Title
Three Video Scores and their Compositional & Notational Strategies

Permalink
https://escholarship.org/uc/item/2rv043m7

Author
Oram, Celeste Alice

Publication Date
2016

Supplemental Material
https://escholarship.org/uc/item/2rv043m7#supplemental

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Three Video Scores and their Compositional & Notational Strategies

A Thesis submitted in partial satisfaction of the requirements for the degree Master of Arts

in

Music

by

Celeste Alice Oram

Committee in charge:

Professor Rand Steiger, Chair
Professor Ricardo Dominguez
Professor Thomas Erbe

2016
The Thesis of Celeste Alice Oram is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2016
# TABLE OF CONTENTS

Signature Page ........................................................................................................ iii

Table of Contents ................................................................................................... iv

List of Supplemental Videos ................................................................................. v

Acknowledgements .................................................................................................. vi

Abstract of the Thesis .............................................................................................. vii

Chapter 1 Video as Score: Mechanics & Matrices .............................................. 1

1.1 Score & Screen / Surface & Interface ......................................................... 8

1.2 The Musical History of Video ..................................................................... 11

1.3 Semiotics of the Screen ............................................................................. 18

1.4 Spitting the Sound-Image Atom ................................................................. 23

Chapter 2 rupture | rapture & $8 \times \infty$: The Mirror Score .................................. 30

Chapter 3 XEROX ROCK: YouTube Pedagogy & Database Aesthetics .......... 42

3.1 Composing the Database ........................................................................... 43

3.2 Disembodied Pedagogy & the Screen ....................................................... 48

3.3 Time in the Digital Dimension ................................................................... 55

Appendix .................................................................................................................. 64

i. rupture | rapture text instructions for performance ...................................... 65

ii. XEROX ROCK text instructions for performance ........................................ 67

iii. $8 \times \infty$ text instructions for performance .................................................... 70

References ............................................................................................................... 75
LIST OF SUPPLEMENTAL VIDEOS

1. oram_01_rupture-rapture_score.mov

rupture \ rapture audio-video score

2. oram_02_rupture-rapture_performance.mov

rupture \ rapture performance video; performed by Hillary Young (voice) & Fiona Digney (percussion), 16 January 2015

3. oram_03_xerox-rock_score.mov

XEROX ROCK video score

4. oram_04_xerox-rock_performance.mov

XEROX ROCK video score plus performance audio & video; performed by wasteLAnd: Nicholas Deyoe (electric guitar), Richard Valitutto (piano) & Justin DeHart (marimba), 23 May 2015

5. oram_05_8x8_score.mov

8 x ∞ audio-video score

6. oram_06_8x8_performance.mov

8 x ∞ video score plus pre-recorded performance audio; performed by the Callithumpian Consort: Sarah Brady (flute), Gary Gorczyca (clarinets), Yukiko Takagi (piano), Mike Williams (percussion), Jeffery Means (percussion), Gabriela Diaz (violin), Stephen Marotto (cello), David Goodchild (double bass), 23 October 2015
ACKNOWLEDGEMENTS

I am grateful for the venturesome investment of time and energy made by the performers who have taken on the scores in this portfolio:

- *rupture | rapture*: Hillary Young & Fiona Digney; Charissa Noble & Christopher Clarino; Emerald Lessley & Chelsea Tinsler
- *XEROX ROCK*: Nicholas Deyoe, Richard Valitutto, Justin DeHart
- *8 x ∞*: Sarah Brady, Gary Gorczyca, Yukiko Takagi, Mike Williams, Jeffery Means, Gabriela Diaz, Stephen Marotto, David Goodchild

The performance of *XEROX ROCK* would not have happened without the initiative and organization of Katharina Rosenberger and wasteLAnd. Thanks also to Nicholas Deyoe, who produced the rehearsal recordings for this piece.

The performance of *8 x ∞* would not have been possible without the generous co-operation of Stephen Drury and Rand Steiger. I am especially grateful to Yukiko Takagi and Stephen and Bill Drury for offering their home studio as a space for recording the audio for this piece.

Final thanks to Jessica Flores and the UCSD production staff, whose magic made the performances of these works so technically elegant; and to Brady Baker, for supplying the equipment these productions demanded.
ABSTRACT OF THE THESIS

Three Video Scores and their Compositional & Notational Strategies

by

Celeste Alice Oram

Master of Arts in Music

University of California, San Diego, 2016

Professor Rand Steiger, Chair

Composing music means score-making; and score-making means systems-making. A musical score is a system—a matrix—of control.

My recent compositional work investigates new media and strategies for managing matrices of control in musical notation: namely, video scores, which prescribe in real-time a musician’s movement and actions, rather than the sound they are expected to produce. The performers’ task is to mirror the actions demonstrated in the video score in such a way that sound is produced on their instrument.
This thesis examines my notational and compositional approach to the medium of the video score, illustrated by three recent works: *rupture \| rapture* for voice and percussion, *8 \times \infty* for mixed octet; and *XEROX ROCK* for electric guitar, piano, and marimba.

The first chapter outlines the theoretical and aesthetic contexts in which I situate my approach to video scores: theories of the interface, the screen, and interactivity; the audio-visual interdependence of the video medium, taking into account early video art; and the semiotics of the screen.

The second chapter discusses two works of the triptych—*rupture \| rapture* and *8 \times \infty*—and the mimetic relationship between score and performer. Both of these scores include not only video but also audio material, the sounds of which the performers are called upon to recreate with their instruments.

The third chapter addresses the third work of the portfolio, *XEROX ROCK*, in reference to the aesthetic and pedagogical qualities of the YouTube videos from which the score’s video footage is sourced.
CHAPTER 1

VIDEO AS SCORE: MECHANICS & MATRICES

I open this thesis with the claim that composers do not create sound. Performing musicians create sound; composers create the scores which induce and influence their actions. A score is, by broad definition, a system of control – a set of constraints which prompt and influence a scenario. By setting such limits, a score bounds what is taken to be ‘a work’: a thing of conscious and unique design.

Sol Lewitt’s score for Wall Drawing #260: On Black Walls, All Two-Part Combinations of White Arcs from Corners and Sides, and White Straight, Not-Straight, and Broken Lines, for instance, specifies constraints for the proliferation of lines in two-dimensional space. A musical score is a set of constraints which prompt and influence the generation of sonic events in time. Yoko Ono’s score for Voice Piece for Soprano, for instance, specifies constraints for a sequence of screams: who screams, where, and in which direction. Insofar as soundwaves, literally, are the audible residue of physical impulses and disturbances, a musical score induces those impulses and disturbances which generate sound. Composing is score-making; and score-making is systems-making.

The medium (or media) via which a score’s constraints are communicated wields significant influence over what can be controlled, as well as the degree of
control. Historically, mainline Western music composition has, since the fifteenth century, employed a standardized, literate (i.e. paper-based) staff notation system optimally calibrated around two axes: one (vertical) specifies systematized, even-tempered pitch; the other (horizontal), metrically organized rhythmic units. But there are countless alternate ways of communicating constraints for a musical work. Musical scores can be verbally or demonstratively transmitted (in the case of oral music traditions), collectively negotiated (in the case of rock bands or improvising ensembles), alluded to through visual abstraction (in the case of graphic scores), expressed through literary instructions (in the case of text scores), or digitally programmed (in the case of computer music). None of these scores is fully determined. Each—to differing degrees—circumscribe a set of multiple possible realizations within the constraints of their control. Even a meticulously programmed computer music work may allow for its performance on different speaker equipment, or in various social, cultural, and relational circumstances. Even in the case of an expansively permissive graphic or text score, its very nature as a score presumes that the score will exert foundational influence over an interpreter’s performance – that, without the score, there would be no event. In all of these cases, the defining characteristics of the sounding musical result are deeply contingent on the parameters and degrees of control which the score-medium is capable of constraining.

In creating the works in this portfolio of compositions, I have explored the medium of video as a tool and platform for score-making. There is vast latitude for
notational strategies within the video medium, as evidenced by the numerous and creative ways other contemporary composers are employing video in their score-making. To mention only a few: the video score for Daniel Portelli’s solo piano piece *Mapping Australia* (2014) combines archival video footage of a 1966 documentary about cartography and land management in Australia—the gestures of whose human figures denote the rhythmic material to be performed—with textual cues describing percussive actions to be performed inside the piano, plus cross-references to an auxiliary sheet music score. In *Animal* (2015), another of Portelli’s works, a choral ensemble follows a video score which comprises graphics of staff-notated pitch material, text and phonemes in scrolling animation, and time-manipulated video footage of a match being struck (which indicates rhythmic material). Mayke Nas’ *Anyone Can Do It* (2006) is scored by a video of animated slides with text instructions to be executed by six unrehearsed audience volunteers. Michael Baldwin’s *this is not natural* (2014), for piano, double bass, and French horn, takes video footage of the performing musicians themselves as source material, but then time-stretches this footage to radical extremes and instructs the performers to re-perform these hyper-protracted actions.

Amidst these varying approaches, at this point it is necessary to explicitly define my own principles and approach to using video as a notational tool. My turn to using video as a notational medium was prompted by a disinclination to take for granted pitch and rhythm as the most salient features of musical composition. A staff notation system, by affording the greatest notational control over pitch and
rhythm, implicitly hierarchizes pitch and rhythm as the parameters afforded the greatest attention – whether in composition, listening, or analysis. As my compositional attention steered instead toward dimensions such as physical gesture, the quality of a performer’s comportment, and the interplay between sound and action, it became necessary to employ a categorically different notational system – one that would afford functional control over those dimensions.¹ Turning to video for making scores offered a way for me to compose less in reference to the arbitrary systems with which musical sound has historically been organized, but more in reference to physical, material realities. As composer and computer musician Curtis Roads argues, “to pick up a gold-tipped fountain pen and inscribe a treble clef and key signature on onionskin staff paper is to import a system of composition. It entails a selection from an enormous range of possible approaches to making music, a palette of sounds, their combinations, a mode of performance, and even the audience. Every subsequent compositional decision articulates the chosen aesthetic” (Roads 326).

The underlying notational strategy of the scores in this portfolio (and my earlier video scores) employs video footage of moving human figures to demonstrate the physical actions which performing musicians are to execute in such a way that sound is produced on their instrument. The video score itself is silent; it has no soundtrack. Performers follow the score much like a Zumba tutorial

¹ It has becomes common practice for composers to employ — within the staff notation system — textual instructions, symbols, and tablatures to notate gesture or action where necessary. However, such notational strategies tend to necessitate additional reinforcement by way of verbal instruction or physical demonstration, and their notation is thoroughly idiosyncratic from composer to composer.
video: they mirror the actions in the video in real-time. The more precisely synchronous their actions are with the video, the more successful their performance. The video footage has been flipped horizontally to facilitate this mirrored relationship: for instance, if the figure in the video moves their hand on the right side of the screen, the performer moves their own right hand. The footage indicates not only the rhythm and force of the gesture, but also its position in relation to the body of the performer and the body of their instrument, the fine motor skills and actions to be executed, and, in some cases, the affect with which the gesture is performed. The performer is explicitly instructed not to re-create the sounds they imagine would be in the video’s soundtrack. Instead, they attend to meticulously re-performing the physical quality and orientation of the gestures in the score, and always with the imperative to produce instrumental sound. This throws up surprising anomalies: there are, after all, some parts of a string instrument which, however vigorously you bow on them, won’t release more than an asthmatic wheeze of a sound.

A primary concern in creating these scores has been that they be auto-didactic: that they avoid deferring to abstracted representational systems in which an interpreter must be briefed and educated. Though it is necessary for the video scores to be supplemented by concise written instructions which describe the performer’s task of synchronous mirroring, the scores avoid ‘mixing media’: there are no diagrams, textual titles, or vestiges of staff notation implanted into the video score itself. While the score is running, therefore, maximal attention is afforded to
reading the physical gestures and actions being represented. The expectation is that a performer learns to better interpret these scores simply by repeatedly interpreting them—encoding them into muscle memory—in much the same way that skill in a video game is developed through repeated gameplays.

Another crucial feature of these video scores is that the video file is fixed and final. Unlike real-time, algorithmically-generated video, which can respond to sound, touch, gesture, and other factors in their live environment, my video scores are a fixed document, ‘composed’ and crafted before being revealed to their interpreters.

Video scores considerably re-arrange the degrees and kinds of control which can be exerted over the sounding musical result. Insofar as video represents a precisely-edited sequence of events-in-time, musical rhythm can still be controlled exactly—arguably, more so than on staff notation because of its unequivocalness—as can form, insofar as the eventual .mov video file is a fixed media which plays from beginning to end without interruption. Where the piece is scored for more than one instrument, split-screen video allows for the execution of polyphony between parts, thereby exerting control over texture.

Pitch, however, in a video score cannot efficiently be specified in reference to a twelve-tone system. It can be specified in reference to the physical dimensions of a musical instrument (e.g. high end vs. low end). It can be specified relationally, with broad strokes of contour and compass. It can at any moment be destabilized by the amount of force or pressure the physical gesture exerts on the sounding
object – especially in the case of string instruments, where excessive or minimal bow pressure can skew the pitch unpredictably. But, fundamentally, pitch is incidental to the sounds the score prompts; pitch is not the object of the performer’s attention, nor does the pitch content remain consistent between each run-through of the score.

In this thesis, discussion will implicitly gravitate towards the particular uses to which I have put video in the scores in this portfolio, rather than grappling with the generic implications of video notation, its performance practice, and distribution – whose scope is broader than this project allows. Chapter One of this thesis will explicate a theoretical framework for video-as-score, taking into account the medium-specificities of the screen and of video, the semiotics of the screen, and the interface between sound and screenal image. Chapters Two and Three will discuss specific works from my portfolio in reference to this framework. Chapter Two will address two works: rupture | rapture (2014), an audio-video score for voice & percussion with live video; and 8 x ∞ (2015), an audio-video score for flute, bass clarinet, piano, two percussionists, violin, cello, & double bass with fixed 8-channel audio media. These two works share common mechanics: their scores comprise not only video but also audio prompts; the video footage used in the score is purpose-filmed, capturing choreographed sequences of actions; the video score remains visible only to the performers, and is not displayed to the audience. Chapter Three will discuss a work whose score involves markedly different mechanics: XEROX ROCK (2015), a video score for electric guitar, piano,
and marimba with fixed 2-channel audio media. This video score comprises found video footage gleaned from online media sharing platforms (chiefly YouTube and Vimeo); the video score itself is not only read and followed by performers, but is simultaneously made visible to the audience in performance.

1.1 SCORE & SCREEN / SURFACE & INTERFACE

A musical score serves as the interface between the maker of the score and the interpreter of it. This interface mediates constraints of control to the bodies and objects influenced by those constraints. The interface-surface of a staff-notation score is most often a sheet of paper: a singular material object, destructible, inert, able to be scribbled on with pencil, cut up and re-arranged, laid out to be observed all at a glance.

The interface-surface of a video score is a screen: a material object which itself has little character, but which is capable of displaying a vast library of media. The media the screen displays is temporal, drip-fed to its viewer frame-by-frame; it runs sequentially unless physically interrupted and navigated by the viewer; it cannot be altered without the aid of specialist software.

In the case of cinema—the earliest mass-distributed use of the screen as a site of a technologically mediated moving image—the screen was an untouchable site of transmission, mass-mediating from one to many. Yet, in the intervening century, the changing nature of the media exhibited on screens (not to mention the
evolving technologies which capture both still and moving images) has changed the dynamics of the screen’s interface. A viewer can now select and control the contents of a screen, implant their own image onto it, even touch it.

Musicologist Holly Rogers makes the case that video—and the screen it inhabits—“cannot be considered any one thing”. Rather than being definable as a fixed medium like painting or sculpture, video is a “culture”: a set of relationships around its uses, “a facilitator of intermedial discourse rather than a genre” (Rogers 39). ‘The screen’ is less and less a site of passively received transmission; the screen is more and more embroiled in a ‘screenal culture’ of active consumption, control, and interaction. A video score is planted firmly within this screenal culture: despite being a fixed media file, its ontology as a score demands human interaction by its implicit invitation to interpret the score. In its context as a musical score, the video-score file is not only a product of a medium but also a node within a network of intermedial relationship and interaction.

Paradoxically, the compounding interactivity of the screen surface/interface exacerbates what Liz Kotz describes as the screen’s “seductive immateriality”: the point at which “the medium itself disappears and leaves us in the presence of the thing represented” (Rogers 372). The greater the volume of contact with the interface, the more likely the interface is taken for granted, permeated to the point of invisibility. Some glitch or hiccup in the system becomes necessary to frustrate the illusion of ‘immersion’ in the screen, and re-remind the user of the screen’s presence as mediating interface. As Michel Serres observes that interfaces “work
because they don’t work. Non-functionality remains essential for functionality. If the exchange succeeds – if it is perfect, optimal, immediate – then the relation erases itself. But if the relation remains there, if it exists, it’s because the exchange has failed” (Galloway 26).

Insofar as a score is an interface for control and constraint, by extrapolation, scores work because they don’t work. Scores maintain their mediating function as a set of constraints which invite varied interpretation precisely because a score cannot prescribe a fully determined set of constraints. And yet the paper surface of the staff-notation interface, like the screen, also hovers on the threshold of disappearing into the thing it represents: the sounding musical work. In the concert music tradition, the ubiquitous presence of the paper score is taken for granted, and the score is oftentimes kept hidden, furtively turned away from the audience’s view. Under these circumstances, the medium of the score dissolves into the much more readily perceivable thing-represented: the musical work, as realised by the performing musician. At the point when a notation system becomes flawlessly efficient in communicating its constraints, the score disappears. Indeed, the continuing practice of learning works by heart—doing away with the score entirely—as a demonstration of a performer’s ultimate mastery suggests that the paper score’s deepest ambition is to disappear.

A video score exists in a constant state of disappearance: each frame disappears the moment it appears. It is a paradox video shares with cinema; Deleuze posits the cinematic image as “present and past, still present and already
past, at once and at the same time. If it was not already past at the same time as present, the present would never pass on” (Deleuze 79). Once a video score’s present has passed on and run its full course, it vanishes entirely, leaving behind only two visible artefacts: the graphical-user-interface object whose clicking makes possible the score’s replay, and the screen, now empty of video media and re-instated as a site of human-computer interface and navigable control. However, the temporality of the video score—the fact that it can never be seen at-a-glance or glossed over—dictates that video scores cannot be memorized (at least not without painstaking labour) and thereby disappear. What’s more, the very function of the video score as a temporal prompt, to be enacted synchronously, dictates that the score must needs be followed: that the score’s temporal presence—its play-through on the screen—is the sine qua non for realizing the piece that the score notates. The video score’s inerasable presence in performance thereby serves as that glitch, that “non-functionality” (as Serres describes) which preserves the mediating relation of the musical score as an interface between composer and performer. Rather than disappearing into the thing it represents, the video score remains present as the visible scaffolding constraining and controlling the musical work.

1.2 THE MUSICAL HISTORY OF VIDEO

Despite the obvious parallels between cinema and video, conflating them as analogous mediums risks overlooking the essential anatomical differences of their
technical mechanics and—crucially—their differing engagement with sound. To present an introductory hypothesis: the cinema screen is an image-surface alongside which sound is produced separately and received as ancillary to the antecedent image. The video screen, on the other hand, is an interface of active intermedial exchange, where a sonified moving image is heard as clearly as it is seen.

The lynchpin argument of Holly Rogers’ history of video art is that video technology’s most striking feature is its enabling of a “synergy” between image and sound “rarely before possible” (Rogers 1). Encoded onto video tape as electromagnetic impulses, sound and image share a material channel, allowing them to be recorded and transmitted simultaneously. Film, by contrast—along with live audiovisual genres such as opera, theatre, and dance—is “intermedial primarily at the level of reception” rather than the moment of creation. The soundtrack of a film must be recorded separately and inscribed onto the optical track of celluloid in post-production. In addition, cinema’s sound production process most often involves a different roster of crew and creative personnel to those who are responsible for the film’s moving image.

Video, on the other hand, is in essence a medium which not only captures a moving image, but which makes a moving image audible. The Sony Portapak, introduced to the market in 1965, was portable and affordable to the independent amateur consumer. Most importantly, by recording sound at the same time and on the same channel as video, it allowed one operator to capture both sound and
image, collapsing authorship into one video-composer. Moreover, as John Belton describes, “film deliberates” (Belton 65); the technological mechanics of cinema production lend themselves to meticulous construction and collage. Conversely, splicing and editing analogue video tape is a notoriously imprecise, laborious, and unreliable (and therefore seldom-executed) process; video therefore exudes a kind of point-and-shoot liveness.²

In that sense, video and music share a comparable ephemerality, in that neither can be laboriously ‘deliberated’ over in the way film can be: a video performance must be rehearsed and prepared, like a musical performance, and then whatever plays out in front of the camera is consigned to capture. When tracing the early adoption of video technology in art-making practices, it warrants mention that several notable artists who took to video in its earliest stages were musicians, many of whom had been working in sonic media before turning to video. Nam June Paik began his career as an experimental composer who studied in Germany and encountered Stockhausen and Cage at the 1958 Darmstadt Summer Courses for New Music. Tony Conrad was a member of La Monte Young’s Theatre of Eternal Music. Steina Vasulka was a classical violinist who performed with the Iceland Symphony Orchestra. Video drew creative approaches that were already attuned to temporality, to liveness, and to the phenomenon of sound.

Steina Vasulka, to investigate one artist in particular, comments that “for me, it is the sound that leads me into the image. Every image has its own sound and

² Digital video has, of course, drastically overwitten that artefact of video technology; on which more in Chapter Three.
in it I attempt to capture something flowing and living. I apply the same principle to art as to playing the violin: with the same attitude of continuous practice, the same concept of composition” (Rogers 28). Vasulka’s video work Violin Power (1970-78) takes explicitly as its subject the act of violin-playing, and the capturing on video of instrumental soundwaves and their generative gestures. Video footage of Vasulka playing the violin is technologically manipulated by the soundwaves which her violin produce: via the Rutt/Etra Scan Processor, the violin’s soundwaves provide a frequency signal which modulates the geometry of the video image, which displays Vasulka performing those very sounds. As the frequency signal spreads and distorts the video’s scan lines, the straight line of Vasulka’s moving violin bow appears to wobble—at frequencies relative to the pitch she is playing—such that the once-straight bow in fact looks like a soundwave itself. Insofar as this method of video processing is executed live, concomitant with the recording of the video, in Violin Power Vasulka is quite literally ‘playing’ the video with her violin, and in doing so, inflecting her own performing image with a visualization of the sounds she produces.

Not only did early video works exploit and elucidate video technology’s inherent intermediality; the presentation of early video works in galleries and art spaces also tended toward intermedial installations, where the arrangement and spatialization of video screens in the gallery was as defining an element of the work as the videotape itself. In contexts such as these, the videotape serves as a kind of ‘score’: it is the work’s fixed, determined component, which then
underscores and influences the undetermined, real-time spatial and social relations which play out as viewers experience the work in time and space. The videotape and video screen function, again, not as ‘works’ within a medium, but rather as nodes within a micro-culture of intermedial interaction, supplying determinate prompts which catalyse and collide with their necessarily live and indeterminate presentation.

Nam June Paik’s iconic performance-piece *TV Cello* (1971), performed by Charlotte Moorman, aptly illustrates these kinds of circumstantial relationships. The ‘score’ for *TV Cello*—in other words, the constraints defining its performance—is the sculpture of three television screens, vertically stacked to resemble the body of a cello and rigged up with a cello bridge and playable strings. The ‘score’ also designates the video content that plays on each of those screens: one fixed collage of other cellists playing cello, plus more indeterminate content—a live television broadcast, and a closed-circuit feed of Moorman playing the TV-cello in real-time. Moorman, as she performs, interacts with this ‘score’—these circumstantial constrains—in such a way that it influences her musical performance on the instrument (the TV-cello), which in this case is itself also ‘the score’. The closed-circuit video feed of Moorman is not only visual content in the performance, but also a microcosm of the entire intermedial network that binds the work in multiple feedback loops of interaction and influence.

Early video art, as these works exemplify, concerned itself with the fusion of sound, action, and moving image, where the screen was the site of their
coalescence. Joan Jonas’ video piece *Duet* (1972) presents another striking illustration of the way in which early video pieces posited the screen as a live interface between sound and image. In *Duet*, Jonas performs a kind of vocal duet with herself – or rather, she performs a duet with her own recorded video image that appears on the video monitor in front of her. Jonas mirrors and echoes her screenal simulacrum by imitating the vocal sounds her video-self makes: howling, plangent cries, almost like a lonely dog. The pre-recorded video of Jonas therefore serves as a kind of ‘score’ – a fixed document which prompts and influences the responses of the live Jonas. The fixed, pre-recorded tape of the ‘video score’ is in this instance imbued with a sense of liveness by virtue of its function as prompt for live re-impersonation.

Crucially, the dynamic of the ‘duet’ between the two Jonases is cemented via sound. Recognizing the same figure on both sides of the screen suggests a certain degree of connection, but hearing the two voices fuse sonically is a far more convincing token of their live duetting interaction. The figures are separated by their differing materiality (photons vs. flesh) and their differing temporal origins (past vs. present), but when their voices are heard simultaneously, in duet, that distance collapses into the ephemeral moment of sonic perception. While the video-Jonas’ image is bound to a quadrilateral screen, her voice can escape the screen and inhabit the live space; we *hear* the voice of each figure coming from the same source, inhabiting the same space, as if both figures were, in fact, comingling live. At the moment the live Jonas stops responding vocally to the video, the bond
between the two figures abruptly drops off as we are reminded of the distance between screen and body. Sound—especially, in this instance, vocal sound recognisably produced by the actions of fleshy human bodies—springs out of the screen’s surface to fuse the video with the real.

From its first adoption as a technology in art-making, video was a site which explored the entangled creation and reception of sound and image. The screen, therefore, under the influence of video, became a surface from which one not only expected to receive images but also sounds. Although of course, in practice, sound emanates not from the screen itself but from speakers elsewhere, the video screen was nevertheless embedded with sonic expectancy. While cinema began life as silent movies on a silent screen, the idea of a silent screen has been antithetical to the medium and culture of video from its genesis. In light of this, it is necessary to reiterate that a video score does not invite the performer to ‘sonify’ it — to create the sounds they imagine the moving image generating, or to accompany the image in a manner analogous to movie-hall pianists providing musical accompaniment to early silent films. The interpretation of a video score therefore does not follow in the cinematic tradition of retroactively supplementing a moving image with ancillary sound. Instead, by compelling the performer to re-enact the action in the video image in such a way that sound is produced, the dynamics of the video score firmly inhabit video culture, wherein sound and image are apprehended synchronously and inseparably.
In a related vein, it is also noteworthy that so much early video art dwelt on the figure and behaviour of human bodies. The body in video artworks is presented as being capable not only of material agency but also sonic agency; the body’s interact with itself and its environment is evidenced by not only visibly but also audibly. The video score, by prescribing action in the pursuit of generating sound, likewise frames the performing body as a subject whose physical and sonic agencies are mutually contingent.

1.3 SEMIOTICS OF THE SCREEN

A synopsis, to open this section, of an illustrative video piece: Vito Acconci’s *Centers* (1971). Acconci faces the camera squarely, the shot framed tightly around his face. And for the video’s full 22 minutes, he points his right index finger dead straight at the camera. His arm is outstretched with an uncomfortably cramped rigidity underneath his chin such that his clenched fist, with its single protruding finger, almost obscures his face. He tires visibly as the minutes wear on but, though his finger falters, it never gives up pointing at the camera – and, therefore, *through* the camera, through the screen, right into the eyes of the viewer. Acconci’s pained squirming and shuffling reinforces the fact that he is not a frozen, statuesque figure—not a *tableau vivant* of Uncle Sam—but rather a live actor consciously performing a continuously-renewing action. He points not at
one moment in protracted time, but at 29.97 moments every second – one for each frame of the NTSC video format.

Although Acconci performed his pointing gesture 45 years ago, the gesture carries the invitation to be read as pointing explicitly at me—or you, or whomever views it—today, in the very moment I encounter the video. For that is the semiotic function of pointing: it denotes conscious attention directed not just at a particular thing but in a particular moment. (Look! Now! Before it disappears!) And yet Acconci’s pointing finger is a tautological gesture in the context of video. By compulsively summoning me, Acconci’s pointing directs my attention to the screen. But my attention has already been focussed there by another thing which ‘points’: the camera. Not for nothing does the same verb describe the orientation of the camera lens, ‘pointing’ attention at its subject. Indeed, Acconci is pointing neither at me, nor the screen; he’s really pointing at the camera, which points back at him, locking them in a kind of indexical Mexican standoff.

Acconci’s fingertip therefore emblemsizes a fundamental property of video: the way in which video, as Anne Wagner describes, “summons you into the present moment” (Wagner 69). An implicit pointing finger is embedded in the dynamics of the video screen’s interface. Whenever I apprehend a pointing finger, it directs my attention to an object now, regardless of the temporal origin of the object or even of the pointing gesture. Insofar as the video camera ‘points’, the subject the camera captures will therefore always be an object of my present attention, despite the fact that videotape always contains bygone events. (Live video feeds, of course, evoke
even more immediacy.) Whether or not Acconci’s finger is in the frame, the video image is always pointing, summoning us into a relentless present.

The pointing finger is one of Peirce’s archetypal examples of an indexical sign, “because it is in dynamical (including spatial) connection both with the individual object… and with the senses or memory of the person for whom it serves as a sign” (Peirce 107). Peirce goes on to claim that “anything which focuses the attention is an index… [indexes] direct the attention to their objects by blind compulsion” (108). Insofar as the video image implies a pointing finger summoning our attention, then the video image is an implicitly indexical sign. As Peirce mandates of indices, video calls upon its witness to “use his powers of observation, and so establish a real connection between his mind and the object” (110).

The Peircean index is sometimes defined—with erroneous narrowness—as a sign which bears material vestiges of its object’s physical influence: a footprint in the sand, for instance. Analogue visual media satisfy this definition of the index: exposing light onto photographic plates or film stock generates an image that is a kind of optical ‘footprint’ of the scene it captures. Several new media theorists thereby make the case that digital technology has brought about the death of the index, because the technical anatomy of digital media’s signs depend on electronic impulses rather than the physical influence of, say, lightwaves on photographic negatives. Lev Manovich posits that, with the advent of digitization, film is “no
longer an indexical media technology, but, rather, a subgenre of painting” that fabricates an illusion of its subject (Manovich 295).

Yet Peirce’s repeated emphasis on the index’s necessary “connection” by way of “attention” between its viewer and its object suggests that the defining condition of the index is not the material influence of its object, but the index’s co-existence in time with both its object and its apprehension. Peirce explains the index is a “sign of its object by virtue of being connected with it as a matter of fact”, not necessarily as a matter of material matter. Kris Paulsen argues that indices are therefore “indicators of the present, not the past”; the meaningful interpretation of an index is parasitically dependent on the index’s present—i.e., temporal—context (Paulsen 46). Where a photograph, or film, indicates the material traces of lightwaves long-past, the electromagnetic (or digital) anatomy of video allows it, with every re-play, to re-conjure stored past events as a sempiternally renewing present.

However, despite the seemingly live presence of video’s apprehension-in-the-moment, Paulsen, in writing about Acconci’s Centers, is quick to stress that this work also exemplifies the way video frustrates the illusion of communion across the interface of the screen; though “Centers calls the viewer present”, it simultaneously “refuses a relationship with her. The viewer sees Acconci, but he cannot see her back despite the overwhelming feeling to the contrary” (Paulsen 77). Acconci appears to point at the viewer, but he also seems to point at the screen: that surface which inhibits any chance of real-time, corporeal interactivity. The screen,
in order to serve as interface between two systems, must also serve as an impermeable boundary between them, defining those systems as incompatible opposites. As Branden Hookway asserts, “the interface comes into being in the maintenance of its contractions” (Hookway 9); or, as Paulsen asserts of the video screen more specifically, “the video screen is a site where things and their opposites touch” (Paulsen 78). It is a characteristic of the index to inhabit both sides of this ontological divide: the index depends on a connection to its object in its present context, and yet in order to maintain its summoning function as an index, it cannot be the object. The indexicality of video images, Paulsen argues, is what affords them “an ontological indeterminacy that enables them to be two things at once – here and there, now and then, generic and specific” (79).

With video thus defined as an indexical medium, the video score, then, comprehensively renovates the semiotics of the musical score. A staff notation score is in essence a system of symbolic signs; as Peirce dictates of symbols, staff notation is connected with its object (the sounding music) “by virtue of the idea of the symbol-using mind, without which no such connection would exist” (Peirce 114). Staff notation comprises agreed-upon visual abstractions which represent to its interpreters specific concepts of sound, be it the concept of a G#, or the concept of mezzo-forte, or the concept of triple meter. Without the interpretation of an educated musical mind who is privy to these concepts and their signifying symbols, staff notation forsakes its semiotic quality.
A video score, because it does not—like staff notation—prescribe the sounding result which a performer should strive to realize, does not depend for its semiotic mechanics on symbols and abstracted representations. The video score is a sign only of itself. The video image’s indexical signs point to the physical actions and gestures of a human figure, compelling the performer-interpreter to apprehend them in the present moment.

1.4 SPLITTING THE SOUND-IMAGE ATOM

Video, Holly Rogers argues, is an inherently audiovisual technology because of its “rare ability to fuse music and image at an intermedial level” (Rogers 2). As previously discussed, video captures sound at the same time it captures images, stores those two kinds of information on the same signal channel, and replays them synchronously. Video therefore carries with it a sonic expectancy—especially the expectation that the video’s sound will in some way be synchronous with the video’s image.

A video score, however, performs a substantial subversion of video’s audiovisual synergy: it removes sound from the equation, leaving behind a silent video image.³ To dissect the sound out of video is drastic surgery; from it emerges an image which appears wounded, amputated, incomplete. Like an amputee might

---
³ Though some of my video scores also employ audio media, I treat the audio material in a way more analogous to cinema than video, in that the audio is conceived, recorded, and edited separately from the video image. The audio material does not comprise the sonic residue of the gesture or action in the video image, nor is the audio score edited to align synchronously with the video score.
report the sensation of a phantom limb where a physical limb once was, watching a silent video—in which the sonic consequences of actions and gestures have been excised from their image—can conjure in the mind of the viewer a phantom soundtrack which completes the incompleteness of the silent moving image. The notational mechanics of the video score, therefore, are driven by the inherent sonic expectancy of the video medium: the silent image compels the generation of sound. Musical scores already bear an ontological imperative to be interpreted; their function as a score insists that it is incomplete until this is so. With a paper score, this imperative is intellectually and culturally embedded into the score-object; but with a video score, the interpretive imperative is embedded into its medium.

It is significant that, in the case of a video score, sound is the element excised and re-constituted, for sound is copiously theorized as a phenomenon which evokes a sense of liveness, of being immersed in the present moment – even if the sounds clearly originate in the past. Frances Dyson, for one, posits that sound “returns the listener to... that feeling of being here now, of experiencing oneself as engulfed, enveloped, absorbed, enmeshed, in short, immersed in an environment. Sound surrounds” (Dyson 4). As previously discussed, the indexical semiotics of a video score serve to summon the score into present apprehension: giving a sense of liveness to a fixed text, and making the pre-recorded past appear present. This semiotic liveness is further compounded by the fact that the medium animating the apprehension of the index is the musicians’ live generation of the ‘live’ phenomenon of sound.
Each of the works in this portfolio, when performed live, presents some kind of foundational audio-visual split. In $8 \times \infty$, the audience witnesses a split between sound and object: the musicians appear onstage without instruments, and perform a sequence of silent gestures in synchronisation with an electronic audio track comprising the pre-recorded sounds of those gestures, performed on musical instruments and quotidian objects. Sound is therefore split from the visible evidence of the sound’s originating gestures, both across time (past vs. present) and medium (fleshy vs. digital). In XEROX ROCK, the audience witnesses a split between the silent actions of the figures in the video score projected onstage, and the synchronously mirrored, sound-generating actions of the live performers. The audience’s visual attention is thus split between two plausible gestural origins of the sound they hear. In rupture | rapture, the audience witnesses a split of scale: the audience sees a live video feed of the percussionist’s hands projected onto the screen onstage, blown up to appear monstrously huge, while at the same time they hear the closely-amplified, delicate sounds made by the vocalist. Via the technological mediation of video projection and amplification, the smallness (both visual and sonic) of the live performers’ actual behaviour is no longer commensurate with the sum total of affective content unfolding onstage.

None of these pieces provides an identifiable locus where sound and image (whether live or mediated) singularly meet: the pieces all present a split, a wound, an incision in the sound-image event. In a traditional live solo piano performance, for example, there is a visible site of audio-visual connection: the instrument of the
piano, which is activated by the live performer in such a way that the sound heard is directly commensurable with the gesture seen. These video score performances, on the other hand, by their excision of a precise sound-image locus, necessitate the formulation of a new equation for perceiving sound alongside image. The wound needs stitching up; the incision needs mending.

Though one could argue various ways in which this split is negotiated in these pieces, I present here the one I take to be most significant. By excising the locus where one single sound meets its one single originating gesture, these pieces enable a profusion of possible mappings of sound to image. Sound and image do not balance out either side of a one-to-one equation, but rather bounce around a combinatorial matrix where initiating gestures and sonic results can be rearranged and interchanged. This is especially the case in the pieces $8 \times \infty$ and XEROX ROCK, where the audio track comprises several superimposed recorded takes of the performers’ interpretation of the video score. These pieces present a surfeit of sonic ‘solutions’ to the gestures the audience observes, any one or combination of which plausibly fulfil the sonic expectations of the visible gesture. XEROX ROCK also presents a surfeit of visual ‘solutions’, by making the digital figures of the video score visible onstage alongside the live performers. A performance of XEROX ROCK can be deemed most successful when the live musicians’ sonic rendition of the score is so faithful, and closely aligned, that the sounds they produce convincingly serve the illusion that they are also the soundtrack to the silent video score. In the case of rupture | rapture, though the work presents only...
the monophonic through-line of live sound, without any pre-recorded alternatives, the discrepancy of scale between the performers’ visible actions and the mediated result opens up a kind of sliding-scale continuum, along whose points the piece is continually being mapped and re-mapped.

In this way, these pieces perform the notational strategy of the video score itself: a video score presents to its interpreter one single sequence of action which—because it does not signify sound in the way staff notation does—could be animated by an infinite number of sonic results. There’s a streak of the virtual sensibility here: exploding the one-to-one sound-image equation into a multi-dimensional matrix ejects the combinatory possibilities out of the constraints of physical, material properties, and into the immaterial potentialities of the perceptual. And ushering in a virtual dynamic ushers in too the dynamic of interactivity. Granted, the video scores are fixed media; their content and course cannot be altered by the influence of their interpreters in the same way a player can steer the narrative of a video game, or a computer user interacts with GUI objects in order to execute operations. In these examples, the evidence of interactivity is visible; in the case of a video score, the consequences of interaction are heard. The video score image serves as the interactive interface for steering the live navigation of sonic possibility – bounded by the programmed constraints of its interfacing system, but permissive of user creativity.

Lev Manovich argues against the fallacy of privileging literal, material interfacing when considering the nature of interactivity:
When we use the concept of ‘interactive media’ exclusively in relation to computer-based media, there is the danger that we will interpret ‘interaction’ literally, equating it with physical interaction between a user and a media object...at the expense of psychological interaction. The psychological process of filling-in, hypothesis formation, recall and identification, which are required for us to comprehend a text or image at all, are mistakenly defined with an objectively existing structure of interactive links (Manovich 57).

A video score, by staking out a sound-image combinatorial matrix, calls for precisely this “psychological process of filling-in”, not only from its interpreter but also its audience. And, after all, a video score demands from its performers not only psychological interaction, but also their physical interaction: the score continuously leads a performer through time, animating their movements. Because the score prompts the execution of action, the process of comprehending and “filling-in” the video score necessitates a somatic heuristic, by which those prompts are tried out and mapped onto the performer’s body.

Finally, it must be noted that an interactive dynamic is not unique to video scores. By Manovich’s formulation, of course, all musical scores are interactive interfaces: their comprehension requires of their performers the psychological processes of recall, identification, and “filling-in” – in other words, interpretation. As discussed early on in this chapter, an interface preserves its functionality via its non-functionality: and a musical score’s interactivity hinges on this characteristic of the interface. Every score, to varying degrees, fails to exhaustively articulate each and every parameter of a musical event; every score is an incomplete representation of the event it prompts. This incompleteness—this non-
functionality—is what affords a musical score its interpretive interactivity. A musical score’s interactivity picks up where the limits of its medium leave off.
CHAPTER 2

rupture | rapture & 8 x ∞ : THE MIRROR SCORE

The video score depends on one central dynamic process: mirroring. The video score notates the physical actions which the performer is to execute; what’s more, it notates these actions via their literal demonstration by human figures – not by abstracted representation, as in the case of tablature. To interpret and perform the score, the performer is therefore required to mirror these actions. The text instructions accompanying the scores explicitly invoke the concept of mirroring: the performers are prompted to “mirror the movement”, or “mirror the actions” of the moving figures in the video score. The performer’s act of interpretation is grounded in the act of mimesis, as they re-enact—in real time—the movements they see in the score. In other words, the interpretation of a video score is contingent on negotiating the nature of mimesis: what exactly the task of ‘mirroring’ demands, and to what degree of exactitude.

The video score presents a visual invocation of the mirror, too: the footage in the score has been ‘mirrored’—flipped around a vertical axis—so that, for instance, whenever the figure in the score moves their hand on the left-hand side of the frame, the performer moves their own left hand. The score for 8 x ∞, a piece for mixed octet (flute, bass clarinet, piano, 2 percussion, violin, cello, and double
bass), often features footage of a figure handling musical instruments; flipping the footage therefore makes the instruments appear ‘the wrong way round’, thereby visually differentiating the video score from, say, a video of an instrumental performer in recital. There is a similar visible ‘strangeness’ in the video score for *rupture | rapture* (voice & percussion): the percussionist’s video score demonstrates actions to be performed on a djembe drum, but these actions are (for the most part) performed not on a drum but on the surface of a wooden table top, and bear negligible resemblance to regular drumming technique. In short, the footage in the video scores for both pieces is embedded with a substantive visual departure from an experience of watching people play musical instruments.

In this way, the performer’s task of ‘mirroring’ the video score engages a different mimetic dynamic from imitative strategies of instrumental pedagogy. In the case of, for example, a violin lesson, a teacher might demonstrate a bow hold or a fingerboard position on their own instrument, exhorting the student to imitate their actions. Because the teacher’s physical alignment—as they play the violin ‘normally’—is directionally opposite to the student’s, this kind of mimesis requires of the student some prior embodied knowledge of playing the violin in order to cognitively translate the action they see into a workable violin technique. An absolute beginner violin student, with no embodied knowledge of what playing the violin ‘feels like’, would have greater difficulty reverse-mapping the teacher’s actions onto their own body. (A more effective pedagogical strategy with absolute beginner students might be for the teacher to physically shape their student’s hands
and arms into a technically ‘correct’ position.) Imitating the observed technique and actions of another musician is essentially a process of filtering what is observed through one’s own already-structured, already-embodied technical knowledge – and then synthesizing the two.

The video score, on the other hand, by presenting demonstrational footage that can be directly mirrored, expects of the performer no prior embodied knowledge of instrumental technique, and instead establishes a mimetic dynamic based on immediate visual observation. In fact, the video score deliberately alienates the embodied technical expertise of experienced instrumentalists with its strange, ‘wrong way round’ appearance. Instead of calling on the performer’s embodied knowledge of instrumental technique to interpret the actions in the score, the mirror-score privileges a visual engagement with the instrument: prompting in the performer a kind of out-of-body experience as they are required to map their own body onto the shapes and movements they observe on the screen.

This notational strategy of alienation attempts to minimize the influence of the performers’ already-established technique on the mimetic exercise, in order to encourage sonic results that depart from the conventional sounds of ‘good instrumental performance’. Altering the manner of playing an instrument inevitably alters the sound it produces, and so the video score seeks to wean its performers off their habitual manner of playing. After all, the performers’ explicit focus is to re-enact the figures’ movement, not to re-create the sounds they speculate the figures in the score are making with their movements. The visually-privileging, mirrored
video score reinforces this focus. Interpreting the score requires some knowledge base; the alienating mirror-score diverts the performers’ knowledge base away from learned instrumental technique—structured around the pursuit of certain instrumental timbres and sonic qualities—and towards immediate visual observation, opening up an expanded (and sometimes surprising) musical palette.

By no means does this renunciation of acquired technique diminish the centrality of the instrument itself as a technical object. In both 8 x ∞ and rupture | rapture, the musical instrument—be it a drum or a flute—is the axis of symmetry grounding the exercise of mimesis; it is the *sine qua non* of the performers’ actions. In other words, the instrument serves as a material interface: a site of tactile interaction where virtual, disembodied information is apprehended, negotiated, and physically manifested into effectual, ‘real-world’ actions. The instrument becomes not only an external frame of reference, setting limits and bounding parameters within which the performer’s actions are executed, but also a knowledge base itself, as the object upon which the performer works out and embeds their interpretation of the score. The instrument serves as a mirror, where one digital body is reflected onto another fleshy body.

Not only musical instruments serve as interfaces in 8 x ∞. The percussionists, rather than playing purpose-built percussion instruments, utilize a collection of quotidian noise-making objects: newspaper, a handheld radio, a dog whistle, a bird call, an abacus, a computer keyboard, two quarters, and two pieces of sidewalk chalk. The common denominator of these objects is that they are all
tools for the transmission of information and data, be it language, currency, numerical values, indices, or symbolic commands. However, when employed as sound-making instruments in the context of $8 \times \infty$, these objects are detached from their ordinary functions, especially as the gestures performed with/on them are not necessarily the gestures required for their ordinary functional operation. (For instance, the computer keyboard is not only typed on, but also brushed and swept with the hands.) Like the musical instruments, these objects are repurposed as a physical, anchoring interface between action and sound. They are employed for their sonic properties, as a tool for aurally articulating the contours and dynamics of the performers’ gestures.

Repurposing quotidian objects as interfaces for present action is a common strategy in early video-art pieces. For example, Gerald Byerley’s *Tape II* uses a roll of sellotape not for its ordinary function as an adhesive, but rather as a marker of specific time and place. Long, straight lines of tape are laid out on the ground on various domestic surfaces—linoleum floors, bathroom tiles, windows, and stone steps—and slowly but methodically pulled taut until it peels off the surface. The tape, in essence, traces a straight line across these everyday surfaces, drawing attention to one discrete point at a time, in a continuous indexical present tense. What is especially notable in *Tape II* is the distinctive ripping sound of the tape as it peels off; different material surfaces produce different sounds. The liveness of presence—of the particular but ever-changing spot on the floor which the tape indicates—is heightened by the sonic immediacy of the sellotape in motion. The
quotidian objects in $8 \times \infty$ function comparably to the tape in Byerley’s *Tape II*: they visually and sonically mark the present tense of the performer’s actions in time, in space, and on materials.

This action/sound interface is broken, however, in live performance, when the musicians appear onstage without instruments and yet perform—silently—the same gestures which generated the sound in the audio track. It was remarkable, in the live performance, to observe the musicians’ physical and psychological discomfort when performing familiar actions in a newly unfamiliar context. Their comportment was significantly altered by the absence of the object; even instrumental actions—presumably ‘second nature’ to experienced professional musicians—appeared awkward and ungainly when executed without the aid of the instrument as a physical frame of reference. The performers’ compulsion to act silently also demonstrably impacted their quality of movement; their tendency to gesticulate histrionically or exaggeratedly was perhaps an attempt to over-compensate for the strangeness of the gestures’ ensuing silence. The absence of the physical object came across as so strange and alienating that, in a sense, it intensified the dependency of the performers’ actions on a sounding physical interface. Paradoxically, even in its absence, the instrument (or repurposed instrument) as object-interface is still the *sine qua non* of the performer’s actions.

Both *rupture* $|$ *rapture* and $8 \times \infty$ also employ audio scores, interpreted independently from the video score. In *rupture* $|$ *rapture*, the vocalist’s part is exclusively an audio score; in $8 \times \infty$, video score episodes alternate with audio
score episodes in all eight parts. The interpretive exercise of the audio score is, like the video score, mimetic; however, rather than the *ocular*-centric dynamic of the video score, the audio score demands meticulous *aural* attention. The performers are therefore instructed not to ‘mirror’, but rather to ‘translate’ or to ‘recreate’. While the video score features actions performed on musical instruments—and therefore directly analogous to the performers’ execution of those actions on their own musical instruments—the audio score features sonic material originating from sources far removed from the sonic forces that will ‘translate’ them. In the case of *rupture | rapture*, for instance, the vocalist’s audio score uses source material of oceans, glaciers, stars and satellites; in *8 x ∞*, the instrumentalists are tasked with re-creating the sounds of computer hard drives, Morse code, the echolocation of bats, and dial-up modems.

Just as the mirrored video score attempts to visually distract the performer from learned instrumental technique and its attendant soundworlds, the audio score’s far-removed originating sources also seeks to tug the performers into the furthest sonic reaches of their instrument’s capabilities. When responding to the audio score, a performer operates less in reference to a stored sonic knowledge base of their own instrument, and is instead enjoined to make their instrument sound like something other than itself.

The audio score’s accompanying instruction of ‘translation’ or ‘recreation’—rather than ‘mirroring’—encapsulates an important concession: that the exercise of mimesis is by definition impossible, and necessitates negotiation. It
is patently clear that a human voice cannot exactly imitate the sound of a comet, and that the task of interpreting the audio score is therefore one of approximation. In the same way, one musician’s exact physical imitation of another musician’s digital bodily simulacra is also impossible, insofar as no two human bodies are alike. The video scores’ text instructions articulate this admission of imperfection, encouraging the performers to mirror the score ‘as closely as possible’, ‘as meticulously as possible’, or ‘as faithfully as possible’. The score bounds an asymptote towards which a performer’s interpretation must aim, while at the same time acknowledging the inherent unreachability of that aim.

The mirror-score, therefore, does not call upon a mimetic dynamic of re-enactment: that is, a concerted attempt to re-create the contextual conditions of the ‘original’ being mirrored, in the same way that ‘re-enactments’ of historical events purport to exhibit the greatest possible authenticity and verisimilitude. Were that so, the score would not compel interpretation or performance: because the most faithful re-enactment would be to simply let the score play out, existing as the performance itself. The mirror-score instead demands a mimetic dynamic of re-performance: adapting the essential substance of the original score to new and varying contextual conditions – different conditions of time, place, instrumental forces, human bodies, technical expertise, the witnessing audience, and the aesthetic valences of the ‘new music’ concert hall – or whatever the aesthetic predilections of the performers undertaking the score.
The mirror-score’s mimetic dynamic, therefore, is fundamentally that of adaptation. That the mirror-score’s chief imperative is the generation of live sound—set out in the text instructions by the condition that the performers act ‘in such a way that sound is produced on their instrument’—further reinforces the mimetic process as an adaptive re-performance. As discussed in Chapter One, the video score is startlingly silent, excised from the sounding consequences of the actions in the footage. A live musician’s performance of the score, by re-constituting the sound of the score’s actions, catapults those actions from literal mimesis into a qualitatively different context.

The expectation that every performance of the score is an adaptive re-performance, reflective of its unique context, is embedded into the performance procedure for $8 \times \infty$. In this piece, the musicians do not perform live instrumental sound in concert; instead, they record eight different performances of the video score, which are compiled into an 8-channel audio track. The audio track playback then constitutes the sonic material ‘performed’ in the live concert event. In concert, the musicians do not perform live sound: they appear onstage without their instruments, and perform the actions in the video score silently, synchronized with their corresponding sounds in the pre-recorded audio track. The audio track which accompanies them is therefore configured as a log of attempts, a kind of database bounding the score’s interpretive standard deviation. The averaging, or analysing, of ‘trends’ in this collection of sonic data is achieved by the continuity of the single, present fleshy figure onstage demonstrating the physical conditions which
generated the collection of sounds heard by the audience. Crucially, the densely textural sonic profile of the audio track is dependent on the expectation that all eight superimposed takes will be different – that they will not be identical, literal ‘mirrorings’ of the score. Rather, the broader the bounds of each adaptation or ‘translation’ of the score, the deeper the audio track’s textural density.

Hillary Young, the vocalist for whom *rupture | rapture* was composed, and who premiered the work, communicated to me in personal e-mail correspondence an illuminating perspective on the adaptive ‘translation’ she undertook in interpreting the score. Young writes:

There is something both more intimate and more urgent about translating... similar to playing ‘telephone’, my task was paradoxical; I knew immediately what the piece was supposed to sound like, but it was actually impossible for me to replicate that, and therefore my assumed knowledge of what the piece would sound like was actually false. The message I was passing along was automatically going to be at least slightly altered.

Young compared her interpretive task of ‘translation’ with linguistic translation, noting that translation necessarily ends up “obscuring the actual meaning of something”. This, of course, is an issue of frequent concern to singers, who in most of their work deal with language and the duty to render that language intelligibly. Young took the approach of “translating the nuance” of the audio score, rather than translating the “actual meaning of something”. Rather than attempting to re-create the actual non-vocal sounds in the audio score, her interpretation privileged an articulation of the audio score’s changing contours of colour, dynamic, frequency, density, and texture.
Young notes how labour-intensive this process was, comparing it to “a language I had to learn to speak through immersion. And immersion is truly how I learned it; there were weeks when I would listen to the score at least ten times a day.” The immaterial nature of the audio score also presented a significant challenge to her as a performer; it “took away the score as an object outside of myself, so that I was forced to build an object internally from a combination of memory, repetition and learning by rote. Initially this was pretty disturbing to me as a performer; I had to rely on something that I couldn’t see or touch.” Ultimately, Young reflects that this forced me to have a much more internal understanding of the score.”

Young also remarked on the task of vocal mimesis in relation to vocal pedagogy, noting that “as a singer, mimesis is a necessary but dangerous part of what we do”. In lessons, a singer is often exhorted to imitate the sound of their teacher’s voice, which necessitates the student to adapt a healthy vocal technique that will allow their own body and instrument to safely reproduce that sound. There is therefore a “fine line” between mimicking sound and mimicking technique: the former can be dangerous if not suitably adapted to the student’s own instrument. Young described how “the audio-video score liberated me” from this concern: “I was never asked to mimic a voice, and that way, the sounds I was making were authentically mine”.

It must be noted that the video score invokes a discussion about the ethics of notation and performance, in light of the video score’s direct and deliberate
control of a performer’s body. Of course, all musical scores—staff notation included—impress themselves onto the performing body by summoning specific actions from that body. Even where a score notates an abstracted representation or suggestion of the sound to be produced, the score still coerces its performers into certain physical and psychological states. The video score, however, by somewhat presumptuously taking over the performer’s bodily agency, could be said to enact this coercion to a more violent extent. In prescribing tasks and actions to be executed, the video score renders its performers as labourers, treating them as bodies whose physical labour is micro-managed in the interests of producing aesthetically ‘profitable’ results.

Perhaps, however, it can also be said that in making so transparent the vectors of power and control structuring notation and performance, the video score is a more honest notational strategy. A staff-notation score, though it ultimately exercises physical and psychological control over a performer, masks the extent of its influence via the proxy of abstracted notation; a performer appears to maintain a certain autonomy via the intellectual agency which they must engage in order to interpret this abstraction. A video score, on the other hand, though it confines the performer’s somatic behaviour within clear limitations, makes plain the extent of its physical control over a performer. Crucially, it maintains the performer’s prerogative to adapt the material in the score into new contexts, around new instruments, and onto new bodies.
CHAPTER 3

XEROX ROCK: YOUTUBE PEDAGODY & DATABASE AESTHETICS

XEROX ROCK—a video score for electric guitar, piano, and marimba—warrants a separate chapter in a discussion of video score mechanics, because the notational strategy of this score differs from both rupture \( \mid \) rapture and \( 8 \times \infty \) in two crucial ways. First, instead of remaining hidden from the audience’s view, in a live performance the video score is made visible to the audience, by way of a large projection screen behind the players. Second, the video score comprises footage not purpose-filmed by the composer but rather gleaned from YouTube, the largest online database of freely accessible videos. The video score footage features people playing either guitar, piano, or marimba/xylophone, in both performance and pedagogical contexts. This footage is edited into three independent video ‘parts’—one for each player in the trio—using the relevant instrumental footage for each instrumental part (e.g. the pianist’s part contains footage of people playing piano). All video parts are presented polyphonically in one video score by way of a two-by-two split screen: the guitar part is positioned top-left, the piano bottom-left, and the marimba top-right. As this split screen matrix yields four video panels, but
there are only three video parts, the bottom-right split-screen panel remains blank in the score.

3.1 COMPOSING THE DATABASE

The performers are required to make audio recordings of eight full run-throughs of the score in rehearsal, ideally at times interspersed throughout the rehearsal process – i.e. not eight takes in one session, but a few takes per rehearsal, with some spot-rehearsing in between. In this way, the recordings serve as a longitudinal document databasing the rehearsal process. Each of the eight audio takes are then superimposed to create a single polyphonic stereo audio track which accompanies the trio in live performance. The eight rehearsal takes are precisely aligned at the very beginning of the score according to placemarking ‘pips’ (high-pitched beeps), but are otherwise not altered in the editing suite to perfect their rhythmic synchronicity. Only gain levels and panning are adjusted; the older the takes, the lower they are mixed in the balance, and the further from the centre they are panned. The most recent take is mixed loudest, and is panned centre. The performers are also required to make a video recording of the final rehearsal run-through. This dress rehearsal video is inserted into the lower-right panel of the video score’s split screen matrix (which was previously blank). The new four-panel video is the score followed by the performers and displayed to the audience in live performance.
The score-reading instruction given to the musicians is to “mirror the movements of the human figures in the video in such a way that sound is produced on their instruments”. As with the other video scores in this portfolio, the performers’ attention is therefore focussed on re-enacting the movement demonstrated in the video score – not reverse-engineering the sounds which they expect the figures in the video footage to be making. To that end, the performers are encouraged to consider only information contained within the score to inform their interpretive decisions. For example, at 01:01 in the piano part, it is quite possible to recognize that the pianist in the video score is playing the exposition of Tchaikovsky’s Piano Concerto No. 1. But the angle of the shot is such that the exact placement of the player’s hands on the keyboard is visually obscured: only a performer’s prior knowledge of the Tchaikovsky Concerto could inform them that the chord being played is a D flat major chord. The performer therefore ought not play a D flat major chord at this moment, because that pitch information is not contained in the score itself; rather, the pitch at this point in the score is indeterminate. However, at a moment such as 03:25 in the piano part, the camera angle gives a bird’s-eye view of the keyboard, thus the video score clearly demonstrates to play a single note C. However, because the full compass of the keyboard is not in the shot, it is not clear exactly which C on the piano keyboard is being played. At this point in the score, while pitch is therefore determinate, register is indeterminate.
These distinctions are audibly staged when the eight rehearsal takes are superimposed into the accompanying stereo track: in the case of the (not-)
Tchaikovsky at 01:01, the amalgamated sonic result is a dense cluster chord whose attack points are temporally aligned and uniformly strong, but whose pitches display no statistical trends across the eight takes, and are therefore essentially random. Conversely, the case of the single note C at 03:25 generates an unusual moment of pitch unison in the audio track, although indeed the pianist plays two different Cs during the course of the eight takes.

Such examples illustrate how the video score for *XEROX ROCK* is deliberately engineered to shift sonic textures through varying depths of focus: the score manipulates the degree of precision with which various musical dimensions are prescribed and therefore realized. *XEROX ROCK* systematically works with and against the video score’s inherent limitations *vis-à-vis* which musical dimensions can be most accurately notated. In the case of pitch, while the video score’s ability to prescribe exact pitches is limited and contingent on other cooperating dimensions (e.g. that the rhythm is slow enough to make out where hands and mallets are placed), the ‘depth’—or compass—of pitch, as fleshed out by several layered rehearsal takes, *can* be controlled. Managing the cropping, angle, and playback speed of the video footage makes available to the composer a sliding scale of notational precision, and therefore a sliding scale between broadly splayed clusters of notes and perfect unison once several renditions are multi-tracked. Within the domain of rhythm, a video score can accurately prescribe well-prepared
single attack points, as well as sequences of predictably periodic attacks. Such musical figures are therefore realized with a greater clarity of rhythmic focus in the multi-layered audio track, creating a heterophonic vertical texture. On the other hand, rapid figures of unpredictable rhythms cannot be re-enacted as accurately by the score’s interpreter, and so the multi-layered audio track renders these passages as messily polyphonic textures.

This strategy of employing sliding scales of musical resolution also takes into account the physical mechanics of each instrument, and the degree to which they exhibit a visible relationship between playing technique and resultant sound. For instance, because pitch positions are arranged much closer together on a guitar fretboard than they are on a piano keyboard, the ability to visually demonstrate precise single pitches is more difficult on a guitar than on a piano – a difficulty compounded by the need to triangulate the guitarist’s left hand’s fretboard position with which string(s) their right hand is picking. However, the immediate visual distinctiveness of different chord shapes on a guitar makes tonal harmony much easier to visually represent on a guitar than on a piano. While a certain camera angle could make the distinction on a piano keyboard between D flat, E flat, and A flat major quite ambiguous (not to mention F sharp and G sharp minor), on a guitar, the differences in hand shape between the open chords of C, D, E, G, and A majors are fairly unequivocal, even from several rows back at a rock concert.

The unique technical mechanics of each instrument therefore generate distinct pitch profiles when the eight rehearsal takes of the score are superimposed:
the piano tends towards a greater focus of pitch in single notes and melodic lines
but clustery chords, while the guitar tends towards discernible chords but
chaotically tangled single-note lines. The marimba, with its monochromatic and
relatively level keyboard layout, is the most ambiguous instrument of the three on
which to visually decipher pitch, and so the marimba features fairly consistently in
the amalgamated audio track as tight pitch clusters. The marimba’s notable
consistency is in dynamics: membranophones’ inherent what-you-see-is-what-you-
get method of sound production makes the relationship between the speed and
quality of the physical attack, and the sounding result, very predictable and
uniform.

I think of this notational method ‘as three-dimensional composition’. Sonic
lines and textures are composed not only in relation to their passage through time,
but also in anticipation of their compounded textures when the score’s various
game-plays are compiled. I also think of this method as ‘statistical score-making’:
instead of representing a single, ‘ideal’ data-point of sound, the score’s notation
acknowledges that it will generate various interpretations, and so strategizes a way
to weight the probable tendencies of how the score will be realized. At several
points in XEROX ROCK the score is deliberately engineered to encourage the
generation of anomalous data points. For instance, at 05:23 the score in all three
parts abruptly freezes after a prolonged passage of rapid finger-work. As there is no
visual cue in the score to pre-empt this abrupt freeze-frame, the performer will
inevitably over-shoot and fail to stop on a dime at the precise moment the video
footage freezes. Despite its appearance, the notation at this point does not, therefore, strategize to compose an abrupt silence, but rather to compose a frayed edge of sound that peters out unpredictably.

The in-built ‘shortcomings’ of a video score—its arguable failure as an exact, precise notational method—allows instead the composition of a database, and the manipulation of that database’s statistical tendencies. When several recorded run-throughs of the score are superimposed and heard at once, they embed into the score’s performance the entire gamut of that score’s interpretive possibility— the whole bell curve of its possible, probable, and certain sonic realizations.

3.2 DIEMBODIED PEDAGOGY & THE SCREEN

It is significant that the primary source footage for the score of XEROX ROCK is a slimly representative snapshot of that redoubtable pedagogical database: YouTube. All three video scores in this composition folio operate by mobilizing the digital screen as a site for transmitting kinaesthetic information, and are therefore pedagogical resources themselves. Operating according to the principle of mirroring—the very first pedagogical method an infant responds to as they mimic the sounds of their mother’s voice—the screenal dynamic of these video scores is an imperative screenality: their demonstrative visual syntax suggests the a pedagogical function, and therefore commands re-enactment. Consumers of moving images and video media develop fluency in parsing the visual conventions
of different genres and functions of moving images. Even if distinguishing markers of content were removed, a television news broadcast is distinct from a sitcom or an aerobics instruction video or a home movie by virtue of its visual syntax.

Genres of moving image media are also usually distinguished by their mode of viewing and distribution—on an airport television screen versus a commercial DVD versus a home television set. Yet YouTube indiscriminately distributes a vast variety of content within exactly the same online platform and interface, thereby conditioning its users’ fluency in distinguishing genre and function of moving images from the image’s visual syntax alone. Indeed, the very fluent YouTube user need only see one video frame—the video’s display thumbnail—to make a decision about whether or not the video is relevant to their search or interests. In order to cut through YouTube’s content glut and attract views, a pedagogical or instructional video must adopt the visual conventions of the instructional video in order to advertise its usefulness to a potential viewer. In Chapter One of this thesis I argue that the video score’s indexical semiotics embeds into the score the imperative to be interpreted. In a similar vein, by employing the front-on, eye-level, headless mid-shot camera angle of YouTube instructional videos, the video score is also embedded with the imperative to be mirrored, to be learned from.

An overwhelmingly common use of YouTube is as a ‘peer-to-peer’ content sharing platform—to the unabated consternation of royalties holders—providing a means to circulate digital files such as music, television, and movies without the centralized regulation of a commercial distributor. Musicologist Kiri Miller, in a
detailed field study on YouTube pedagogy, argues that the analogous ‘amateur-to-
amateur’ culture of YouTube’s amateur instructional videos performs a similar
decentralizing redistribution of information networks, where the information being
distributed is not digital files, but rather “kinaesthetic knowledge” (Miller 184).
Despite the ever-increasing volumes of human knowledge which can be digitized
and data-fied, the astounding number of ‘How To Chop Vegetables’ videos on
YouTube attests to the significant quantity of essential human knowledge that
cannot simply be digitized, downloaded, and installed. Miller observes, “we still
have bodies and we still have to practice…we still have to learn from other
people’s bodies, finding a way to comprehend their kinaesthetic knowledge and
make it our own.” In the case of YouTube instructional videos, the screen is an
interface for “body-to-body transmission” where the act and method of
transmission is in fact disembodied (Miller 183). Like teleportation, in order to
travel through a screen, embodied knowledge must first be disembodied – but in a
way that preserves the imperative that it be re-embodied at the other end.

One way of attempting to overcome the disembodied gap in the screenal
transmission of embodied knowledge is to over-compensate with exhaustive detail.
Miller notes that, in one guitar lesson video she studied, the teacher spends eleven
minutes explaining how to hold a guitar pick, during which time he meticulously
describes the physical feel of the pick between the fingers. She speculates that, in a
face-to-face guitar lesson, a teacher would not go into such overwhelming detail
(Miller 155). The YouTube video teacher will likely never interact in real-time or
direct contact with their students beyond receiving comments and messages, to which—as an unpaid amateur—they are not strictly beholden to respond. While a face-to-face teacher can simply identify their individual student’s difficulties, and address them for only as long as it takes for them to be overcome, an effective video-teacher must address in their lesson every possible solution to every possible problem their numerous online students will encounter. More advanced (or impatient) students can easily skip past the parts of the video that are superfluous to their learning needs. Thence comes a YouTube tutorial’s painstaking detail, as a means of “compensating for the restrictions of this virtual lesson interface” (Miller 155).

A YouTube instructional video encourages exhaustive attention to detail not only at the teacher’s end, where kinaesthetic information is disembodied and translated into a screenal format, but also at the point of re-embodiment: the student’s end. By creating a fixed, easily navigable document of kinaesthetic information, the digital tutorial offers the student unlimited capacity to pause, rewind, zoom, slow down, and re-watch the video. Indeed, the syntax of YouTube instructional videos depends on this; rarely does a video teacher assume their video will be watched only in real-time. For example, the YouTube video I consulted to guide me through assembling an Ikea bed frame was scripted and edited in such a way that I was instructed to pause the video while I took the time to execute a certain step, and then to rewind the video to re-watch the instruction before repeating that step on each of the other three corners of the bed. In an extreme
example, I once taught a fairly beginner piano student, Ben, who, despite barely being able to master written sheet music for two hands together, with devoted labour and enthusiasm taught himself to play Elton John’s *Pinball Wizard* cover by watching a YouTube video and working out each precise finger placement and hand position.

In this way, the infinitely navigable digital video format presents its kinaesthetic information as an optimal ideal, which *can* be asymptotically reached given sufficient time, labour, and scrutiny. Each repeated viewing narrows the distance between disembodied screenal representation and re-embodied fleshy performance, between expert and novice. But of course, not all YouTube learners are as devoted as Ben; at a certain point, the student will determine that they have extracted enough kinaesthetic information from the video tutorial to satisfactorily achieve the results they set out to accomplish, to the standard they personally set for themselves – whether it be playing Elton John or chopping onions. At this point the student is likely to retire the instructional video as their source of expertise and information, and thenceforth rely on a synthesis of the information newly-acquired from the video, and their own heuristic adoption of it, to execute the task at hand.

In this way, whenever kinaesthetic information is transmitted from body to body, it inevitably morphs as a result of being tempered to the different physical architectures of the receiving body.

The video score of *XEROX ROCK* necessitates a similar negotiation between flawless replication and interpretive approximation. Most of the video
score’s footage is shot clearly enough that a performer could painstakingly parse the video frame by frame, zooming into hands and keys and strings, to decipher the exact musical material performed. This approach would, however, involve an absurd amount of time and labour, with diminishing returns past a certain point: it would yield less interesting musical results, because such slavish devotion to the score would not generate the textural depth that the score’s deliberate indeterminacy strategizes to orchestrate. Indeed, this approach would exhume the often recognizable musical material being played in the video footage—be it Liszt preludes or Van Halen licks—rendering the score a collage of quotations from other pieces, rather than a generator of unique, original musical material.

However, the score’s necessary interpretive imprecision does not diminish or undermine its authority as a text to be studied. Insufficient attention to the score’s video footage risks not generating enough homogeneity or focus in the eight-layered audio track to produce satisfyingly cohesive musical textures; just like an ill-rehearsed staff-notated score, it risks sounding ‘messy’. The performers of XEROX ROCK are therefore entrusted with making the judgement: with how much detail must the score be examined in order to yield interesting musical results?

One relevant factor in making this judgement is the knowledge that, in live performance, the musicians’ success at interpreting the video score can be scrutinized by the audience: the score is projected in plain view on the stage, serving as a constant source of comparison. And because a video score—unlike
staff notation—does not rely on abstracted codes in which its interpreter must be educated, the score quickly makes itself ‘readable’ to a lay audience. Because score’s fairly straightforward notational strategy of physical mirroring can be followed by the audience more easily than a staff-notation score could be, the threshold of ‘mastery’, or expertise, separating audience and performer recedes. The performer’s familiarity with the score, having studied and rehearsed it, still endows them with more mastery of the score’s performance than the audience; however an audience member can acquire sufficient mastery over the reading of the score to inform a critical eye with which to judge the success of the musicians’ performance – and not only their live performance, but also their past performances captured on tape in the eight-layer audio track. One easily recognizable token of success is the presence of synchresis: the convincing illusion that the sound produced by the musicians is fused with the silent video footage. Michel Chion describes the phenomenon of synchresis as an “irresistible mental fusion”: so powerful a perceptual phenomenon that when it’s present, it is ‘irresistibly’ present (Chion, xix). Little critical know-how is necessary to determine the presence or absence of synchresis, especially when the audience almost certainly arrives in the performance moment with an already high degree of fluency in audio-visual media. The visible video score, in a sense, authorizes the audience to be an armchair critic.

A similar spirit of amateur criticism is also alive and well in YouTube’s amateur-to-amateur music culture. Miller, citing the critical comments that amateur music tutorial videos attract from fellow amateurs, argues that, in an amateur-to-
amateur network of knowledge transmission, “established ideas about expertise, authenticity, and authority become subject to public debate grounded in idiosyncratic individual experience” (Miller 218). YouTube’s distribution interface—which, in a sense, presents video content simply as a vehicle for interactive user evaluation in the form of ‘likes’ and comments—shapes a culture of discourse in which the critical eye is no longer the prerogative of educated amateurs, but is actively fostered in all users. What counts for expertise and mastery is not judged according to the authority or institutional standing of the information’s source; instead, it is something to be contested and negotiated by the users who consume that information. In a similar way, a video score inhabits an ambiguous and contingent critical frame: without an established performance practice, what makes a performance of a video score ‘successful’ is under continuing negotiation and re-evaluation.

3.3 TIME IN THE DIGITAL DIMENSION

Chapter One of this thesis discussed the technical mechanics and operating structures of the video medium in the context of analog video: sound and images encoded onto electromagnetic tape, offering the video-maker only limited ability to splice, cut, and edit footage in the artful way that the cinematic medium of celluloid film allows. Thus far only passing acknowledgement has been made of the fact that the three video scores in this portfolio are of course digital video, not
analog: their source footage has been captured as digital data, and edited in digital video software that enables comprehensive manipulation and re-collaging from a vast databank of footage. With the assimilation of both video and cinema into the digital dimension, the material and technological distinction between the two mediums has been all but lost. The distinction between the concepts of ‘video’ and ‘film’ persist instead through the cultures and networks of production, consumption and distribution.

Despite this significant technological evolution, it is still relevant to frame the aesthetic mechanics of video scores in relation to video’s original technical mechanics. I would argue that the other two pieces in this portfolio, \( \textit{rupture} \) \( \| \) \( \textit{rapture} \) and \( 8 \times \infty \), maintain a demonstrable analog video aesthetic because they do not make use of editing capabilities unique to digital video; in other words, these two video scores could conceivably have been created with analog video without significant functional changes. Both scores use footage that I, an independently operating artist, was able to produce single-handedly; compiling the score required no more than sequencing this footage, which dubbing one videotape to another makes perfectly achievable. In \( 8 \times \infty \), the exactitude with which events are timed between the eight separate ‘parts’ in the score would be much more difficult to execute in analogue video, requiring careful rehearsal and onscreen performance rather than being easily tweaked in post-production – but the translation from digital media would not be impossible.
**XEROX ROCK**, however, could only be a product of digital video. It draws its material from a huge library of footage originating from numerous contexts; this contextual breadth could only be collected (by the independent artist, rather than a studio-supported production crew) via a resource like YouTube, which makes an enormous volume of visual content freely available to the video-maker. The aesthetic texture of **XEROX ROCK** depends on this contextual glut. It is important that the score uses footage of both professional concert musicians and bedroom amateurs, and that it includes videos from both performance and pedagogical contexts. In this way, the score levels distinctions between ‘preparation time’ and ‘performance time’, visually counterpointing the audio track’s analogous compression of rehearsal and performance time into the same sonic event.

An incidental, but no less interesting, by-product of siphoning footage off YouTube is the synopsis it reveals of each instrument’s unique subculture, as presented and negotiated in a globalized, peer-to-peer sphere such as YouTube. The guitar footage tends conspicuously towards male amateurs performing in domestic, to-the-camera settings. Several of the amateur guitar videos boast high-end prosumer production values, implying the amateur guitarist’s considerable financial investment in both their hobby and their personal promotion of it. The piano footage tends towards professional musicians performing in live concert settings – a symptom of the increasing demand for professional pianists in a competitive field to carefully curate their own online presence. The marimba footage largely originates from the official channels of percussion instrument
manufacturers, which feature sponsored videos of percussionists endorsing and demonstrating a manufacturer’s instruments and/or mallets – a reminder that, because of the sheer expense of the instrument, mallet percussion is one of the least available musical outlets for the amateur.

Most notably, although the score’s heterogeneous video database visually invokes a wide range of musical styles (e.g. rock guitar vs. classical piano vs. virtuoso percussion repertoire), the score’s interpretive imperative—that only the figures’ movements be re-enacted, and not the video’s supposed soundtrack—largely erases these stylistic distinctions from the sounding result. The originating context of the footage in the score lingers not in easily (and aurally) identifiable markers of musical style, but in the visual residue of extra-musical factors associated with that musical style: markers of people, place, profit, and production.

Not only does XEROX ROCK depend on the breadth of the digital video archive, what’s more, the score’s rhythmic and temporal profile could only be achieved via digital editing software. The montage-effect of rapidly cutting between heterogeneous footage is a cinematic convention that arose from the easy manipulability of celluloid, but would significantly challenge the capabilities of analog video production. XEROX ROCK’s rhythmic language also depends on digital video’s capacity to be slowed down, sped up, frozen, looped, and reversed – to orchestrate a mutually contingent distortion of time and the body. Only rarely in the score is the footage let to play out in real time, in embodied time: the time it took for the action to be comfortably executed in fleshy space. Instead, the
temporal profile of the footage is tightly controlled within a digital space, where it is divorced from the physical considerations of the acting body and instead recalibrated in anticipation of the action’s satisfying sonification when it is re-embodied. In other words, the score depends on the ability to manually ‘musicalize’ video material whose temporal profile might not have been, in the first instance, visibly ‘musical’. Lev Manovich synopsizes the distinction between analog and digital film practices with the epithet that “digital filmmakers work with ‘elastic reality’” (Manovich 301). The reality of time in XEROX ROCK is treated with substantial elasticity, just as the distinctions between various real-world musical contexts are pulled and stretched.

One characteristic rhythmic device in XEROX ROCK is the loop: repeating a limited excerpt of video footage a number of times in direct succession, most often edited alternately in reverse playback so as to create a figure-of-eight effect. The loop’s repetitiousness is reminiscent of the pedagogical technique of imitation-and-repetition: for instance, the student repeating a task after the teacher in order to imitate their expertise, or the student being compelled to repeat certain tasks ad nauseum as a strategy for their eventual mastery. As has already been discussed, a digital teacher is endlessly repeatable and, as Miller describes, “infinitely patient” with the student’s own repeated attempts. The repetitious looping in the video score therefore reinforces the score’s pedagogical quality by its implicit imperative that each subsequent repetition be more successful than the last (Miller 165). In this way, the loop orchestrates horizontally in time what the eight layered rehearsal
takes orchestrate and collapse into a vertical texture: sustained repetition as a device for incrementally but demonstrably approaching an optimal realization of the score.

Also relevant to the way the loop functions in *XEROX ROCK* is Lev Manovich’s reminder that the loop—now a ubiquitous rhythmic device in temporal digital media art—began life as an artefact of limited computer memory. Early computer games, for instance, in the absence of sufficient processing power to continually animate characters in real time, relied on looping a limited repertoire of actions in order to depict characters in motion – for example, looping one step over and over serves as an adequate representation of walking (Manovich 316). Just as looping a video game maximises limited computer memory, looping in a video score serves as a strategy for maximising limited human memory. One crucial hard-wired constraint of the video score is that it is drip-fed to a performer in real-time: it does not supply an overall mnemonic map in the way a paper score can be studied and parsed outside the ‘musical time’ of the score itself. A performer is therefore reliant on their own memory of the score for any kind of formal navigation or pre-emption of forthcoming material.

With this as the case, a constant onslaught of new, unique score material risks overloading a performer’s mental ‘processing power’ as they scramble against the clock, as it were, to interpret and respond to the score’s constant stream of prompts. Much ‘cheaper’ is to re-use recent material still fresh in the performer’s ‘random access memory’ so that, instead of continuously interpreting prompts, a
performer has processing power available to ‘groove’ along with the video score: to pursue mastery of the material, and begin to employ artistic license in its interpretation. In this way, the video score prototypes a cybernetic approach to notational strategies and musical interpretation.

Another technique whereby the video score manipulates an experience of time is via the recurring motif of the freeze-frame, which—especially due to the videos’ demonstrable origin from the internet—calls to mind the buffering glitch: the failure of digital video data to be re-constituted in real time. The buffering glitch is a telling reminder of the video’s digitality: while analog video is re-constituted in real-time by mechanical means—tape running over a reader—meaning any playback glitch is the result of a material snag, digital video is substantiated into a moving image via computational processes, and its glitches come about through imperfections and bottlenecks in data flow. Those literate in digital media know to read the buffering glitch not as a pause in the media, but as a pause in the interface; in other words, it’s not the video that stops – only its playback. The buffering glitch is cognitively excised from one’s experience of the media’s temporal momentum, and therefore seems to happen ‘outside time’.

Employing the glitch as a rhythmic motif in the video score thus introduces a certain ambiguity to the dimension of time in performance – should the freeze-frame be read as a pause in the media, or as a pause in the interface? In other words, is the freeze-frame a pause in the score—i.e. a ‘musical’ pause, like a rest or caesura—or a pause in the performers’ reading of the score, much like a page turn
or a break between movements? These ambiguities serve as visible reminders of the video’s function as a score and an interface: when the video stops in its tracks, so do the performers, thereby reinforcing the video score as a constant object of the performers’ negotiation and interpretation, rather than their mere accompaniment. Kris Paulsen argues that, where interfaces are involved, visibility is synonymous with dysfunction: “we forget the ‘good’ interface precisely because it is transparent” (Paulsen 104). In this sense, freeze-framing the video engineers a deliberate rupture in the transparency of the score so that its presence—and influence over the performing musicians—can be re-asserted. That the freeze-frames orchestrate the distinctive musical punctuation mark of silence—a glitch in the performance’s sonic as well as visual fabric—also supports their function as prompts to take stock, in performance, of the networks of interaction and influence between score and performer, moving image and sound.

Despite these temporal punctures deliberately revealing the video score’s relational guts, the video score nonetheless equivocates between its utility as an interface and its aesthetic appeal as a moving image. Lev Manovich describes the overlapping functions of images employed in interfaces, oscillating between fictitious illusion and practical utility as “the new role of image as image-interface competes with its older role as representation” (Manovich 290). In a similar way, the moving images of a video score slip between functioning as interface—a site for communicating information, in this case musical notation—and functioning as illusionistic representations. The figures in the video score possess an illusionistic
quality, seeming to take on a role as performers themselves; after all, the tasks they execute are uncannily similar to the tasks executed by the live musicians as they ape the actions in the video score, and their actions are substantiated with sound. By virtue of using the uncanny image of the human figure as its notational device, a video score preserves a degree of autonomy as a representational aesthetic object over and above its utility as a score.
APPENDIX

i. rupture \ rapture

TEXT INSTRUCTIONS FOR PERFORMANCE

[next two pages]
Performance Notes

Voice

The vocalist follows the audio track. Her voice is closely amplified.

Her task is to translate, as meticulously as possible, the sounds in the audio track into vocal sounds. With her mouth, teeth, tongue, lips, and voice, she activates a broad spectrum of percussive and pitched sound. While richly varied, the sounds also seem secretive.

She performs wearing headphones so as to remain tightly synchronized with the score. She may also watch the percussion score on a laptop screen. The ‘pips’ that accompany the countdown at the beginning of the score are for timing only, and are not vocalized. Single ‘pips’ also occur as cues at the end of a significant silence, shortly before the next sound; these pips are likewise unvocalized.

Percussion

The percussionist follows the video track. She performs on a circular drum of her choice, which has a diameter approximately equivalent to double her handspan. Her drum is closely amplified.

Her task is to mirror, as meticulously as possible, the movement of the hands in the video, using the drum as a surface. While observing the movement scrupulously, she also has an ear attuned to coaxing a rich variety of sounds out of the drum.

During performance, she follows the score on a small screen so as to remain tightly synchronized with the score. She may also choose to listen to the vocal score through headphones.
The two performers each require a laptop and a pair of headphones from which to ‘read’ the score in performance. The two laptops are connected with an ethernet cable so that the score playback can be synchronized between the laptops with a screen-sharing app.

There is a video camera trained on the percussionist’s hands in a tightly-framed, level, extreme-close-up shot (similar to the framing of the score). This live video feed is projected (probably via HDMI, depending on the camera’s output specs) onto a screen on the back wall of the stage. The camera is gaffer-taped upside-down on the tripod so that the projected image appears upside-down.

If specialized lighting states are available, the stage is as dark as possible, with a tightly focussed, low-intensity spot on each performer.

To view this piece in performance, visit:
http://vimeo.com/117703673
ii. XEROX ROCK

TEXT INSTRUCTIONS FOR PERFORMANCE

[next two pages]
**XEROX ROCK**

Celeste Oram

*a video score for electric guitar, piano & xylophone with fixed A-V media*

---

**Performance Notes**

The task of the ensemble is to mirror the movement of the human figures in the video score in such a way that sound is produced on their instruments.

The guitarist follows the footage in the top-left panel; the xylophonist follows the footage in the top-right panel; the pianist follows the footage in the lower-left panel. The lower-right panel is deliberately left blank in the score.

The ensemble strives to synchronize their actions precisely in time with the video score itself. At all times, the ensemble mirrors only the actions of humans in the score. Mechanistic/robotic actions are not mirrored by the live musicians; this footage essentially functions as a rest.

The guitarist uses a clean, bright tone with minimal reverb. The xylophonist uses four mallets. The pianist activates the pedals only when footage in the score features feet.

*score available at [http://vimeo.com/celesteoram/xeroxrockscore](http://vimeo.com/celesteoram/xeroxrockscore) with the password 'nikelanddime'*

---

**Tech-Spec Notes**

In performance, the ensemble is accompanied by a fixed audio track, compiled from recordings taken of each full run-through of the score in rehearsal. These recordings are palimpsested/multi-tracked/layered on top of one another in simultaneous synchronization. The most recent takes are mixed at a higher level, with older takes mixed at progressively
lower levels. This audio track is played back in performance such that it ‘shadows’, or synchronizes with, the live ensemble; and at such a level that the loudest take in the audio track is roughly the same level as the live sound.

It is recommended that, when recording each rehearsal take, the countdown pips at the start of the score are made audible in the recording. This will facilitate easier synchronization of the takes when compiling the multi-tracked audio.

In performance, the score is projected onto the back wall of the stage so it is visible to the audience. In the version of the score visible to the audience, the lower-right panel of the score is filled in with video footage of the ensemble in their most recent rehearsal before the performance, synchronized so that their actions align temporally with the corresponding actions in the score itself. The ensemble themselves read the score off laptop screens; multiple laptop screens can be synchronized via a screen-sharing app, using an ethernet cable to network the laptops.

[NB: The composer will undertake to compile both the accompanying audio track, as well as the performance-ready version of the video score.]
APPENDIX

iii. $\delta x \to \infty$

TEXT INSTRUCTIONS FOR PERFORMANCE

[next four pages]
Celeste Oram

an audio-video score for flute, bass clarinet, piano, 2 percussion, violin, cello, & double bass with fixed audio

11’
Performance Notes

There are two phases involved in this work. The second phase is the live performance in concert. The first phase is recording the audio material that makes up the fixed audio track necessary for the live performance.

Phase I

The first phase is rehearsed and prepared not together as an ensemble, but individually. Each musician is assigned a package of seven audio-video files which serve as scores for the work. Of these seven scores, five are common to the entire ensemble: the ‘ritornellos’, which contain primarily audio material. The remaining two scores are unique to each instrument, and contain video material without any audio soundtrack. In these video-only scores, there are two kinds of video footage: one kind features the musical instrument, and one kind does not.

Whenever a musical instrument does appear in the footage, the musicians’ task is to faithfully mirror the movement of the figure in the video; resultant sound may or may not be produced. This footage will periodically demonstrate for the musician to shift between different positions or ways of holding their instrument. Once a new playing position has been established, the musician continues to hold that position until another episode featuring the instrument demonstrates a change of position. Think of it as a game of ‘Simon Says’, in which players only respond if Simon says so; the musicians only shift the position of their instrument if the instrument in the video score shifts position.

Whenever there is no musical instrument in the video footage, the musicians’ task is to mirror the actions of the figure in the video in such a way that sound is produced with either their instrument or body. This may, in practice, involve adapting the position or direction of the action so that it interacts with the instrument. Note that all video score footage—both with and without instruments—is presented from a mirror-image perspective: the score is therefore followed like a literal mirror, or an aerobics instructor (i.e. if the hand on the right-hand side of the screen moves, the musician moves their right hand).

When it comes to the five ritornellos, or audio scores, the musicians’ task is to recreate the sounds as faithfully as possible with their instrument. These audio score episodes are sometimes accompanied by visual representations of the sounds, to serve as an aide-memoire.
To paraphrase: While responding to video footage with instruments, the musicians invest in executing movement. While responding to audio ritornellos, and video footage without instruments, the musicians invest in executing sound.

Once the scores are sufficiently rehearsed, each musician individually records eight full takes of all seven scores. These can be recorded in one three-hour session per musician, which the composer will undertake to direct. Note that only audio material is recorded during these sessions; no video filming is required.

During the recording process, each of the seven scores is recorded in one take, and all eight takes of each score are recorded successively. The musicians choose in which order to record the seven scores. (e.g. all eight takes of the first video score are recorded, then all eight takes of the second video score, then all eight takes of the first ritornello, etc.) The two percussionists do not record the first ritornello audio score. The three short pips which begin each score are there to facilitate easy synchronization of audio post-production, and do not require any response from the musicians.

PHASE II

For phase II, the recording scores from phase I have been re-arranged into one continuous audio-video score, with a unique part for each musician. The recordings from phase I have been assembled into a fixed audio track which directly synchronizes with the video material that prompted those sounds.

In concert, all eight musicians perform a live rendition of their own score together as an ensemble. The fixed audio track is played aloud in the auditorium, audible to both musicians and audience. However, the musicians appear onstage without their instruments, and therefore their responses to the score differ from phase I:

>> When footage with musical instruments appears, the musicians do not respond to it; instead, they remain still.
>> When footage without musical instruments appears, the musicians faithfully mirror the movement of the figure in the video. The performed movements should be tightly synchronized with the sounds in the audio track.
>> When the ritornello episodes appear, the musicians perform the score exactly as they had performed it in phase I, as if holding and handling their real instruments. They remain seated, even if their actions when performing with their own instruments involved otherwise.
>> At all times, the musicians’ movement appears more natural, graceful, controlled, and intentional than it does robotic, awkward, or aimless.

[ NB: The clarinet, percussion, and double bass will notice there are brief episodes in their parts when the audio track features another instrument. At these moments the musicians should faithfully mirror the movement of the figure in the video, as usual. ]
To paraphrase: Musicians do not respond to the video footage which features instruments. While responding to ritornellos, and video footage without instruments, the musicians invest in executing movement with conviction and rhythmic accuracy.

Each musician is seated; each of their chairs is of a visibly different type. They are arranged onstage in a shallow semicircle, seated in alphabetical order of surname from stage right to stage left (e.g. from stage left to stage right: Williams > Takagi > Means > Marotto > Gorczyca > Goodchild > Diaz > Brady). Each musician reads the score from the screen of a laptop or tablet. However, the ritornellos’ original audio tracks are not included in the performance scores (i.e. the performers do not require headphones or speaker monitors). Instead, during the ritornellos, the musicians follow both the visual aide-memoires onscreen, and the aggregated sound of the fixed audio track playing aloud in the auditorium. The ritornellos in the performances scores are accompanied by an audio mix of all eight instruments; to aid in preparation and rehearsal, soundfiles of each individual instrument track for each ritomello are available in the media downloads collection in a folder named “individual ritornello tracks”.

Media downloads collection at:
https://drive.google.com/folderview?id=0B_VsrSpgG_wbVG9UV0g0T1JSNGM&usp=drive_web
REFERENCES


