay inherent in the setup still denies many common stylistic conventions (such as harmonic chord sequences).

Research Questions: Can the distributed environment support different styles of music and improvisation? Specifically, how does it transform rhythmic, groove based drone music and how do specific delay amounts impact the music?

Technical Notes: (not changed from before, just reiterated) 2103 (Participant D), big screen projection, Mac w/ 10 gig, HDMI, blackmagic video card, MOTU 828, 2 mics (AT Pro 35 on didj, Crown 700 for talk), Jacktrip for audio (512/44k), Max to send video and audio 1005: (Smith) Mac w/ 29" screen, blackmagic video card, Micron 10 gig, HDMI Canon camera, using mac audio, Crown 700.

Report: Played 4 times, once at lowest latency playing free to rhythmic to accel to decel to end, once trying to play rhythmic, once with an extra 60 ms of latency introduced, once with an extra 400 ms of latency introduced and artificial feed-back (~ -8 db at each end) introduced.

From Participant D: just like the uncanny valley theory in robotics a middle ground of uncanny delay made it very hard to play together. However, more delay made for solid rhythmic alignment and most musical comfort. At the lowest latency settings true synchrony was hard and took mental effort, but at higher settings it was easy to lock in at the period of the round-trip and jam.

B.2.8 February 3rd, 2011

Locations: NCSA 2103, RPI (experimental telematics lab)

Participants: UIUC: Ben Smith & Participant B. RPI: Participants E, F, and G, and crew (5 undergrads, comprising the Titinnabulate ensemble)

Abstract: This session comprised many more musicians, but still focused on interactions between pairs of musicians, now displaced by a greater physical distance. It was very successful, despite the fact that few of us had ever met in person (between the two institutions). For this session I composed a series of short "etudes" (see Appendix 3) that guided the experiments, focusing the playing on a number of previously identified questions and issues.

Research Questions: I composed a set of etudes for this session that focused on specific issues that had been identified from the previous sessions, above. The etudes are included in Appendix 3. The additional consideration of how to manage and facilitate a session with ten people was taken into account in the composition of the etudes.

Technical notes: Two mics, projector, conferenceXP video and 2 channel JackTrip audio in 2103. At RPI: many lines from different players (acoustic and direct/electric instruments), conference XP video, compressed. In NCSA the camera was placed on the west wall but we were forced to project the video on the north wall, thus we were either facing RPI (so they could see us), or looking at RPI (and they saw us from the side).

Report: The video connection was not especially good, we've had better going to NY city, showing lots of artifacts and frozen pixels. They weren't especially coordinated in setting things up so focused testing was not going to happen for the video. At least it was functional. Hopefully better next time. The delay seemed to be about half a second each way (or slightly quicker).

Audio checking took a long time, since they had to play each instrument and I had to give a thumb sign for more or less. In the end there were still some imbalances between instruments. We need a better way to do that (they were sending to us through an aux/bus out, not their main mix which they used to monitor their sound in their lab). We believe the audio was turning around at ~ 40 ms each way, maybe slightly longer.

We played two improvisations that took most of an hour(!), with pauses for discussion (which was very light, as I had expected).

Improv I: The first I called "Tuning" and directed the players to play together, in pairs (across the connection), trying to match pitch, energy, and timbre. We had both sides rotate players so that everyone got a chance to play in some combination. (this meant that Participant B and I switched several times while RPI only went around the ring once) The first couple players at RPI went quickly, simply intoning an A (440). By the end they were getting quiet varied and aiming more for the matching energy and timbre parts.

Reflections from others: the first couple people were sad that they had not experimented more and wished they had been at the end. They observed that no one looked at the monitor while playing, but most people looked at the screen when they were just listening.

From Me: the later players actually got the idea of the improv, feeling out the energy of the other player. I will call it "warm up" next time! Everyone seemed very responsive and were listening, although some of the instruments (the theremin, analog synth, and iPod) were either hard to change gears or the players were not very conversant on them. This became especially apparent in the second improv.

Improv 2: "Contrast" — players were directed to play in pairs as before (round robin), now playing either in complement or complete contrast, switching at any moment they saw fit.

From Others: The latency was so low as to not be an issue, it was just playing, not a big deal ("we've been doing this since last summer," from an undergrad, sarcastically: "we could try dial-up and then we'd

have some fun”). Participant E: the latency was perceptible, but “nice,” producing a soupy, fluid space of interaction and reactions.

From Me: some were better than others about making clear “complement” and “contrast” distinctions, probably in proportion to their skill as improvisors, not connected to the telematic aspects. We probably would have played worse in person . . . is it possible the focus required to play in this setting causes people to achieve higher? This is my new hypothesis. I can agree as far as my own playing is concerned, having the audio amplified but having the person removed causes me to listen much more intently and focus more. Not that this is universal or that better music can’t be achieved in person (I still believe the reverse, in fact), but that in this case it caused everyone to work better, play better. I’ve been in these class jams before, and they generally are bad. This was actually quite good, all things considered.

The improvs were long. Next time I’ll suggest tighter time limits (at even 2 minutes a person, at RPI, it would take 16 minutes).

It became more apparent that some of the instruments, or the players, couldn’t change very quickly, as the alternation of matching and not-matching was on a long period (like 1 minute), while others (the vocalist) were switching every few seconds, at times.

I am also learning that I cannot get most people to observe and think about the telematic nature. But I can get them to do things that expose elements to my eye. Thus it seems ineffective to try to use them as fellow researchers, but rather to employ the other musicians as test cases, setting them up to accomplish things and observing their progress.

We, Participant B and I, used the video on several occasions to note that an off-mic discussion was taking place. Without the video I would not have known what had happened—if they’d heard me, or if something had disconnected. During the actual playing it facilitated “continue” and “stop” moments—although no one but me noticed this!!

Regarding “Intimate Amplification:” Telematic chamber music places the ear of the remote participant inside, or right up against the instrument. This is never the case in acoustic music, and even amplified, co-present is different. Having ones only connection be through a microphone held against the lips of the vocalist, or in the bell of the saxophone, is a position few humans assume regularly.

RPI’s setup: they sat in a circle, several with their backs to the camera or completely blocked by someone else, with tables and lots of gear. We stood two in the center of the big room, with 2 mics, moving in and out of the camera as we played (i.e. treating it as a performance and a video recording). I felt, basically, as if RPI was not respecting the interaction, as they provided nothing to look at. Also, they were each small in the image (since there were 8 of them), while we 2 took up the whole image (thus being quite large in the

frame). I would prefer them to stage it as a performance, although apparently for them this is just a class in a cluttered studio. Just as performing musicians, especially classically trained ones, dress in a certain way and have practices (even rituals) of behavior during concerts, I approach telematic sessions similarly. For me this is a performance, even when it is also an experiment.

B.2.9 February 9th, 2011

Locations: NCSA 2103, NCSA 1005

Participants: Ben (violin), Participant H (hand percussion, laptop/samples)

Abstract: This session started out as an exploratory session with a new musician, and brought issues of audio balance and mixing to the fore. This was aggravated due to the differing needs of the two musicians, one that required local sound monitors, and the other that played acoustically.

Technical Notes: between NCSA 2103 and 1005, same as previous with Participant D. Video and audio delays appeared about the same (around 50 ms, mostly due to using the Mac internal DAC in 1005—it doesn't have a firewire 400 port). Participant H's audio was played back directly in the room (in order to give him a monitor of his electronic and quiet acoustic sounds).

We played three sessions, one from a drone through a free jam, adding delay and feedback towards the end, one based on a birimbau groove (highly rhythmic), and one completely free in which Participant H explored new electronic sounds.

Participant H said it felt comfortable and musical. He also spoke of being more or less focused (“in tune”) at different times, but felt that we were in fact “in tune” very well at some points.

For me audio balance became a focal concern. Due to Participant H needing audio monitors the levels were set somewhat oddly—his mic (for percussion) was very hot but the electronics were not, all of which was somewhat quiet at my end while I was very loud at his end. As a result I found myself playing very lightly, yet Participant H could hear me very well (relating to my “intimate telematics” theory). However when I was playing up and down to blend with what I was hearing it was coming out very different at Participant H's end (video evidence of this). I would play loudly to fill in a chord, dominating the sound in 2103, then back off to provide a subtle coloring and it would then fit into the mix in 2103.

A presentational performance would require extensive tech rehearsal and/or a live sound engineer with feeds from each source for the monitors and the house.

Rhythmic playing was no problem, I suspect because Participant H picked the period of the delay out of thin air when he started our groove improv. Thus it sounded quite tight, although again phased by one note.

B.2.10 February 10th, 2011. 2-4 p.m.

Locations: NCSA, 2103 and RPI (session #2)

Participants: Ben Smith, Participants A, B, C, and I (all at NCSA). Participants F and G and students at RPI.

Abstract: The continuation of the first session with RPI, however this time technical difficulties broke the video connection. This, on top of the already slightly stilted communication interactions made organizing and executing the session very difficult. However, proceeding through the Telematic Etudes (Appendix 3), uncovered some interesting questions regarding the ability to play together rhythmically, specifically: how can one know when the remote musician is really playing at the rate of the latency, or just tuning you out?

Research Questions: Continued exploration of the Telematic Etudes (see Appendix 3), seeking to explore the musical possibilities of distributed duets.

Technical Notes: Despite the technical success of the first RPI session (2/3/2011), we were unable to get bi-directional video working. I suspect it was on their end since nothing changed on ours, and they have very strict firewall protocols. As a result they were able to see us, but we could not see them. Part way through all video was lost.

We also attempted multi-channel audio through Jack (4 channels), which resulted in an uneven mix with a lot of clipping and audio artifacts. We needed more time to check out the audio, while we spent most of the 30 minutes of setup trying to get the video up.

Report: We played two improvs, one ‘warm-up,’ in which the musicians went in pairs (duets comprised of one musician at NCSA and one at RPI) round-robin, trying to listen as intently as possible to their remote partner. The results were of varying quality. The second was an attempt at playing metrically, but most of the musicians did not grasp the concept and played very amorphous things. The highlights were Participant C and the RPI guitar player jamming on a groove and myself and Participant F playing a quasi-swing riff. Afterwards I had their vocalist clap with me so that I could hear the round-trip delay time. Her claps were so variable I either have to conclude that there was a huge amount of audio jitter (sometimes she was even ahead of me!) or that she can’t clap a beat to save her life. The times that it settled it appeared we had a ~100ms round-trip time.

The take away thought of the day was one of discomfort and instability. [Participant E had recently expressed dissatisfaction about another collaboration over email and thus sat in a corner at RPI for this whole event, unlike the previous event where Participant E had played along.] The technical problems were very unsettling and not being able to see RPI made talking very hard. Also they all “spoke” through the one vocalist, which basically meant they did not speak at all. It was thus impossible to gauge their enthusiasm

or engagement. Video is very helpful for these moments!

Participant C observed that the lack of video made it very hard to perceive their musical “attitude,” i.e. if they were trying to play rhythmically or not, louder, softer, etc. His conclusion is that delayed video is fine, so long as it is fluid and the remote participants are visible and observable.

B.2.11 February 14th, 2011. 11-1 p.m.

Locations: NCSA 2103, NCSA 1005

Participants: Ben (mandolin & violin in 1005), Participant I (violin in 2103)

Abstract: This session was very spontaneous, allowing the participants to work from the research questions posed previously as well as the Telematic Etudes.

Research Questions: Can aesthetically “pleasing” music be produced over a network connection?

Technical notes: same setup as before. We found a faulty XLR cable that was producing some strange effects (apparently echoing the audio back over the network).

Report: The blend was great. Having 2 violins play together was very stimulating and satisfying, reaffirming the validity of the setup. No significant aesthetic advances were made other than to confirm the ability to play together and the support of harmonically rich, approachable music.

Towards the end of the video recording of this session there are some very accessible (i.e. “nice”) moments that demonstrate the stylistic freedom and musical potentials.

B.2.12 February 16th, 2011. 3-4:30 p.m.

Locations: NCSA 2103, Krannert Center for the Performing Arts (KCPA), Dance Rehearsal Room (DRK)

Present: Ben Smith and Participant I (NCSA), Participant A, C, and three (maybe more) other musicians and dancers (at KCPA)

Abstract: This session involved a dance technology class at KCPA, introducing the students to the issues of telematics. While video was not employed the musicians were able to achieve several successful improvisations using audio alone.

Research Questions: What are the implications for rhythmic improvisations in a telematic setting?

Technical Notes: For the first connection between KCPA and DRK we were unable to get video working (we tried to use Unreal Media Server software, which we are all somewhat unfamiliar with). JackTrip was used for the audio, working great with only one instance of brief buffer under-run. Latency appeared to be ~120ms round-trip.

At NCSA we used a single Crown 700 mic, MOTU 828, Mac Pro tower with 10 gig Myrinet NIC running JackTrip (2-channel) to send and receive audio. The MOTU was connected to the room PA (large sub, 2 JBL Ions). In KCPA they used a Mac pro laptop to send/receive audio (with a 100mbps or 1gbps NIC), the internal mac audio interface, a small mic, and the room's PA system (large speakers hung from the ceiling ~40 feet from the musicians).

Report: The class was very interested and approached it very analytically. We tried several combinations of musicians, playing variously pulsed (the KCPA crowd were all playing hand drums) patterns, speeding up and slowing down and getting a sense of the “togetherness” of the connected places.

Participant A and Smith had a very interesting progression, during a short rhythmic duet, where she held down a very steady 3-3-2 (eighth note) pattern and I matched it on the mandolin. We played at roughly the pulse of the round-trip time. After a minute she shifted, sounding one eighth behind me. I then took her pulse as the ‘true’ meter, and started switching back and forth between playing an eighth ahead of her, and playing with her (in which case Smith sounded one eighth behind her in KCPA). The sensation was of sending my sound in advance in order to line up at her end, versus accepting her sound and lining up at my end.

Thus I imagine a musician could easily learn to play ahead, as orchestral tuba players do, and be very happy feeling in sync, yet actually being an eighth of a second ahead of the remote players. This would be a huge shift in classical musical practice!

I had a chance to alter the delay amount (with a bit of audio feedback going in each direction, so the musicians could hear the delay) while they played which verified my intuitions. At several points they very clearly shifted the pulse to match the delay, changing the music in a very affective manner according to the length of the transmission time. This was a very strong sound, easily applicable to composition.

The KCPA musicians had not played over a telematic link before and commented that it felt very viable and satisfying, that they could really get used to playing in the telematic domain.

B.2.13 February 23rd, 2010. 1-2:30 p.m.

Locations: NCSA 2103, 1005

Participants: Participant I (1005, violin), Participant A (2103, flute), Ben Smith (observer, moving between both rooms)

Abstract: This session briefly attempted to make a co-present vs. distributed musical comparison, which revealed the inapplicability of such a research model. Several more prescribed improvisations were played, which led in interesting musical directions, exploring unison playing, once again, and the ability to

coordinate changes. Audio-video synchrony was very good, allowing for dramatic gestural cues.

Research Questions: How much can the audio delay be reduced without producing distracting transmission artifacts? Can an improvisation rehearsed face-to-face be compared to a distributed version of the same thing? Can musicians coordinate tempo gradual tempo changes effectively? What are the effects of asking musicians to play a score fragment together, in unison, and move back and forth from improvisation to score fragments?

Tech notes: set Jack buffers to 64 samples (1.45 ms long at 44.1kHz sample rate). However there were frequent blips (quiet click sounds), probably due to buffer under-run or packet loss. The delay was almost imperceptible, but the blips were annoying, to me. The musicians didn't seem to mind or notice the clicks and enjoyed the short round-trip time.

For the first time we used MOTU 828s on both ends (thanks to a firewire 800-400 converter), which also is contributing to a great round-trip latency (imperceptible, so on the order of 20 ms).

Video was very slow both ways at first, but after an hour (2 pm) it became very fluid. By this point I had set both ends to 320x240, 25 fps, in an attempt to reduce the delay (it was originally around one second!). At 2 p.m. it appeared to be in sync with the audio, so probably around 50 ms one way.

Report: improv 1: started as a planned piece practiced once in co-present proximity. Then played in a telematic displacement.

improv 2: started with idea of accelerating and decelerating over long periods (initially only one arc, but it kept on going at great length).

improv 3: used score fragment as formal delineator. I provided both musicians with the same short written score and asked them to play it in unison, with lots of rubato, between more freely improvised sections. They ended up playing in the form ABACA where the score fragment served as the A sections.

Both reported achieving an almost meditative state the longer the session went on!! They also both spoke of being able to do "this" for a very long time (asking for clarification was answered: "'this' as in playing together and playing telematically"). Both of these points indicate further the immersive-ness of the environment, and again possibly heightened by the need to listen and focus so thoroughly.

Participant I is very gestural, and tries to cue things (or his playing style is just very physical). It didn't work when the delay was large and frames were dropping, but it became very nice and effective with the low delay. In the last two improvs they were able to coordinate very well from Participant I's gestures. It raises the question of which is a higher priority: lowest latency audio, or audio-video synchrony? This is probably a compositional issue, as some pieces may require one or the other.

In my opinion the score fragment improvisation was the most compelling piece they played. Instinctively,

or perhaps by necessity, they drew a lot from the motivic material of the score fragment, producing a much more cohesive piece. Further listening, to the videos, may give clues as to how they coordinated returning to the score fragment. It seemed that each hinted at the return several times before the other finally caught on and joined in.

Participant I requested that Participant A's camera be moved so as to be more in front, as it had started at a 3 o'clock position relative to her facing the screen. We moved it to be closer to 1 o'clock and he was much more satisfied that they were able to "look at each other."

B.2.14 March 7, 2011, Monday. 1-2 p.m.

Participants: Participant J (cello), Ben Smith (violin)

Locations: NCSA 2103 and 1005

Abstract: Introduced another new musician and attempted some scripted, compositional improvisations. Various modes of reconciling feedback in the system were approached.

Research Questions: with a new person involved, how do the musicians reconcile playing together over a distributed connection, how does longer delay effect what is played, what are the effects of recording and looping everything that is played, and what are the results of setting a small amount of feedback in the system?

Technical Notes: Set the Jack buffers at 512 samples, producing a noticeable audio delay, but still very short (60-80 ms). After tweaking TCP settings on both machines I was able to send 1280x720 both ways with almost no difference from the audio (maybe another 20-40 ms behind the audio). It was very exciting! Note: crank the TCP buffer space to the max allowed on the machine (which is 1 MB in 2103 and 2 MB in 1005), and set window frame to 8.

Report: We played four improvisations. One with the system as it was, one with extra delay introduced, one with a delay loop running at 30 seconds, one with the delay of the system increased and loopback enabled.

For the first piece we played quite comfortably and felt in-sync at both ends. There was a slight amount of push and pull, but otherwise was very much like co-present jamming.

The second piece we tried to play at the period of the delay, rhythmically. Participant J afterwards observed that he basically decided to be internally consistent with his rhythm and "tune you out" (to Smith). He then commented that this was interesting, that one person could have a groovy experience while the other had to be generous and play alone, play for the other person. Thus we were probably not at the period of the delay, but Smith's end appeared aligned.

The third piece was based on my attempt to confuse which "copy" is the real one. That is, if all the

sound we hear of the other is a digital reproduction than all reproductions should sound equally plausibly “real.” With a loop set at 30 seconds (with only a few db of decay per iteration) the sound quickly built to a huge density where the actually real player was buried. In the end it was either a) easy to tell what the new material, and thus the actually real playing, was or b) it was so chaotic that it was all equally unreal. Perhaps it needs to be not a loop but separate play heads reading out of one buffer—allowing for different time spacings and changes of density. Unfortunately here all the copies became equally unreal, and not convincing.

The final piece had ~250ms of delay and the audio from each player going back in to the loop with -5 db attenuation. The sound was then a “reverberated” echo, at the period of the delay. We were able to play both freely and rhythmically together, moving from thin to very full textures. Participant J commented that it was the most satisfying of our improvisations, due to the “naturalness” of the sound.

B.2.15 March 11, 2011, Friday. 3:30 p.m.

Participants: Participant K (voice), Ben Smith (violin)

Locations: NCSA 2103 and 1005

Abstract: Introduced another new musician to the setup and explored several composed improvisations. The strongest result was a layering, looping piece that attempted to confuse the perception of the “live” player with their replicated sounds.

Research Questions: same as previous session with another new musician.

Technical Notes: same as previous session.

Report: After introducing Participant K to the setup we played several improvisations along the lines of the previous session with Participant J. Participant K took to the environment immediately, listening well and producing some very compelling vocalizations and musical passages. Our rendition of the 30 second building loop piece was very exciting, almost frightening. As it looped again and again we both found ourselves picking up on sounds we heard and reinforcing them again and again. Perhaps having two instruments in the same range made the texture more cohesive. Also, our musical gestures were more abstract, non-tonal often and certainly non-harmonic. The end effect was as a long series of waves that built and built, crashing and repeating every 30 seconds (because we joined our previous selves after a small delay each time the layering-in was feathered, producing very continuous crescendos.)

Appendix C

Appendix 3

C.1 Telematic Etudes

The following are a series of experiments in telematic improvisation and performance, examining the musical relationships between two or more distributed performers. They are intended as frameworks for exploration, to be pursued as long as the participants desire. While many have a stated aim or objective, the true goal is to open the realm for creative research and discovery. Open, general wording is left to the interpretation of the individual participants.

Tuning

Two musicians play long, full, steady tones. Their goal is to align themselves and their instruments in pitch, energy, timbre. The musicians should stop when both are content and satisfied.

Complement

Two musicians improvise together, playing complementary sounds, textures, melodies, material. Allow the improvisation to evolve by staying as close together, musically, as possible.

Contrast

Two musicians improvise together choosing at any moment to play material as similar as possible to that of the other player, or as contrasting as possible. Each musician may choose and change at any point, staying together or separating after short bursts or after long periods.

Meter

Two musicians begin playing rhythmically consistent material, attempting to play together as accurately as possible. Each musician should always speed up (within reason) to “catch” the other player and match their rhythm. The goal is to arrive at a place of stability, where both are playing in rhythmic synchrony (although each will hear the other offset by one or two notes).

Synchronic Flow I

Two musicians alternate between *Meter* and *Complement*, creating an alternating texture of rhythmic convergence and textural synchrony. Begin with *Meter*, aiming for rhythmic stability. Once achieved, both

players switch to a different texture and begin *Complement*. At any moment either player may begin *Meter* again, searching again for rhythmic stability, repeating this sectional alternation in a reborn “rondo” form.

Quartet I

Four musicians, two at each location, comprise a two duet ensemble. Duo A is made up of one musician at each location, while the other two form Duo B. Duo B begins the improvisation, maintaining an accompanimental role throughout, working together to support and reinforce Duo A. Duo A joins, providing the solo/lead/melodic component, working together to shape and guide the improvisation.

[Note: the audio for each musician may be sent through a single channel, or as two separate channels. In the later case they should be mixed and spatialized analogous to their original setup, i.e. gently, not distinctly separated.]

Quartet II

Four musicians, two at each location, form an improvisatory quartet. At any point each musician may choose between *Complement* and *Contrast* with any of the other musicians. The choices are informed through listening locally and across the network. All four musicians come together here as a single dynamic unit, responding quickly and fluidly.

Synchronic Flow II

Three or more musicians begin with *Complement*. Gradually, all move towards *Meter*, attempting to play in rhythmic synchrony with all the other musicians. Each should always anticipate and accelerate to “catch” the other musicians, as much as possible. Once rhythmic stability is achieved the ensemble returns to *Complement* and ends.

Dialogues

Any number of Duets are formed between pairs of musicians at any location (i.e. no duet is made up of co-located musicians). Each duo plays as a unit, listening foremost to one another, followed by the other duos. Every duo improvises their own voice and dialogue, drawing on material from other duos or distinguishing themselves with unique material. Every musician listens continually, waiting when there is more to be heard, playing when everything has been heard.

References

- Ahmad, I. “Multi-view video: get ready for next-generation television.” *Distributed Systems Online, IEEE* 8, 3: (2007) 6–6.
- Ascott, R., and E.A. Shanken. *Telematic embrace: Visionary theories of art, technology, and consciousness*. Univ of California Pr, 2007.
- Barbosa, Á. “Displaced soundscapes: A survey of network systems for music and sonic art creation.” *Leonardo Music Journal* 13, 1: (2003) 53–59.
- Bartlem, E. “Reshaping Spectatorship: Immersive and Distributed Aesthetics.” *Fibreculture Journal* 7.
- Biocca, F. “The cyborg’s dilemma: Embodiment in virtual environments.” In *Cognitive Technology, 1997. ‘Humanizing the Information Age’. Proceedings., Second International Conference on*. IEEE, 1997, 12–26.
- Blum, D., and G. Quartet. *The art of quartet playing: The Guarneri quartet in conversation with David Blum*. Cornell Univ Pr, 1987.
- Braasch, J. “The Telematic Music System: Affordances for a New Instrument to Shape the Music of Tomorrow.” *Contemporary Music Review* 28, 4: (2009) 421–432.
- Cáceres, J.P., and C. Chafe. “JackTrip: Under the hood of an engine for network audio.” *Journal of New Music Research* 39, 3: (2010) 183–187.
- Cáceres, J.P., R. Hamilton, D. Iyer, C. Chafe, and G. Wang. “To the edge with china: Explorations in network performance.” In *ARTECH 2008: Proceedings of the 4th International Conference on Digital Arts*. 2008, 61–66.
- Chafe, C., and M. Gurevich. “Network time delay and ensemble accuracy: Effects of latency, asymmetry.” In *Proceedings of the AES 117th Convention*. 2004, 6208.
- Chew, E., A. Sawchuk, C. Tanoue, and R. Zimmermann. “Segmental Tempo Analysis of Performances in User-Centered Experiments in the Distributed Immersive Performance Project.” In *Proceedings of the Sound and Music Computing Conference, Salerno, Italy*. 2005a.
- Chew, E., A. Sawchuk, R. Zimmermann, T.P. Duo, C. Kyriakakis, C. Papadopoulos, A. François, and A. Volk. “Distributed Immersive Performance: Ameliorating the psychophysical effects of network latency.” *Proceedings of the Corporation of Education Networks in California (CENIC)* .
- Chew, E., AA Sawchuk, R. Zimmerman, V. Stoyanova, I. Tosheff, C. Kyriakakis, C. Papadopoulos, ARJ François, and A. Volk. “Distributed immersive performance.” *Proceedings of the Annual National Association of the Schools of Music (NASM), San Diego, CA* .
- Chew, E., R. Zimmermann, A.A. Sawchuk, C. Kyriakakis, C. Papadopoulos, ARJ François, G. Kim, A. Rizzo, and A. Volk. “Musical interaction at a distance: Distributed immersive performance.” In *Proceedings of the MusicNetwork Fourth Open Workshop on Integration of Music in Multimedia Applications, September*. 2004b, 15–16.

- Chew, E., R. Zimmermann, AA Sawchuk, C. Papadopoulos, C. Kyriakakis, C. Tanoue, D. Desai, M. Pawar, R. Sinha, and W. Meyer. "A second report on the user experiments in the distributed immersive performance project." In *Proceedings of the 5th Open Workshop of MUSICNETWORK: Integration of Music in Multimedia Applications*. Citeseer, 2005c.
- Cooperstock, Jeremy. "History of Spatially Distributed Performance.", 2011, Accessed June 8th, 2011. <http://www.cim.mcgill.ca/sre/projects/rtnm/history.html>.
- Csikszentmihalyi, M. *Flow: The psychology of optimal experience: Steps toward enhancing the quality of life*. Harper Collins Publishers, 1991.
- Dallow, P. "Representing creativeness: Practice-based approaches to research in creative arts." *Art, Design & Communication in Higher Education/The Journal of the UK Learning & Teaching Support Network's Subject Centre for Art, Design & Communication* 2, 1: (2003) 49–66.
- Dessen, Michael. "New Polyphonies: Score Streams, Improvisation and Telepresence." *Leonardo Music Journal* -: (2010) 21–23. http://www.mitpressjournals.org/doi/abs/10.1162/LMJ_a_00007.
- Dresser, M. "Telematics.", 2008, Accessed Dec 22, 2010. <http://www.allaboutjazz.com/php/article.php?id=30198>.
- van Eijk, R.L.J. *Audio-visual synchrony perception*. Unpublished doctoral dissertation, Technische Universiteit Eindhoven, 2008.
- FFMpeg. "FFMpeg.", . <http://www.ffmpeg.org/>.
- Frayling, C. "Research in art and design." *Royal College of Art Research Papers* 1, 1: (1993) 1–5.
- Giddens, A. *The consequences of modernity*. Cambridge: Polity, 1990.
- Gresham-Lancaster, S. "The aesthetics and history of the hub: The effects of changing technology on network computer music." *Leonardo Music Journal* 8: (1998) 39–44.
- Haseman, B.C. "A manifesto for performative research." *Media International Australia Incorporating Culture and Policy: quarterly journal of media research and resources* , 118: (2006) 98–106.
- Held, Richard. "Telepresence." *The Journal of the Acoustical Society of America* 92, 4: (1992) 2458. <http://link.aip.org/link/?JAS/92/2458/2>.
- Jeong, B., L. Renambot, R. Jagodic, R. Singh, J. Aguilera, A. Johnson, and J. Leigh. "High-performance dynamic graphics streaming for scalable adaptive graphics environment." In *Proceedings of the 2006 ACM/IEEE conference on Supercomputing*. ACM, 2006, 108–es.
- Krumhansl, C.L. "Rhythm and pitch in music cognition." *Psychological Bulletin* 126, 1: (2000) 159–179.
- Lerdahl, F., and R. Jackendoff. *A generative theory of tonal music*. The MIT Press, 1996.
- Lombard, M., and T. Ditton. "At the heart of it all: The concept of presence." *Journal of Computer-Mediated Communication* 3, 2: (1997) 0–0.
- Mantovani, G., and G. Riva. "'Real' Presence: How Different ontologies generate different criteria for presence, telepresence, and virtual presence." *Presence* 8, 5: (1999) 540–550.
- Mori, M. "The uncanny valley." *Energy* 7, 4: (1970) 33–35.
- Munster, A., and G. Lovink. "Theses on Distributed Aesthetics. Or, What a Network is Not." *Fibreculture Journal* 7.
- Nora, S., and A. Minc. *Computerization of society: a report to the President of France*. MIT Press, Cambridge, MA, 1980.

- Oliveros, P., S. Weaver, M. Dresser, J. Pitcher, J. Braasch, and C. Chafe. “Telematic Music: Six Perspectives.” *Leonardo Music Journal* 19: (2009) 95–96.
- Olmos, A., M. Brulé, N. Bouillot, M. Benovoy, J. Blum, H. Sun, N.W. Lund, and J.R. Cooperstock. “Exploring the role of latency and orchestra placement on the networked performance of a distributed opera.” In *12th Annual International Workshop on Presence, Los Angeles*. 2009.
- Pastor, M.A., and J. Artieda. *Time, internal clocks, and movement*. North Holland, 1996.
- Pettey, G, C.C. Bracken, and B. Rubenking. “No Vision is Complete without Sound: The Sonification of Telepresence.” In *Proceedings of the 11th Annual International Workshop on Presence*. Padova, 2008.
- Phelan, P. *Unmarked: The politics of performance*. Burns & Oates, 1993.
- Pöppel, E. “Reconstruction of subjective time on the basis of hierarchically organized processing system.” *Advances in Psychology* 115: (1996) 165–185.
- . “A hierarchical model of temporal perception.” *Trends in Cognitive Sciences* 1, 2: (1997) 56–61.
- Reeves, S., S. Benford, C. O’Malley, and M. Fraser. “Designing the spectator experience.” In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2005, 741–750.
- Renambot, L., A. Rao, R. Singh, B. Jeong, N. Krishnaprasad, V. Vishwanath, V. Chandrasekhar, N. Schwarz, A. Spale, C. Zhang, et al. “Sage: the scalable adaptive graphics environment.” In *Proceedings of WACE*. Citeseer, 2004, volume 2004.
- Sawchuk, A.A., E. Chew, R. Zimmermann, C. Papadopoulos, and C. Kyriakakis. “From remote media immersion to distributed immersive performance.” In *Proceedings of the 2003 ACM SIGMM workshop on Experiential telepresence*. ACM, 2003, 110–120.
- Schiphorst, T. *The Varieties of User Experience*. Ph.D. thesis, The University of Plymouth, 2009.
- Schuett, N. *The effects of latency on ensemble performance*. Undergraduate honors thesis, Stanford, 2002.
- Tofts, D. “tsk tsk tsk and Beyond: Anticipating Distributed Aesthetics.” *FibreCulture Journal* 7.
- Turino, T. *Music as social life: the politics of participation*. University of Chicago Press, 2008.
- Weinberg, G. “The aesthetics, history, and future challenges of interconnected music networks.” In *Proceedings of the International Computer Music Association Conference*. 2002, 349–356.
- Zimmerman, J., J. Forlizzi, and S. Evenson. “Research through design as a method for interaction design research in HCI.” In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2007, 493–502.