Reflecting Music Through Movement
A Body-Syntonic Approach to Playing [with] the Piano

Xiao Xiao

Bachelor of Science in Computer Science and Engineering
Massachusetts Institute of Technology, 2009

Master of Science in Media Arts and Sciences
Massachusetts Institute of Technology, 2011

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Media Arts and Sciences at the
Massachusetts Institute of Technology
September 2016

© 2016 Massachusetts Institute of Technology. All rights reserved.
Reflecting Music Through Movement
A Body-Syntonic Approach to Playing [with] the Piano

Xiao Xiao

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning, on September 7, 2016
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Media Arts and Sciences
at the Massachusetts Institute of Technology

Abstract

This thesis introduces and examines methods for the capture and reproduction of music on the piano that maintain a tight coupling between the body, the sound, and the physical instrument. For expert musicians, the body plays an indispensable role in the physical act of playing and the understanding of both musical structure and expressivity. However, many music learning technologies mistakenly assume that playing music is “playing the notes” and neglect the role of the body in the development of the musical mind.

Drawing from research in telepresence, tangible interfaces, and augmented reality, I propose to bring the body back into the picture, literally and metaphorically, by augmenting a digital player piano with projection-mapping. My platform synchronizes dynamic imagery with the piano’s moving keys and acoustic sounds.

I here focus on two main projects: MirrorFugue and Andante. Inspired by reflections on the lacquered surfaces of a grand piano, MirrorFugue simulates the presence of a virtual pianist whose reflection is actually playing the physically moving keys. It encourages anyone to take the seat left empty at the piano, to feel in his or her own body how music is expressed through the body of the performer, and to play along. Andante presents music as miniature figures that appear to walk and dance on the piano keyboard, physically striking a key with each stop. It conveys the expressivity in rhythms and phrases as well as musical structures through the bodies and movements of the figures. Both installations are designed as immersive “sandboxes” for the playful exploration of musical ideas.

Beyond my projects, this thesis explores the parallels between music learning and learning in general. I discuss the connections between theories of music learning (particularly Dalcroze Eurhythmics) with theories of general mental development (Jean Piaget, Jerome Bruner, Seymour Papert, and Marvin Minsky), as well as how strategies from music learning could inform the art of learning in general.

Thesis Supervisor: Hiroshi Ishii
Title: Jerome Weisner Professor of Media Arts and Sciences
Reflecting Music Through Movement
A Body-Syntonic Approach to Playing [with] the Piano

Xiao Xiao

The following people served as readers for this thesis:

Signature redacted

Reader: 

Ken Perlin
Professor of Computer Science
New York University

Signature redacted

Reader: 

Edith Ackermann
Professor of Developmental Psychology
Université d'Aix en Provence, France
Visiting Professor of Architecture
Massachusetts Institute of Technology

Signature redacted

Reader: 

Donal Fox
Composer and Pianist
“All experts (artists and scientists) accumulate towers of skills, building new ones on top of older ones. At every stage we become novices again, finding new goals (and hence new obstacles) and then seeking ways to deal with them”

– Marvin Minsky, *The Future Merging of Science, Art, and Psychology*

... in loving memory
# Table of Contents

## Acknowledgements

4

## Introduction

10

- Music Learning
- Music Learning + Technology
- Music Learning + Learning at Large
- Roadmap

## I: Problems of Music Learning

11

- Chapter 1: The Pop-Ed Music Culture
  - Learning music = Learning Pieces
  - Learning Pieces = Learning Notes
  - Practice = Repetition
  - Key to Success = Talent

12

- Chapter 2: Taxonomy of Music Learning Systems
  - Learn to Play Music
  - Learning through Playing with Music

13

- Chapter 3: [Technological] Origins of Pop-Ed Music
  - The Birth of the Composer
  - The Machine and its Instructions
  - The Age of Disembodied Sound
  - Technologies for Possibilities

## II: Designing Interactive Technologies

17

- Chapter 4: The Piano as a Platform for Invention
  - Piano
  - Computing

18

- Chapter 5: MirrorFugue: Conjuring the Recorded Pianist
  - Re-Presenting the Body
  - From Reverie to Reality
  - Implementation
  - Selected Observations
  - Potential Interactions
  - User Study

21

- Chapter 6: Andante: Microworld of the Keyboard
  - Origins
  - Implementation
  - Content
  - Interactions
  - User Study

22

## III: Music Learning and Beyond

25

- Chapter 7: Dalcroze, Design, and Mental Development
  - Genesis of the Musical Mind
  - Design Framework
  - Connection with Constructionism

26

- Chapter 8: Mathetic Principles from Music
  - Conditions for Building the Mind
  - Building and Debugging the Mind
  - Meta-Mathetics: Learning Across Domains

27

- Chapter 9: As We May Play: A Vision and Variations
  - Humanist Computing
  - FUGO: A Platform for Playing with Music
  - Beyond Music

30

## Conclusions

112
## Appendix

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Meditations from the Practice Room</td>
<td>117</td>
</tr>
<tr>
<td>B. Invisible Cities of the Creative Mind</td>
<td>125</td>
</tr>
<tr>
<td>C. Reflections on Marvin Minsky and Music</td>
<td>132</td>
</tr>
</tbody>
</table>

## Notes

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notes</td>
<td>138</td>
</tr>
</tbody>
</table>

## Bibliography

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bibliography</td>
<td>149</td>
</tr>
</tbody>
</table>
Acknowledgements

First order of thanks goes to my advisor, Hiroshi Ishii, for accepting me into the Tangible Media Group, for supporting my research, and for continuing to push my thinking to new frontiers. I am incredibly grateful to have been your student (since a UROP in 2008!) and for the chance to play intellectual pingpong with you for all these years.

My committee:

Ken Perlin, every conversation with you leaves my mind buzzing with new ideas. Thank you for being an inspiring mentor and a great friend (you knew that all of this was just an excuse for us to hang out!). I can’t wait to see what the Future Reality holds!

Edith Ackermann, your warmth, your wisdom, and your childlike spirit have been a constant inspiration, especially in my darkest days in the tunnel of this thesis. I cannot thank you enough for believing in my ideas, for all your insights on the mind/body/heart, and for showing me how to sift the gold from the sand.

Donal Fox, so much of this thesis would not have been possible without everything I have learned from you, from the Art of playing the piano to connecting the dots between music and everything else. Your mentorship

Xiao Xiao | PhD Dissertation
has been instrumental to my mental and emotional development these past few years, and I am forever grateful to have you in my life.

My collaborators:
To Yamaha and to Roger’s Pianos for so generously supplying the instrument to enable my research. And to the MirrorFugue Pianists: Donal Fox, Marvin Minsky, Vijay Iyer, Ryuichi Sakamoto, Allen Toussaint, Jon Cleary, Ron Markham, Nick Sanders, Alisa Ishii, Lena Ishii, Joe Paradiso, Peter Godart, Thiago Vittorio, Gretchen Schadebrodt, Diana Borshcheva, Michael Gold—thank you for allowing me to capture your soul(ful playing). To Nick, especially—remember when we were kids competing in piano competitions? Look how at where we are now!

Pablo Puentes, I learned so much from our conversations and from watching you in action teaching the children. Thanks for introducing me to Dalcroze Eurhythmics! I wish that circumstances would have allowed us to continue our experiments together, but I’m excited for what comes out of your PhD in Sydney!

Hanoi Hantrakul, your knowledge of music history played a big role in shaping of Chapter 3. But more than that, I enjoyed talking with you about everything from music to physics to Buddhism. Hope to finally visit Thailand one day. (P.S. May you always be “shapeshifting like Optimus Prime”).

Arthur Petron, thanks for your help designing with the MirrorFugue truss (which has been on 3 continents!) and for co-hosting the thinking across domains workshop with me.

Anna Pereira, Basheer Tome, Alan Kwan, Paula Aguilera and Jonathan Williams—thanks for all your help on my publications and videos. It has been an honor to have worked with all of you.

DonDerek Haddad, Penny Webb, Dan Fitzgerald, Akito van Troyer, Rébecca Kleinberger, Thomas Sanchez—it was fun to collaborate with all of you on the shape display projects!

Thanks to the fantastic, hardworking UROPs and UAP who went above and beyond what was asked of them: Lauren Gust, Kiran Wattamwar, Christina Sun, Kojo Acquah, Cattalyya Nuengsigkapian, Zack Ullenhuth.

The Tangible Media Group:
Members of the Tangible Media Group, past and present: Keywon, Daniel L., Sean, Jinha, Dávid, Tony D., Lining, Jifei, Cati, Amanda, Shunichi, Chigira, Yusuke, Shinji, Tommy Jamie, Austin, Sam, Sheng-Kai, David Rose, Ken, Luke, Viirj, Penny, Daniel F., Udayan, Wen, Ryuma, Basheer, Felix, Philipp, Clark—It has been an honor to work alongside all of you. I miss sitting through group meetings with you already!

And many thanks to our amazing admins: Sarah Summers, Mary Tran Niskala, Natalia Villegas for their tireless work behind the scenes making sure research goes smoothly.

Two special shoutouts for the beginning and end of my TMG journey: To Leonardo Bonanni for initiating me into the group, first as a UROP, then as a Master’s and PhD student. And to Daniel Levine for staying up with me the last night before my thesis submission proofreading and and polishing.

The MIT Media Lab Community:

I count my lucky stars to be a part of this incredible network of such bright stars! Alas, there are too many stars in these skies to name them all here, but some particular constellations that have been especially meaningful to me over the years include:

Mitch Resnick—I still have many fond memories of being in your freshman advising seminar. Thank you for giving me my first opportunity to UROP at the Media Lab and for introducing me to the ideas of Seymour Papert. What I learned from you has been formative to my own ideas about learning.

Neri Oxman—You are an inspiration, a paragon of what it means to live for your artistic vision. Thank you for believing in me as an artist. It means the world.

Tod Machover, Pattie Maes, Joe Paradiso—Thank you for your kindness and support over the years, and for allowing me to participate in your classes. I’ve learned so much from each of you!

Linda Peterson and Keira Horowitz—Thanks for your patience and compassion in helping me navigate the quagmires of thesis-land!

Joi Ito—It has been a joy to watch you transform the lab! I was so excited to meet you during the interview in April 2011, and you have completely blown away my (already high) expectations! Thank you for bringing so
many diverse points of view to the lab. Our discussions about metaphors across domains and the time I spend in your Awareness class have been highly influential to many of the ideas in this thesis.

Maurice Ashley—I am grateful for the opportunity to get to know a real life chess grandmaster! It has been a pleasure to compare ways of thinking between music and chess with you.

Idit Harel—You are the *deus ex machina* of this thesis. Had you not stepped in a month ago, treated me to yoga and sushi, and talked me through my ideas, I don’t even know what I would be submitting right now. Your intervention brought new life to my writing, gave it a new groove, and showed it a new song to sing. For all this I cannot thank you enough.

**Outside the lab:**

To everyone who gave me an opportunity to speak and to present my work—Danielle Duplin and Matta Saia with TEDxBoston (and Mason and Danny for helping me craft my first major talk), Daniel Obodovski with TEDxSan Diego (and Kate Peters for her coaching), Saul Kaplan with BIF, Aspen Ideas Festival, Mr. Kawabe and Mr. Kubo with Intersect by Lexus, Tara DePorte with the Human Impacts Institute Boston Climate Salon, the Boston Arts Academy, and Kongo with EMW.

To Ruth Lingford and Dylan Ladds with the animation class at Harvard, thanks for introducing me to the art of Animation! Andante could not have happened without what this class!

Phyllis Taylor, The Historic New Orleans Collection, Ken Schwartz at Tulane—Thank you for the opportunities to share my work in New Orleans, my beloved hometown. And Sharon Litwin, it was a pleasure working with you on the MirrorFugue exhibit. Your energy never ceased to impress me. May you rest in peace.

**Friends and family:**

Tauntaun, my bff and intellectual catalyst. Your feedback and your own wealth of thinking has been an invaluable to the growth of this thesis. Your openness to experience, fearlessness, and boundless compassion inspires me to no end.
Reflecting Music Through Movement

Gina, my oldest friend, we have known each other for longer than freshmen have been alive! Thank you for being there for me always, and may we continue to grow together.

4th WAR family, especially Shanying, TR and Mich, Fogg, Canada, Rosser, GSE—The honor is entirely mine to have lived amongst you all. And juggling is totally embodied cognition!!!

Friends from lab and outside of lab, past and present, from near and far including Layla, Khalida, Rickie Ashman, Peter Godart, Sandra, Nan, Gershon, Edwina, Jie, Steve, Maya, Nina, Pip, Darthur, Rébecca, Akito, Dávid, Ani, Brian, Jeremy, Pragun, Leanne, Nicole, Juan, Dan Chen, Amit Zoran, Amna, Anika, Denise, Clément, Ira, Kevin Chiu, Katja, Katia, Mei, Robb, Sarah Fdili-Alaoui, Nuno, Emilienne, Bruna, Muffin, Jeff Lieberman, Christian, Savannah, Spencer, Novysan, Sunanda, Xin, Becky, Cranor, Raphael, Dhaval, Joscha, Jake—thank you for existing!

Special thanks to Usman for reading and commenting on parts of this thesis (all the way from Paris!) and to Chrisoula for lending an ear and for suggesting the term “mental infrastructure.”

R. Hinton—I have not forgotten that it was you who lent a hand to help me build the very first projection system for MirrorFugue. You helped open my eyes to a new world of aesthetics, and your relentless pursuit of your art gave me confidence to pursue mine. Our paths may have diverged, but I still believe in you, now as ever, and I wish you nothing but the best.

Chris—From you I first learned about the art of hacking code, and it was you who first inspired me to study computer science. It has been a joy to know you for all these years. (P.S. Thanks for the Dropbox space for our group!)

Ivy Street Crew—Gloria, Margaret, Cynthia, Mary, Barry, Richard, Ruth, Sally, Maddie, Ruben, Rodrigo—I am grateful to be a part of your circle. Our Saturday nights together with Marvin over the past year and a half has inspired much of this thesis. Thanks for listening to me play the piano and for the insightful conversations. I am especially grateful to Gloria for making Ivy Street a home away from home.

爸爸妈妈, 外公外婆, 嘟嘟小姨丫丫—Thank you for your love and support. 外公外婆—Thank you for raising me and for giving me an early start in playing
the piano. 爸爸, now that my thesis is done, maybe I can finally read your “Piagetian” accounts of when I was a baby!

東东, mon beau phénicien— يالا حبيبي—laissons-nous rêver de l’infini, de l’univers le plus profound qui n’est que pour nous (et pour les enfants, qui passent 看懂这一句)。“
Introduction

At the most recent Interaction Design for Children (IDC) conference, I was told an anecdote from a previous IDC panel featuring Seymour Papert, Marvin Minsky, and Alan Kay. A 9-year-old girl had asked about effective ways to learn, and one of the tips that Minsky shared was to learn something very well and use it to help you learn everything else. For me, this prototypical subject has been music.

While learning any subject yields insights for learning others, several unique properties of music make it particularly rich in this regard. Neuroscientists have observed that playing music coordinates more regions of the brain than any other activity. Even without scanning the brain, it is clear that playing music involves physical, intellectual, emotional, and social dimension. Moreover, almost all humans have a natural affinity for music. As Minsky once observed in a research proposal, “in our cultures and many others, people spend more time interacting with music than any other waking activity.”

Unfortunately, though almost everyone loves music, learning to play music is generally believed to be very difficult for all but a “talented” few. These apparent difficulties not only block access to a rich world of creative expression but also to a treasure-trove of insights about learning at large.
This thesis has three overarching goals. First is to determine how to learn music more effectively and enjoyably, from the earliest stages all the way to high levels of expertise. Second is to design new interactive technologies for music learning based on these approaches. And third is to connect strategies for music learning to learning just about everything else.

**Music Learning**

Many who learn and teach music follow a set of default assumptions shared by the general public, where learning music is often equated with learning to play pieces written by composers. This typically involves long hours of repetitive practice to correctly play the notes of a piece, which is assumed to be a necessary step before learning to be expressive.

Having learned to play the piano like this for 18 years, I argue that music seems difficult because this way of learning makes it difficult, not to mention ineffective and unpleasant. Fortunately, there are other ways, which I learned through my mentor composer/pianist Donal Fox. Seven years ago, I began lessons with Donal to learn how to improvise, and ended up unlearning and relearning everything I thought I knew about playing the piano, and about music. From these lessons, I learned that music is not about the notes but about expression, and that expression must be understood through the body. Learning to be expressive does not happen through mindless repetition but through listening, imitating, and experimenting with sound. Expression is about translating what impresses on us—by bringing forth in our own ways.

This way of learning music isn’t new. It has been discovered and taught across the world for as long as music has existed. Most musicians labelled as “talented” learn through similar approaches, even if they cannot always describe what they do. And there is a substantial literature that does describe these better ways to learn music. This thesis has borrowed many ideas from the music educator Eugène Jacques-Dalcroze, which closely relates to the ideas of others such as Willems, Kodaly, Orff, and Suzuki. For the piano specifically, we could consult books by Abby Whiteside, George Kochevesky, Heinrich Neuhaus, and Josef Lhevinne. More recently bassist Victor Wooten, bassoonist David McGill, guitarist Gerald Klickstein, and jazz pianist Kenny Warner among others, have written excellent books on learning musical expression and better way to practice.
If alternate approaches have always existed and are so well-documented, we might ask why mainstream music learning became so problematic, and why it is still so flawed. But more importantly, we could raise the question of how to supplant the problematic defaults of music learning with better approaches. Answering these questions takes us to the realm of both techniques and technology, and to the reasons behind the second goal of this thesis.

**Music Learning + Technology**

Following Marshall McLuhan’s argument that technology changes culture, this thesis puts forth a theory that technologies for music dissemination in the West—namely, modern music notation, the piano, and audio recording—gradually changed the cultural conception of music from expressive performance to disembodied sound. And it was this change in people’s conception of music that let to the preoccupation with notes at the expense of expression in mainstream music learning. Given that technology catalyzes cultural change, I then explore the potential of digital, interactive technologies to promote approaches of learning music that bring the body back into the picture.

By digital interactive technologies, I mean technologies having to do with modern-day computing, but not limited to the current predominant form of modern-day computers (e.g. desktop and laptops, tablets and smartphones). The true power of computing comes from the ability to digitally represent information, and to simulate and manipulate not only numbers but also words, images, and sound. Digital representation allows machines to store vast amounts of information using very little physical space. It also enables the rapid dissemination of information over distance, as well as real-time interactivity.

A number of interactive systems for music learning have been developed, in research and commercially. Most could be categorized under two main camps: one which takes for granted the default assumptions of music learning, and the other which rejects them. The first camp tries to simplify the process of learning to play pieces from the written score by identifying the “pain points” of the typical process and designing technological “solutions”. For example, to target the difficulty of reading standard notation, piano learning systems show users which notes to play using colored blocks. Though these systems may help students learn the notes more quickly, they reinforce the problematic notion that getting the notes right is more important than conveying a musical sequence, or stream.
Designers from the other camp gives people tools to lower the floor of musical expression and creative experimentation. However, these tools are often limited in their musical potential and do not provide a felt ground and clear path to higher levels of expertise.

My research explores how interactive technologies could help people learn to play the piano with greater expressivity. I focus on the piano in part because I am a pianist but also because the piano is a well-established “sandbox” for playing (with) music that offers a relatively low floor with a very high ceiling. This thesis describes two main projects that augment a digitally controlled player piano with in-situ projection: MirrorFugue and Andante. MirrorFugue simulates the presence of a virtual pianist whose reflection is actually playing the physically moving keys, thus encouraging anyone to take the seat left empty on the piano, feel the music as if in the doppelgänger’s shoes, and possibly playing along. Andante, for its part, depicts music phrases as miniature light silhouettes that walk and dance on the keyboard. The figures visually echo the expressive shape of musical lines, as well as the bodily sensation of a pianist’s fingers moving across the keys. Both projects act as immersive “sandboxes” for the understanding of music through the body and the playful exploration of musical ideas.

These projects envision what music learning could look like in the future, taking advantage of technologies likely to be pervasive in the coming years. They serve as vehicles to raise questions and demonstrate ideas about music learning through the observation of user interactions and descriptions of potential usage, or play, scenarios. To generalize beyond my specific projects, this thesis then presents two design frameworks for music learning to guide other creators of interactive technologies.

**Music Learning + Learning at Large**

The final portion of this thesis attempts to connect the dots between learning music and learning at large. The focus on correct notes in music parallels the focus on correct answers in school subjects, particularly in math.

To help children understand math beyond the correct answers, mathematician and educator Seymour Papert developed the Logo programming language to rethink how learners can be initiated to “feeling their ways” into geometry using their own knowledge about navigating in space. In substance, children use Logo to program the movements of a
“turtle” by teaching it (giving it instructions) to draw geometric shapes, either on the screen or in the physical world. Papert coined the term “body-syntonicity” to describe a strategy that helped children program the Logo turtle by imagining themselves moving in the place of the turtle. “Body-syntonicity” thus refers to a type of thinking that uses the knowledge and sense of one’s own body, and “syntonicity” literally means “in tune”. MirrorFugue and Andante, as well as the approach to music learning advocated by this thesis could therefore be described as “body-syntonic”.

It is important to recognize that “body-syntonicity” is not only about the body but also the connection between body-based knowledge, or “embodied cognition”, and other ways of knowing. To program the Logo turtle, children must learn to describe—i.e. translate—the imagined movements of their own bodies using a precise, formal, and quantitative language that can be “understood” by the computer. Similarly, learning music deeply means connecting many different ways of knowing, not only the embodied and the analytical but also the emotional and the interpersonal.

This thesis also draws on theories from three other influential thinkers to explore the parallels between learning music and learning at large. First is the psychologist Jean Piaget, who described stages of human cognitive development from the embodied and concrete thinking of early childhood to the formal and abstract thinking of adults. Second is Jerome Bruner, who articulated three types or modes of representations—enactive, iconic, and symbolic—which respectively correspond to thinking (or getting to know) through the body, in images, and in words and symbols. All three modes appear with the beginnings of the “symbolic function” starting at age 2, when human infants become able to evoke things at a distance, to engage in pretense or role play and to take their own “lived” experience as objects-to-think-with.

Last but not least is Marvin Minsky, whose concept of the mind as built from a society of smaller minds, or agencies, each responsible for different types of thought, or “intelligent” action-patterns (minded conducts) came in great part from his own experiences learning how to improvise on the piano. Minsky’s notion of thinking not only includes abstract, logical thought but also encompasses thought processes having to do with the body, the imagination, and emotions. Minsky always advocated the importance of connecting different parts of the mind, as exemplified in his
famous saying that “You don’t understand anything until you learn it more than one way”.

This thesis connects the dots between the ideas of Papert, Piaget, Bruner, and Minsky in the domain of music. I connect general cognitive development to musical development, drawing from the ideas of Dalcroze and from personal experience, to make clear the inextricable links between motion, emotion, imagination, and symbolic representation in music learning and beyond.

**Roadmap**

The main body of this thesis is divided into three parts, each corresponding roughly to one of the three overarching goals.

**Part I** examines the problems of music learning. Chapter 1 dissects the popular but counterproductive ideas people have about how to learn music and presents alternative perspectives. Chapter 2 reviews existing music learning technologies and identifies the space in which my projects seek to contribute. Chapter 3 traces the history of how popular ideas about music learning came to be, linking their onset to how technologies for music dissemination changed the cultural conception of music from expressive performance to disembodied sound.

**Part II** delves into the question of designing interactive technologies. Chapter 4 offers some background on the evolution of the piano and the computer, drawing parallels between the two domains and justifying my choice of platform. Chapters 5 and 6 then respectively detail the design, implementation, application scenarios, and evaluation of my two projects, MirrorFugue and Andante.

**Part III** generalizes the main ideas of this thesis beyond my specific projects. Chapter 7 presents a conceptual framework of my work based on Dalcroze Eurhythmics and connects the music learning ideas of Dalcroze to ideas about learning from Papert, Piaget, Bruner, Minsky, and others. Chapter 8 discusses strategies for learning at large derived from music. Chapter 9 imagines a vision for the future, presenting sketches for future versions of my piano learning projects and prototypes of other systems that explore similar ideas in other domains.

Finally, the **Appendix** offers three essays about music and the mind. They cover many of the same ideas as discussed in the main body of the thesis but offer a more personal perspective. **Appendix A** is an essay I wrote for
Joi Ito and Tenzin Priyadarshi’s *Principles of Awareness* class from fall 2014, which elaborates some of the ideas in Chapter 8 on the connection between music and meditation. Appendix B recounts an extended metaphor of learning as building “invisible cities” in the mind. The text and images have been adapted from a talk I gave at TEDxSan Diego in 2014. Appendix C reflects on Marvin Minsky’s relationship with music based on the moments we shared in the past year playing [with] the piano.
Chapter 1: The Pop-Ed Music Culture

Music is not easy. To learn it well takes a lot of time, a lot of focus, and a whole lot of love. That said, anyone can learn music, no matter how "naturally gifted" one may be, as long as one makes the time, finds the focus, and has a love for the art.

Of course, not everyone will choose to make the time to learn music if their interests are elsewhere. After all, learning anything well takes time and focus, and people naturally direct their greatest efforts to their greatest passions.

Unfortunately, many people who do love music and are willing to devote the time and focus feel that they cannot learn music. Some try without success while others give up before even trying. All of these people conclude at some point that they are just not "made" to learn music.

What leads to these sorts of conclusions which prevent people from learning music is a set of widespread, misguided views about what music learning entails and how to go about it. Seymour Papert referred to these sorts of ideas as Pop-Ed Culture. Below are 4 pop-ed ideas from the
world of music. I lay out each view, tease apart its problems, and offer alternative perspectives.

**Learning music = Learning Pieces**

For many people, learning music involves enrolling in lessons on an instrument (often the piano), where the goal is to play pieces of increasing difficulty. Pieces are usually written by one of a handful of famous classical composers, usually long dead, and written instructions for the player to follow.

Learning to play involves, on one hand, learning to decode the symbols of the written score, and on the other hand, developing the physical facility on the instrument to play what the composer has intended. To help with building this facility, students are often assigned a series of purely technical exercises, such as scales and arpeggios up and down the keyboard.

In many ways, the typical music lesson resembles how school typically teaches math\(^4\). Students are only taught to diligently follow other people’s rules and instructions, without the chance to make the math or the music their own. Just as schoolchildren are not shown how to think like mathematicians, those who take music lessons are rarely shown how to think like musicians.

One way to thinking like a musician is through music composition, which involves discovering, arranging, and inventing patterns, much like the thinking of certain mathematicians. But another, just as essential side of music is about real-time expression through the body, like an actor onstage. As with acting, the raw script may or may not originate with the performers themselves, but great performers, both actors and musicians alike, own the stage by infusing who they are into what they play, far beyond simply following instructions.

The performative side is another way to approach music creation—through improvisation. Just as improv actors can invent and express what they say and do in real-time, musicians can also extemporize, either solo or in “conversation” with others\(^5\). Today, the art of music improvisation is most well-known in jazz, but historically, learning to “speak” music was a core component of all music training, even in the Western “classical” tradition\(^6\). In fact, many of our greatest classical composers—Bach, Beethoven, Chopin—were skilled improvisers, and improvising, or “playing around” with music (on keyboards, no less!) was an essential
part of how they arrived at their musical ideas. We could even go so far to argue that many classical composers could not have composed what they did had they not played and played with music.

Thus, learning to “think” like a musician involves learning to play music and to play with music. Playing music is first about feeling music through the body. This involves knowing how to carry a beat, how to “sing” a phrase, and how to channel emotions through movement. Playing with music is not about memorizing music theory rules in the abstract but about building structures in the mind that embody the knowledge. It involves learning to hear underlying patterns when listening to any piece, imagining music as sound in the mind, and physically playing what one “hears” (with expression, of course!). We could think of it as what it feels like to be the music and to be inside the music.

**Learning Pieces = Learning Notes**

All good musicians (composers, performers, improvisers alike) have learned from other musicians, which includes learning to play and to play with the works of great composers. Studying existing music offers the opportunity to embody a wide range of emotions and characters, as well as a wealth of musical patterns to borrow, rework, and reuse.

Unfortunately, pop-ed ideas severely limit how people engage with the pieces they study. For most amateur students and their teachers, learning a piece is about learning to play the correct notes. After figuring out how to decode the score and how to produce various pitches on the instrument, practice becomes the process of “programming” the body to play the notes of a given piece according to instruction. One by one, notes are read from the score and played in sequence. After repeating the sequence enough times to achieve some fluency, the student then proceeds to add expression by following the score’s instructions for dynamics and articulation. To get an idea of how a piece is “supposed” to sound, plenty of audio recordings exist as references for “correct” interpretations. Aside from accidental mistakes (which must be fixed!), the student never strays from what is written, and nothing is said about how the music is put together beyond the score.

I recently gave an impromptu lesson to a young relative at a family reunion, who was working on a Chopin Ballade. Despite having studied the piano for over 10 years and despite being able to play through the
piece’s considerably difficult technical passages, my relative struggled to tell me what key the piece was in (g minor).

I know that my relative loves music. He spends time to practice every day and even transcribes the themes of his favorite video games to play on the piano. My story is not meant to undercut any of what he has achieved. Rather, it shows how the pervasive idea of playing music as “playing the correct notes” could limit even those with passion from reaching a deeper understanding of music beyond the notes.

What do I mean about understanding music beyond the notes? We could start to think about it through the intellectual, emotional, and physical components of any piece (composed or improvised). On the intellectual side, we could ask ourselves questions such as what are the main themes, where are the sections, and how do the themes unfold through the sections. This is not unlike analyzing an argument or a story. Of course, we could also get into questions specific to the language of music, such as what key the piece is in, when the key changes, and how the changes progress.

Emotionally, we want to probe what the piece is trying to say, which is not unlike reading poetry. We could ask questions such as what feelings or images the piece evokes, or what characters it conjures through the texture and shape of sounds in time.

Learning to play music also involves a physical component. This includes working out which techniques a piece uses (e.g. fast runs, octaves, trills) identifying which sections need more attention on physical coordination, and designing practice methods to gradually work up the techniques the piece requires.

A deep understanding of music means not only delving into each of these three components but also exploring the connections between them, such as how structure informs expression, and how expression informs technique. In the end, it is as Marvin Minsky once said—“you don’t understanding anything unless you learn it more than one way”\textsuperscript{17}. Learning music means knowing it in many ways and approaching it from many perspectives. And once you really know something, be it a math theorem or a music piece, it becomes your own.
Reflecting Music Through Movement

Practice = Repetition

Many of us have grown up with familiar sayings like “practice makes perfect!” or “the way to Carnegie Hall is to practice practice practice!” These sayings imply that to practice means to repeat, and if you spend enough time repeating a piece of music, you will one day be able to play it flawlessly. These sayings are, of course, rooted in the previously discussed pop-ed ideas, which equate learning music with learning to play pieces, and learning pieces with learning the notes.

When learning notes, pop-ed culture often assumes that the primary difficulty is the long process to cement muscle memory to accurately and consistently play the notes as written. Any glitch that occurs is assumed to be a result of imperfect muscle memory, and the remedy is always more rounds of repetition.

Spending so much time repeating easily becomes boring, a big reason why so many people quit. Those who do grind through years of repetitive practice often bear their experiences with a touch of pride. Many go on to become music teachers and proceed to instruct their own students on the merits of “no pain, no gain” — as if practice were something to diligently suffer through before reaping the rewards of playing.

But repetitive practice is not even an effective way to learn! It might suffice for learning to play the notes, but rote repetition never leads to a deeper, more connected understanding of music.

To see why, consider the following analogy from neuroscientist Alvaro Pascual-Leone, who compares how the brain behaves to a ski slope. Every time we think or do something, we ski down the slope and leave a track, which deepens the more times we repeat the thought or action. The deeper the track, the more likely we are to go down the same way we have gone before, but we also lose the ability to slightly alter our path each time or to change course mid-way. This is what happens from rote repetition. And if we somehow stumble out of such a track on the way down, it is near impossible to smoothly find our way back. (This is why so many classical performers are terrified of memory slips!)

Continuing the ski slope analogy, we could envision a better way to practice as covering the entire hill with a network of interconnected paths. This way, no performance is ever the same. We could choose which way to go at any given moment, even go off the beaten path to explore new parts of the slope. The network of paths for one piece might overlap and...
interconnect with the networks of other pieces and patterns. We could then speed up learning new pieces by building off existing paths and even create improvisations by exploring the paths of our entire hill, guided by what we feel in the moment.

Practicing to build such a network of paths takes time and is not easy. It is a lot to build up in the mind, which requires care and attention. But unlike mindless, repetitive practice, it is anything but boring! In fact, it could be fun if you know how to go about it.

Good music practice is exactly what Seymour Papert calls “hard fun”, which he demonstrates through the classic example of video games. According to Papert, people are not drawn to play video games because they are easy. In fact, the most captivating games are usually quite challenging, where players have to solve hard problems, think on their feet, and of course, learn in the process.

One property of successful video games is that their design helps players to stay in a “flow state”. Coined by Mihaly Csikszentmihalyi as a state of optimal engagement, flow occurs when the activity at hand just exceeds one’s current ability so that one is challenged without being frustrated, and fully engaged without room for boredom.

Knowing this, we could learn to practice in a way to maintain our own flow state. To do so, it helps to think of practice not as a chore to suffer through but as a game to play. Marvin Minsky described playing music as the temporal equivalent of playing with blocks in space. Practice could thus be seen as the game of building a tower with sounds in time. Like its physical counterpart, our musical tower needs a solid foundation. This means always playing with ease, expression, and holding a consistent pulse. To build, we gradually add more layers and more details without shaking our founding of pulse and phrasing.

To keep our minds engaged, we must often go beyond what is explicitly written. Given a physical challenge, we could avoid mindlessly repeating for the sake of muscle memory by inventing our own exercises that target the core of the problem. Inventing exercises (e.g. mini-games) to practice is also the key to making any music you play your very own.

**Key to Success = Talent**

Most people who follow the pop-ed approach do not reach a high level in music. Many can usually only play the few pieces that they have recently
practiced but not much else. Few can play by ear, improvise, compose, or even sight read very well, if at all, and what they can play does not sparkle and shine like the playing of the masters. From many amateurs’ perspectives, what the masters (performers, composers, and improvisers) can do seems like magic, where the masters are seen to possess special “talents”, mysterious qualities that cannot be explained.

The belief in talent as the magical ingredient for music learning comes from an assumption that the pop-ed method (learning pieces, learn notes, repeat) is the only way to learn music. If everyone followed the same method, the difference between those who reach great heights and those who do not must then logically lie in differences between people themselves. Marvin Minsky observes that this view allows people to “comfort themselves” and to “excuse their own deficiencies” by saying that experts simply “come endowed with all the qualities we don’t possess”.

However, the pop-ed approach is not the only way to learn music, and as I have demonstrated in this chapter, it is highly problematic—shallow, boring, and inefficient. Becoming a true musician—one who plays and plays with music, who understands it in many ways—requires building up and connecting so many different parts of the mind that no one comes endowed with all the knowledge structures ready-made. Differences in “natural-born” individual abilities certainly exist; some people might have have a better memory or more physical coordination. But no matter these differences, everyone still needs to build their own mental structures to understand music. As Minsky points out, what appears on the outside as “talent” is really a product of having “exceptionally effective ways to learn”. And luckily for us, effective ways of learning can be learned!

* * *

Alternate ways of learning music do exist but most are not very mainstream. The Suzuki method, one of the most well known, is often distorted by pop-ed culture to emphasize rote memorization, losing the spirit of its origin.

These alternate approaches also exist among experts, in classical and in jazz, but few have the chance of studying with the experts without first reaching a certain level of proficiency. Unfortunately, not many reach that level through pop-ed methods. Some experts have written books on how to approach their craft, but to seek out these books in the first place means already looking past the assumptions of the pop-ed ways.
To bring the art of music learning to a wider audience, I look to the power of interactive computing technologies. Just as Seymour Papert envisioned Logo as a way for anyone to learn how to think like a mathematician, I envision technologies for anyone to learn how to "think" like a musician—in the head, in the body, and in the heart.
Chapter 2: Taxonomy of Music Learning Systems

Many interactive technologies for music learning have been developed, both in academic research and the commercial sector. This chapter first reviews existing work in two broad categories: systems aimed to support learning to play music ("play-systems") and systems that enable playing with music ("play-with-systems"). I then identify where my work aims to contribute.

Learn to Play Music

A large number of music-learning systems aim to help students learn to play an instrument, where learning usually means learning to play the correct notes of existing pieces. In 1994, Casio first commercialized electronic keyboards with light-up keys that indicated the notes of a melody. The use of computers in music learning research was explored even earlier by Dannenberg et al. in 1989, who developed a system to train novices on the piano, with the goal of replacing the work of a teacher in the first year of instruction. Dannenberg’s system combined ideas of Computer Aided Instruction (CAI) and Intelligent Tutoring Systems (ITS). The CAI portion presented a series of lessons to the student, and the ITS portion gave feedback on student’s playing. Feedback involved pointing out mistakes and asking students to try again, which then resulted in either remedial instruction or the assignment of new lessons to complete.

Some combination of CAI and ITS has become a popular framework for play-systems. A new wave of interest in using the computer to teach music followed the 2005 release of the hit game Guitar Hero. For the piano, projects with more CAI focus include the work of Kitamura, and
Mukai\textsuperscript{29}, both which use the computer to generate phrases for student to practice. On the ITS side, projects from Shirmohammadi\textsuperscript{30}, Goebl\textsuperscript{31}, and Akinaga\textsuperscript{32} all record and analyze MIDI data from students, where MIDI captures note strikes, releases velocity, and pedaling. More recently, Takagawa\textsuperscript{33} as well as Rogers\textsuperscript{34} brought previously developed ideas of CAI and ITS from the computer screen to the physical world, where descending blocks indicating notes of a piece are projected on and around an electronic keyboard. The projections enable students to more easily observe and practice along with the notes of a piece. The systems are also capable of giving feedback on student playing, using the representation of scrolling blocks to point out inaccuracies of pitch and timing.

One commercial game for piano training is Synthesia, which analyze input from a MIDI keyboard according to whether users play in time to the falling blocks\textsuperscript{35}. More mature versions of CAI + ITS tools have been commercially released by Yousician\textsuperscript{36} (for guitar and piano) and by JoyTunes\textsuperscript{37} (for piano) as mobile and browser-based apps. These apps feature progressive lessons that show student how to play different notes (or chords on guitar) and how to read notation, as well as a feedback algorithm that grades students on accuracy of notes and rhythm. Both provide stimulating game environments and offer an element of social playing. Positive reviews from users suggest that these apps are useful for helping novices build a basic facility on an instrument and for learning to read notation (or tablature)\textsuperscript{38}.

But playing music is not just about hitting the right notes at the right times. Since music is played through the body, student need to develop a sense of how to use their bodies to make music. Video tutorials, such as recordings of master classes, have long existed as references for the curious. In recent years, video tutorials have become more accessible than ever, thanks to Youtube, which has enabled musicians from all over the world to upload videos of their own playing and demonstrations\textsuperscript{39}.
Research on new methods to “instruct the body” to play the piano includes the use of virtual hands that demonstrate correct fingering and wearable devices that vibrate each fingering based on a sequence of notes to play. In contrast, earlier versions of my own project MirrorFugue considers not only the sequence of fingers to use but also how the entire hand moves during playing. First developed in 2009, the first version of MirrorFugue projected video of a remote teacher’s hands on the fallboard of a student’s piano. Preliminary studies demonstrated that seeing a teacher’s hands aligned and to-scale with one’s own hands not only helped beginners learn to play melodies faster and more accurately but also helped them learn how to move their hands like their teacher. In 2011, a similar idea was commercialized with the Ion Piano Apprentice, which added the video of a teacher’s hands on an iPad to a miniature electronic keyboard with light-up keys.

Another side of learning to use the body is understanding your own body’s movements. A 2003 article in a piano teacher’s periodical called for instructors to use video cameras during lessons to give students feedback on their movements after they play. Researchers have since developed systems that use motion capture and sensors to capture and analyze students’ movements for piano as well as other instruments. Feedback has been provided post-performance through video, 3D models, and comparisons between student and master playing, as well as in real-time through sonification.

Learning to play expressively is one of the least explored topics of existing play-systems. Though researchers of play-systems have long acknowledged the importance of playing with expression, most CAI and ITS have a rather limited view of “expression” as following instructions for dynamics and articulation. One approach has attempted to graphically “represent” expression for students to follow, and to evaluate student’s playing on “accuracy” of expression. Other researchers have claimed that knowing how to shape melodies or how to play with a strong sense of rhythm are “advanced” skills outside of the scope of beginners, who should only focus on playing the correct notes.

An indirect approach to develop expression is via auto-accompaniment, which supplies a musical backdrop that could help learners get into the mood of what they play. Auto-accompaniment systems have existed since the 1970s, with electric organs that came equipped with drum machine and common chord progressions. Casio release its first electronic keyboard with auto-accompaniment in 1981. More recently, researcher...
at IRCAM have developed auto-accompaniment systems capable of following a human performer’s timing. A similar technology has been commercialized into a practice app for mobile devices called Cadenza, which was released in 2013 by Sonation.

A different approach to accompaniment was proposed by Oshima in 2002, through a system designed to enable parents without any musical background to accompany their children during practice. Parents can play any note on the piano to determine the expressive timing of a predetermined accompaniment track. This same idea was used in another practice tool by the same researchers and is featured in one mode of the commercially available Smule Magic Piano app.

The idea of “simplifying” what is played to focus on expression has been more recently explored by the Magic Instruments Guitar, which replaces strings with a grid of buttons for different chords. The buttons simplify hand positions to play each cord to one note to press so that a novice could focus on practicing the rhythm and expressive timing of their strums.

**Learning through Playing with Music**

While technologies that support learning to play music generally follow the approaches of Computer Aided Instruction and Intelligent Tutoring Systems, technologies that enable playing around with music tend to follow a Constructionist learning philosophy. These systems do not attempt to instruct students on what is “correct” or to correct their “mistakes”. Instead, they seek to promote learning through self-guided exploration, experimentation, and play.

Constructionist music toys are designed to give children positive early experiences playing with music and to lower the floor to creative musical participation. Such toys include MaKey MaKey, Singing-fingers, Drawdio, and the LittleBits Korg Synthesizer Kit. While these project make it fun to experiment with sound and help children learn in the process, it is important to make the distinction between learning in general and learning about music. For some projects, such as MaKey MaKey and the LittleBits Synthesizer Kit, music is used to engage children in tinkering and making activities. Children might learn about instrument design or the science of sound through these activities, but less about how to play music or how music works from the perspective of a musician.
Other musical toys are more deliberately designed for children to learn about music, such as the Toy Symphony projects from Tod Machover’s group at the MIT Media Lab. These includes the BeatBug, a hand-held instrument shaped like a bug that enables a user to record a tapped rhythm and explore different timbres, pitches, and phrasing based on spatial manipulation of the bug and its “tentacles”. A network of BeatBugs allow multiple people to collaborative create and share layered rhythms. Another from the Toy Symphony series, Music Shapers, enables the user to control the contour, timbre, density, and structure through a soft, squeezable interface. A more recent project, SoundStrand enables users to learn about musical structure by manipulating physical blocks that represent pre-composed musical segments to create variation in harmonic and rhythmic content.

Constructionist environments have also been developed for music composition, offering rich visual metaphors to help users make their own music by playing with patterns. In Impromptu, users can experiment with creating melodies by putting together “tuneblocks”, which are graphical blocks that represent short melodic fragments. Hyperscore enables users to define the shape and overall form of a piece by drawing lines on a grid representing pitch and time. Lines are automatically converted to music using rules of Western harmony built-into the program, thus enabling users to focus on higher-level creative expression. Melody Morph is an iPad app that allows users to create their own compositions/instruments by spatially groups of “bells” with different pitches and timbres, which could be played by tapping.

* * *

With few exceptions, existing systems for music learning could be broadly categorized into those that support playing music and those that support playing with music.

Play-systems have focused on teaching and assessing students on the skill of playing a traditional instrument based on the assumptions of pop-ed music culture. Namely, playing music is seen as playing notes according to instructions; the body’s role is mainly one of mechanics and muscle memory; and expression, when considered, is often seen as an extension of following directions.

In contrast, play-with systems encourage open ended exploration with different aspects of music and learning through creative experimentation and improvisation. However, there tends to be a sharp divide between
systems that support physical playing (on custom, newly designed instruments) and systems that support composition (on a computer screen). Moreover, though these systems lower the barrier of entry (i.e. lower the floor) to musical creativity, they don't necessarily pave a clear path for higher levels of expertise (i.e. allow for a high ceiling or wide walls).

My work seeks to bridge the gap between the play and play-with systems. To do so, I explore how systems could support learning to play an instrument beyond the notes, and how to play with music on an instrument. Additionally, I wish to explore how different elements of music could be learned, neither in isolation, nor on custom one-off interfaces, but within the context one instrument. Ultimately, my goal is not only to lower the floor but also to provide clear paths toward both an infinitely high ceiling and infinitely wide walls. For insights on how to design such systems, the next chapter looks into history at how problems of music learning came to be.
Chapter 3: [Technological] Origins of Pop-Ed Music

As discussed in chapter one, the pop-ed culture of music assumes that learning music means learning from scores written by composers, often on the piano, and occasionally consulting audio recordings as references. Many people take these methods and tools for granted, but it is interesting to note that the recording, the piano, and even musical scores have not always existed. All three are technologies that were invented to extend human capabilities of making and sharing music.

According to Marshall McLuhan, any widespread extension of ourselves (which he calls “medium”, plural “media”), results in profound personal and social consequences. McLuhan’s seminal book Understanding Media, from which came the famous saying “the medium is the message”, analyses the consequences of technologies throughout history. Examples include not only inventions that readily come to mind as “technology” (e.g. the radio, the airplane, and the television) but also inventions of the past that are no longer seen as “technological” (e.g. roads, money, and numbers), which have nevertheless left profound marks on individual and cultural perception, understanding, and appreciation.

Following McLuhan’s approach this chapter puts forth a theory that technologies for music dissemination in the West—namely, modern music notation, the piano, and audio recordings—gradually changed people’s
notions of music from expressive performance to disembodied sound. I argue that this change was the seed of pop-ed music culture.

**The Birth of the Composer**

The development of modern notation differentiated the West from other musical traditions. Historian Craig Wright observes that notation made the West an ‘odd culture out’, owing to its adoption of a symbolic and quantitative representation. Other cultures described music differently using writing—if they did so at all. The Japanese, for example, employed characters to describe the gestures of their shakuhachi, while the Chinese referenced animal metaphors when notating gu-qin technique.

Interestingly, Western music notation did begin as a description of gesture, with neither precise pitch nor rhythm. To aid recall, medieval sacred music employed neumes (from the ancient Greek pneuma) describing patterns of breath. These descriptions grew to incorporate relative heights between neumes, which described general melodic contours, or patterns of breath. More quantitative specifications of pitch did not come until the development of staff lines and clefs, around the 12th century. By the late 13th century, the same symbols for pitch began to be used to indicate rhythm on a metered grid so as to ensure voice alignment in the increasingly complex polyphonic arrangements of sacred songs.

The precise indication of both pitch and rhythm radically transformed Western notation from shorthands to evoke embodied memories of songs one already knows, to a system of abstract symbols from which one may interpret sounds one has never heard. Without the precise notation of pitch and rhythm, songs were taught and learned primarily as an oral tradition through the players’ and audience’s direct engagement and whole body-participation. Demonstration and imitation transmitted melodic contour along with expressive shape, and the ambiguity within notation not only allowed but encouraged variation between performances. In fact, improvisation was both commonly practiced and highly valued. Musicians aspired to authentic expression and creative ingenuity, for, in an oral tradition, a successful performance was simply one that ‘spoke’.

The emergence of notation enabled a new and distinct breed of musician: the composer. Like an architect, the composer could arrange the entire blueprint of a piece, from the pitch and duration of a single note to entire sections of motivic development. Against detailed blueprints, playing
could then be evaluated for accuracy of execution, thus introducing a new
criterion for performance: correctness. Craig Wright attributes the
eventual ‘decline of the performer-improviser’ to the ‘advent of the
composer’; but the decisive shift in the role-division did not take place
until well into the 19th century, when modern notation came into contact
with the mass adoption of a new technology—the piano.

The Machine and its Instructions

Of all musical instruments, the piano and its keyboard cousins stand apart
as mechanical contraptions. Other instruments feature a constant direct
connection between human gesture and sound. The woodwind and brass
families are played with breath, while string instruments are plucked or
bowed. In contrast, the keyboard operates as an artificial interface with a
sophisticated back-end machinery. Keyboards introduced a degree of
abstraction, with keys mapped to a ‘linear’ arrangement of discrete
pitches, where adjacent keys differ in pitch by a chromatic ‘half-step’. This
design enabled a single person to play more complex music than ever
possible before, but severed the direct expressive link between the body’s
gesture and the instrument’s sound.

The piano in particular uses a hammer mechanism to strike its strings,
allowing for a full spectrum of volumes from soft to loud. For this reason,
it was originally called the pianoforte, from the Italian dynamic markings
piano (soft) and forte (loud). The pianoforte improved on its precursor
the harpsichord, which had no capacity for varying dynamics, but
compromised on the continuous sound control of the clavichord in
exchange for a wide range volumes. The hammer mechanism limits
control of sound to the initial moment of key strike, making it uniquely
easy to play the literal, correct notes of a score with little attention to their
expressive shape. Recall that notation was originally meant for the human
voice, which flows between each note with natural expression. Attaining
the same fluidity of expression on the piano above and beyond the score’s
explicit indications requires careful listening, a clear imagination of
sounds, and the capacity to translate the imagined sounds into physical
movements on the instrument.

In the century that followed the piano’s debut, keyboardists gradually
discovered how to make music with this new technology, informing
instrument-makers as they gradually refined their designs. In fact, many of
the greatest keyboardists were also composers, who often developed new
techniques and patterns through improvisation, which they then notated.
and published as scores. In the early days, not many people had access to a piano, such access being limited mostly to families of musicians, and to the upper classes that could hire them. Thus knowledge of how to listen, to play, and to invent on this new instrument passed from person to person in parallel with the distribution of written compositions.

By the mid-19th century, the design of the piano, as well as the expansion of its body of repertoire, had reached a certain maturity, coinciding with the rise of the European middle class. With their newly disposable income and leisure time, middle-class families flocked to obtain pianos for their homes, and lessons for their children. The piano’s versatility as a general-purpose instrument made it the way for people to experience music outside the concert hall. Along with the surge in piano sales, sheet music became a booming business. Publishers offered piano transcriptions for each new opera and symphony, as well as many solo pieces for entertainment in the parlor.

The popularity of sheet music further widened the gap between performer and composer, and attitudes toward the written score underwent a decisive shift. Previously, notation had served as guideline rather than as command. Performers commonly took liberties with the composer’s score to suit their own needs, and composers explicitly left cadenzas unwritten, allowing the soloist room to invent. Beethoven began writing out complete cadenzas, and thereafter instrumentalists habitually played composed cadenzas from memory. At the same time, the rise of the music-academia reinforced the new relationship between performer and composer. Historian Richard Taruskin describes conservatories that arose to train musicians to ‘reproduce the letter of text with a perfection no one had ever previously aspired to’ and where improvisation was ‘scorned outright’63. Musicologists, particularly of the Austro-German school, championed the notion of the score as High Art, akin to a ‘work’ of literature.

The late 19th century solidified the familiar paradigm of the composer-creator and the performer-interpreter. Performers gradually abandoned the improvisation of new material on stage, but their interpretations were far from calcified, as they often are today. The present culture of classical performance and its obsession with literal correctness did not emerge until well into the 20th century. Its onset may be linked to the widespread adoption of audio-recording technology, which provided a medium to disseminate “ideal” disseminations of the classical canon.
The Age of Disembodied Sound

Invented in 1877, Thomas Edison's phonograph was originally envisioned for speech reproduction, with applications for dictation in the office and for recording the last words of the dying. The inventor of the gramophone, an early competitor to the phonograph, began to market music discs in 1889. Henceforth, recorded music grew in popularity, propelled by the introduction of radio broadcasting in the 20th century. Audio recording transformed how people interacted with music, as well as the changing culture of performance. Before, music could only be heard live, with the piano recital being a popular concert-format since the days of Franz Liszt. Great concert pianists were celebrities, winning the hearts of audiences with their expression and exuberance. Personal voice and showmanship counted for more than the occasional wrong notes, which were readily forgiven and forgotten.

Recording technology enabled performances to be captured, bought and replayed anywhere—and at any time. It also enabled the career of a new breed of performer: the recording artist. According to Glenn Gould, who exemplified this new breed of pianists, audio recordings presented music in its purest form as crystallized sound: free from the 'theatrics' of the stage. In other words, the recording afforded the listener a new perspective into a piece of music. It stripped away the presence of the performer to focus attention purely on the sound, which can be replayed again and again. Audio recording enabled musicians to craft the 'perfect' performance, eliminating all mistakes by splicing together the best of several takes. In this new medium, these corrections are desirable and often necessary, as repetition amplifies the presence of even minor blemishes.

Unfortunately, the same preoccupation with eliminating all errors has also spread into concert-playing. As recording enabled the rapid distribution of performances across borders, they began to assume the mantle of a de facto 'standard' embodying how pieces should sound. Audiences, habituated to the familiar, flawless sounds of their records, have also grown to expect the same flawless execution live.

These trends marked the onset of our current preoccupation with correctness: the fixation with faithfully reproducing a 'perfect' performance as heard in recordings and instructed by score. Rather than listening deeply to the quality of the sound, most amateurs prioritize surface flaws as the most obvious difference between their playing and...
that of the recorded standard. Among those who make their way into the ranks of the professionals, many are too afraid to deviate away from the accepted interpretations that may be heard in the great recordings. Harold Schoenberg notes that there is now a ‘deadening uniformity’ in today’s piano-playing, such that it is ‘next to impossible to distinguish Pianist A from Pianist B from Pianist C’.

Technologies for Possibilities

This chapter traced how the once-inextricable links between performer, instrument, and sound gradually became severed due to historic technologies for music dissemination. Though the unintended consequence of this cultural change became the seed of our current problems with music learning, these same technologies also opened new doors, new experiences, and new creative possibilities.

The score created a new medium in which to communicate more intricate music than ever before. Entire four-voice fugues could have existed in Bach’s mind, but if he wasn’t able to write them down, we would not be able to play them today, four hundred years later, and continue to find inspiration from them. The piano, too, enabled greater sharing than was ever before possible. In a way, it was the “desktop-workstation” of its day—a general purpose device with which to explore, learn, and perform an infinite array of music. Just as the personal computer democratized computing from the labs of the technical elite, the piano democratized “art music” from the salons of the aristocracy to almost every middle-class home. More recently, audio recordings have enabled anyone to hear music from every corner of the world, anytime, just with the touch of a button.

The next chapters will combine emerging paradigms from human computer interaction with the realm of music. Beyond targeting the root problems of music learning, I will also explore the potentials of my designs for expressive music making and sharing.
II: Designing Interactive Technologies

Chapter 4: The Piano as a Platform for Invention

Over the course of my PhD, I spent much time sitting between my computer’s keyboard and the keyboard of a grand piano. By day, I would write code on my computer for my projects. By night, I would turn around to play the piano.

My research brings these two worlds together into an interactive platform for a digitally controlled, augmented player piano. The acoustic sound and moving keys of the piano, overlaid with projected pixels, engage the audio, visual, and haptic modalities, resulting in a sensorially rich experience firmly rooted in the physical world.

On my platform, I envisioned, implemented, and explored two main projects, MirrorFugue and Andante, both with applications for learning and artistic creation. Before bringing my projects to the stage, I first wish to illuminate a background from the world of the piano, and from the world of computing technologies.

In broad strokes, the piano side paints an overview of new developments, both technical and artistic, focusing on the 20th century to the present day. On the side of computing, I recount a brief history of modern-day computer technologies and summarize key ideas about the future of computing that brings together the physical and the digital. This thesis explores the convergence of a rich artistic tradition with future visions of digital, interactive technologies. It examines how music could bring more
depth and meaning to interactive systems and how emerging technologies could bring fresh perspectives to an established art.

**Piano**

The early years of the 20th century marked the heyday of the piano in Europe and America. The 19th century had brought a series of technical improvement to the instrument, including felt-covered hammers, the cast iron frame, and the technique of overstringing, where two rows of strings are crossed in the piano's interior to bring out overtones in the higher registers.

![Action of a modern grand piano from Steinway & Sons](image)

Before the radios and records, the piano was still the primary way for families to hear music in their homes. For those who could not play well, the player piano and widespread availability of piano rolls offered a way to hear the performance of virtuosos outside the concert hall, albeit at a much-reduced quality.

I want to emphasize here not just the popularity of the piano but how the instrument became a vehicle for artistic invention and ingenuity. Harold Schonberg points out that not all pianists are composers, but with few exceptions (e.g. Wagner, Verdi, Berlioz) all of our great composers of the past were pianists (or, like J. S. Bach, played the keyboard before the piano existed). Frédéric Chopin, for instance, not only wrote hauntingly beautiful music but also made significant contributions to piano technique, inventing new ways of physically calling forth never before heard qualities of sound. Chopin did not think of technique as mechanical exercises to be mindlessly repeated as we do today. His collections of Études, or study pieces, demonstrated new ways of playing each within a meaningful, musical context.

![Inside of contemporary concert grand](image)

Until the latter part of the 19th century, composer-pianists frequently performed, and even improvised in public. Beethoven made his name as an improviser before being immortalized as a composer, and Franz Liszt
was the first of several generations of touring virtuosos, the precursor to the modern “rock star”. By the start of the 20th century, composers were glad to leave concertizing to the professional performers, for they wished to travel in new lands beyond the established territories of Western melody and harmony. The piano became an important vehicle for their explorations.

The piano enabled Debussy, Ravel and others in France to experiment with new sonorities through the use of new scales and harmonies. Inspired by the Symbolist literary movement, their music seem to evoke the ineffable, and to reverberate that which cannot be directly expressed. Debussy imagined a piano without hammers, and invented new pedaling techniques to blend individual notes into hazy washes of colors from which emerged other-worldly scenes. In Vienna, the layout of the piano became the grounds on which the Serialists like Schoenberg, Berg and Webern invented mathematical methods to generate new sonorities. In central and eastern Europe, Bartók, Stravinsky, Prokofiev also explored new harmonic constructions on the piano, but for them the piano was percussive, a true hammerclavier, to be played with a precise tone and propulsive rhythms.

Like Chopin, Bartók also wrote study pieces to introduce his harmonic and pianistic ideas, which are compiled in his six progressive volumes called Mikrokosmos. In an introduction to the first book, Bartók’s son Peter described how his father encouraged him to improvise with different patterns on the piano as a way to learn to play. In fact, Mikrokosmos began as a compilation of these patterns. These compositions were not meant as compositions to be repeated and memorized but rather as micro-words(!) in which to experiment and to explore.

Meanwhile, in America, great strides were made in the beginning of Jazz among pianists like Jelly Roll Morton and Fats Waller in New Orleans. What started out as dance music in nightclubs of the sultry city began a new art form, combining classical harmonies and swinging rhythms, with an emphasis on improvisation as a means for creation.

After World War II, avant-garde composers like Stockhausen, Boulez, and Berio began creating the first music for electronic instruments but continued the serialist tradition on the piano, inventing new composition techniques involving numbers, proportions, and an element of chance. Others explored new extremes of physical playing techniques such as plucking strings, knocking on the wood, and prepared piano, where...
different objects are placed on and between the strings to create new sounds. Busotti required his performer to disassemble a piano during a piece, and the pianist of John Cage’s 4 minutes 33 seconds plays nothing at all for that duration of time.

At the same time, Jazz continued to make inroads, pushing far beyond classical harmony, and innovating in rhythmic complexities through improvisation. The piano became a mainstay instrument in jazz trios, quartets, quintets, and beyond, where pianists such as Thelonius Monk, Bill Evans, Oscar Peterson, and Herbie Hancock and many, many others each created their own distinct style.

The 20th century expanded to new horizons of musical making thanks to the power of electricity. Invented in 1920, the eponymous instrument of Leon Theremin demonstrated the potential of oscillator circuits to produce music. Freed from the constraints of mechanical sound production, electronic instruments could adopt any interface for players to control the sound-producing circuits. The theremin used in-air movements of both hands to control the frequency and amplitude of its oscillators, but it was notoriously hard to play and master. Due to the ubiquity, versatility, and its well-established body of playing techniques the piano keyboard became the interface to a number of new electronic instruments, including the Rhodes piano, the Hammond Organ, and the synthesizers of Robert Moog. These new instruments lent their distinctive sounds to popular music of the 60s, 70s and 80s.

In the present day, the piano and its keyboard cousins remain at the top of the world’s most popular instruments. All of the music I have describe continue to be studied and continue to inspire new generations of musicians. New synthesizers and electronic music controllers are still designed based on the piano keyboard. Meanwhile, the traditional acoustic piano is still a popular choice around the world, especially in China, where it has undergone a boom in popularity. Just in 2015, the premier piano manufacturer Steinway and Sons released a new, high-resolution player piano, Spirio, which is capable of digitally reproducing the precise touch of a pianist’s performance.

From this overview, we could draw two conclusions about the piano as a platform for design. The most obvious is that even though we are no longer in the piano’s heyday where everyone owned one and learn to play it, the piano is still very relevant as a musical instrument, and it is thus worthwhile to design tools that help people learn it. Learning the piano

Figure 4.5: A pioneer builder of keyboard-based electronic instruments is composer/pianist Raymond Scott, shown here with his Clavivox.

Figure 4.6: The Roli Seaboard: a “piano” with soft keys that allows continuous sound control like note bends and vibratos.
means opening the doors to many genres of music, plus the potential for skill transfer to any other instrument in the keyboard family.

A more significant conclusion is that pianos in particular (and keyboards in general) open new doors to musical invention. Why were so many great composers pianists? Because the piano was the ideal tool to experiment and explore by making concrete the patterns of sounds in their imagination. Playing with sound is not just about rearranging patterns in the purely mathematical sense. As Jeanne Bamburger’s work with MusicLogo demonstrates, anyone can make patterns that sound somewhat interesting and make logical sense. What differentiates the great musicians—composers, improvisers, and performers—is not the patterns they can write or play but the relationships between their body, and their instrument, the sounds they create, as well as their imagination.

Though we receive pieces as “notes” on a page, music does not exist as notes in the mind of a musician, but as sounds connected with the movements to create them. Through physical movements on the keyboard, composers discover different qualities and patterns of sound, which they can then imagine and arrange into their compositions. Physically hearing the sounds allow composers to feel their music “resonate” in their own bodies, and through empathy, also feel how the music will resonate in the body of the listener. This process of music creation is in fact not so different from how a great performer practices, finding the movements to produce sounds that resonate their own bodies in syntony with the feelings they wish to express, and making adjustments based on what they hear.

Recall my argument from Chapter 3, that not understanding the connection between movements of the body on the instrument and the sound led many amateur pianists to believe that playing music is correctly playing notes that a composer has written. This chapter reveals another piece of the puzzle—those who become composers and improvisers, even after the rise of pop-ed music culture, were those who understood the link between the body, instrument, sound, and the imagination. And for many composers, the instrument of choice, the one that best enables them to explore and to express musical ideas, has been the piano.

Therefore, my wish to help people learn to play the piano is not only because the piano is widely played. By learning the piano, I do not just mean learning to play existing pieces. Rather, I wish to harness interactive technologies to help anyone who sits down at a keyboard to understand
the relationship between their own body, their instrument, the sounds they create, and their imagination, so that they too can explore and express musical ideas as our great composers and improvisers have for over 300 years. In other words, I wish to bring the “keys” of musical invention, from the world of the elite and the “talented” to everyone who has a love for music.

Computing

The word computer comes from the verb “to compute”, indicating its origin as a tool for calculation. The earliest mechanical tools for calculation were pioneered by Wilhelm Schickard, Blaise Pascal, Gottfried Wilhelm Leibniz and others in the 1600s, when machines for music-making like the harpsichord and the clavichord were already played all across Europe. The first commercially calculating machine, the arithmometer, was an updated version of Leibniz’s mechanical multiplier and was not produced until the 1820s.

The idea for a programmable calculating machine, the precursor of our modern-day programmable computers, was conceived by Charles Babbage in 1834 and was known as the Analytical Engine. Though a fully functioning prototype of the Analytical Engine was never built, we do have a detailed description of how the machine was to function based on Babbage’s own description and extensive notes from his friend and collaborator Ada Lovelace. Like the Jacquard loom (and later the player piano), programs for the analytical engine were stored on punched paper. Individual cards were chained together with ribbon, and the machine was to read each card and execute its instruction, which sometimes involved skipping forward and looping back to various cards on the chain.

In the 1930s, a renewed effort was put forth into building working machines for complex calculations. At MIT, Vannevar Bush designed and implemented a mechanical differential analyzer, which was extensively used during World War II for solving ballistics-related differential equations. Also in the 1930s, IBM began to manufacture a series of commercial machines that performed calculations and punched results onto paper cards. These machines combined Herman Hollerith’s punched card tabulation idea with programmability via plugboards, thus realizing the vision of Charles Babbage.

In Chinese, there exist two different words for “computer”. One is ji-suan-ji, or “calculating machine”. The other is dian-nao, which literally
means “electronic brain”. The machines that we call computers today seem worlds apart from the calculating devices of yesteryear. But beneath everything we do on the computer—word processing, web browsing, 3D-modeling, music-playing, video-conferencing, and much, much more—these “electronic brains” are still simply calculating.

Every capability of the computer could ultimately be reduced to manipulations of numbers, otherwise known as data or information. Behind the scenes, all computers only store data, perform operations on data, transmit data, and receive data—all based on the functions we program. Modern computers are capable of high-level functionality based on low-level calculations (or computation) thanks to the foundational work of mathematicians such as Alan Turing, Alonzo Church, Claude Shannon, and John von Neumann in the 1940s and 50s. A detailed description of how their theories enabled modern-day computing is beyond the scope of this thesis. Rather, I will describe the implications of modern-day computing through three key concepts: digital representation, connectivity, and interactivity.

The modern-day computer is a digital machine\(^\text{80}\). The word “digital” denotes an encoding of numbers in binary, or base 2. Numbers that we use for everyday calculations are in decimal notation, or base 10, where the position of each digit has ten possible values (from 0 to 9). In base 2, each position could only take on two values, typically represented by 0 or 1. Any decimal number could be converted into binary. For example, the number two is 01, ten is 110, and twenty-one is 10101.

The advantage of binary is that it could be easily represented with electricity. A binary number becomes a sequence of bits, where each bit could mean on or off, high or low voltage. Information encoded as bits could be manipulated by electronic circuits to produce new results, which could then be transmitted or stored. Thanks to billions of transistors inside, which act as tiny switches that could take on binary states, the modern-day computer is able to perform massive amounts of calculations at very high speeds. My laptop has 16 gigabytes of working memory (a byte = 8 bits, and giga = \(10^9\)) and a 2.6 gigahertz operating frequency, which means the ability to flip a single bit \(2.6 \times 10^9\) times per second when running programs to calculate!

The bit is the basic unit of information for the computer. Any physical phenomenon may be digitized by breaking it down into minuscule subdivisions and assigning it a value in binary. For example, a picture
may be recorded as a two-dimensional grid by specifying the color value of each point. A sound wave may be recorded and reconstructed by taking samples in time and assigning a digital value to each sample. A video is simply many frames of images, which could each be digitized.

Before the digital age, a physical substrate was required for the capture, storage and transmission of information. In the case of music, patterns of sound were first notated as symbols on paper. Later, technologies enabled recordings of performances on piano rolls, where notes were punched into holes on paper, and audio discs, where sound waves were etched onto vinyl. In contrast, a digital music file is simply a sequence of binary numbers encoding samples of the original sound-wave, which may be easily accessed, stored, and copied on computers, external disks, and in the “cloud”.

With the rise of the World Wide Web, which connects vast networks of computers around the world, we can now instantaneously transmit and receive bits of information from anywhere, at any time. The confluence of digital representation with the connectivity of the web has profoundly changed how people access, create, organize, and share information. These changes are especially profound for the world of music. Streaming services like Spotify enable public access of a huge collection of audio recordings from every genre. Youtube houses videos of various artists, from amateurs to stars, as well as video tutorials from teachers around the world. On IMSLP, virtually score from the Western classical canon is archived and freely available.

Thanks to the power of digital representation and connectivity, the computer has become the ultimate protean tool: a general-purpose device capable of representing, manipulating, storing, and transmitting anything that may be converted to bits. Exactly how the computer could and should be used remains an open area of research. In each moment, there is almost infinite choice in what may be presented to the user and what the user might do. These questions are considered by the field of Human Computer Interaction.

Currently, most of what we consider to be “computers”— desktops, laptops, tablets— follow the interaction paradigm of the Graphical User Interface (GUI). Information is presented predominantly as images and text in rectangular screens. The user may select and move objects with a mouse and input text through a keyboard. The GUI has enabled a wide range of tasks on the computer, but it is sensorially impoverished
compared to interactions in the physical world. Attention is focused primarily on visual information in a two-dimensional viewport, and user input is limited to a small set of repetitive hand movements.

But computers need not follow always follow the familiar form factors of keyboard, mouse, and screen. In fact, movements within Human Computer Interaction research have increasingly called for an expansion of bodily engagement, and for computing to move beyond the screen into the physical world. As early as the 1980's, Myron Kruger pioneered full-body interactions with life-sized silhouettes of human-figures projected on the wall. The field of Augmented Reality (AR) exploits the innately human understanding of space by overlaying information directly onto the physical world. Though AR seeks to bring information into the physical world, its most common implementations focus more on the availability of information and less on its physicality. Common techniques for AR include head-mounted displays and using mobile devices (e.g. phones and tablets) as a window through which to see virtual information superimposed onto the physical world.

In contrast, the vision of Tangible User Interfaces (TUI) focuses more on the direct manipulation of physical objects, taking more advantage of innately human sensory-motor capabilities. Three key concepts from TUI research have been particularly inspiring for my work. First is the idea of I/O coincidence, where the physical world instead of the screen is the locus of interaction. Physical objects serve as input handles for the underlying digital model, and information is displayed on or around the objects via projection. An example of this idea could be found in URP, an urban planning workbench, where users move around physical models of buildings, which update the shadow and wind simulations projected onto the scene.
MusicBottles and I/O Brush stand for another branch of TUI research, which imbues everyday physical objects with new capabilities. MusicBottles gives the illusion of sound stored in glass bottles, like perfume. Lifting the stoppers of a bottle releases the recording resting within. With the slogan “world as palette”, I/O Brush enables users to “paint” with colors and textures of the physical world. These projects take advantage of familiar gestures associated with the objects’ common usage to create surprising and delightful interactions that merge the physical and digital.

Research in TUI has also explored actuated objects as a modality of interaction. Actuated interfaces allows both input and output as physical movement, which is more engaging than pixels alone. Recent research has focused on the potential of pin-based shaped displays as a workbench to dynamically simulate forms, movement, and material properties.

Interestingly, we could regard the player piano as the world’s very first “actuated user interface”. Present-day player pianos such as the Yamaha Disklavier and the high resolution Steinway Spirio are fully digitized and are able to stream archived performances over the Internet. Digital player pianos are an ideal embodiment of Mark Weiser’s vision of “Ubiquitous Computing”, where computers integrate invisibly into everyday objects and integrate seamlessly with our everyday lives in the physical world.

Inspired by research from tangible interfaces, my work seeks to augment the digital player piano into a platform for music learning and creation. The next chapters detail my two main projects, MirrorFugue and Andante.
Chapter 5: MirrorFugue: Conjuring the Recorded Pianist

The central question behind my work asks how interactive environments can help people understand how to “think” like a musician, where “thinking” includes not just the “mental” activities of arranging patterns but also the more concrete, body-syntonic forms of thought. Specifically, I see great potential for interactive systems to help learners understand what it “feels like” to play music, and to develop a “feel” for musical structures. Introduced in the previous chapter, my research platform combines projection mapping with a player piano. The two main projects on my piano platform, MirrorFugue and Andante, respectively explore each of these dimensions of musical “thinking”.

This chapter details MirrorFugue, which evokes the illusion of a virtual “reflection” that appears to play the physical piano. MirrorFugue invites anyone to sit at the empty piano bench, to “resonate” with the performance, and to play along. Through these interactions, MirrorFugue helps musicians of all levels (including absolute novices and young children) understand what the virtual pianist plays in terms of their own bodies.

The following sections summarize my work around MirrorFugue over the course of my time at the MIT Media Lab. I begin by presenting related work on how to evoke human presence and then delve into the story of how MirrorFugue came to be, from the initial idea through a series of prototypes. I then detail MirrorFugue’s latest implementation and present selected observations from demonstrations and exhibitions of MirrorFugue, as well as the full space of potential scenarios related to
learning and artistic expression. Finally, I describe an informal user study on the project’s ability to evoke a sense of presence.

**Re-Presenting the Body**

Within telepresence research and digital arts, there have been two main approaches to convey the presence of someone who is not physically present. One approach uses video, on a screen or projected into the physical world. The other uses the movement of physical objects to suggest “possession” by an invisible spirit. Here, I review examples from both approaches as well as their combination.

**Figurative Representations**

Using video to depict human presence has been much explored in the field of Computer Supported Collaborative Work (CSCW). Interfaces for remote collaboration often feature a video stream of a remote user on a screen, and extensive research has revealed ways in which choice of framing, size and placement of the video effect empathy and the feeling of co-presence. Other studies have shown that including the upper-body in the video dramatically improves collaborators' empathy over a head-only display and that approximately life-sized video is more effective at simulating presence than smaller sizes.

Several researchers have explored how the arrangement of video screens can simulate real-life spatial relationships between distant users. Ishii’s ClearBoard employs the metaphor of speaking through a glass pane, and Buxton’s Hydra situated video surrogates of remote users around a table to simulate real-life meetings.

Artists have projected video of human figures directly into the physical world to create powerful statements engaging the viewer in contemplative...
experiences. Naimark's displacements and Wodizcko's projections on public architecture both play on the concept of human presence.

**Gesture as Movement**

Another approach abandons the image in favor of animating objects in the world to suggest the presence of an invisible being. Objects may move as if under the direct touch of an invisible person, or they may appear to be infused with an invisible spirit. While the movement of animated physical objects carry significant meaning, it is difficult to imagine the identity of the ghostly presence from seeing the objects alone without explanation. A classic example of disembodied presence is the player piano. Its modern incarnation, the Yamaha Disklavier, captures and re-renders a pianist's touch and tone in near perfect detail.

**Dancing Pixels and Dynamic Form**

Player pianos have been used in conjunction with video for both practical and artistic applications. The Sync-a-Vision by PianoDisc features a built-in screen at the music stand which plays video of performances synchronized with the moving keys. Both remote lessons and concerts have employed the Disklavier to transmit playing over distance along with a video feed shown on a separate screen.

Recent work with pin-based shape displays enable a more general platform for the combination of digital pixels and physical form. A particularly compelling example is Leithinger and Follmer's vision of Physical Telepresence, where actuated pins combined with projection and a vertical screen renders the physical presence of a remote collaborator.

**From Reverie to Reality**

The final version of MirrorFugue combines projection mapping with the moving keys of a digital player piano, but it took me several attempts to arrive at this implementation.

The image of what was to become MirrorFugue first came to me when I was practicing the piano one night in fall of 2008. My gaze fell to the reflection of my hands on the vertical surface in front of the keys, and I imagined in the place of my own reflected hands the hands of another, as if the lacquered surfaces of the piano were a portal to another world, or another time.
A year later, I started as a master’s student at the MIT Media Lab and set out to make that image a reality. Not having any idea how to emulate a virtual reflection, I acquired two digital keyboards and a webcam and began experimenting with different projection techniques of my hands at the keyboard. Initial research compared three configurations. In shadow mode a stream of my hands from one keyboard was projected directly onto the keys of the other. For the two other configurations, I constructed a vertical back-projection surface in front of the keys. In reflection mode, the streamed keyboard was projected on this surface, aligned and flipped vertically to emulate the reflection. Finally, organ mode maintained the same orientation of the streamed keyboard on the projection surface, mimicking the tiered keyboards of an organ.

Frankly speaking, none of these configurations approached the poetic image I first had in mind. They did, however, raise interesting research questions. For instance, how to convey the gesture of a music performance and how to share space across distance. These questions harkened back to classic research in CSCW, such as ClearBoard whose spatial configuration enabled collaborators to communicate not only through speech and drawing but also through gesture and gaze.

My master’s research on MirrorFugue extended ideas from ClearBoard and other CSCW projects into the realm of musical collaborations on the piano. The main application was lessons over distance, and my first set of studies showed how video of a remote teacher’s hands aligned and to-scale at the “interaction locus” enabled students to learn through imitating the teacher’s hands. Of the three configurations, organ mode was the most preferred for the learning applications. It allowed students to learn music not only as sequences of correct notes but as a sequence of embodied actions, including fingering, hand position changes, and expressive gestures.

To complete my masters research, I built a second version of MirrorFugue, which added video of a pianist’s upper body to emulate the reflection on the music stand of pianos. Whereas the earlier prototypes focused almost exclusively on function, this time I took great care on the
aesthetic details. I built a pair of wooden cases to hold the projection system, which I meticulously sanded and painted to mimicked a pair of upright pianos. For content, I recorded a selection of pieces that I was working on at the time (some Bach and Beethoven). It was after watching my recorded performances on this version of MirrorFugue that Yamaha agreed to loan me a Disklavier piano for my PhD research.

To build the next version of MirrorFugue, I mounted a projector and started by porting the setup of the previous version, constructing projection surfaces from painted plywood for the fallboard and music stand. I transformed my office into a custom recording studio, complete with lights, and recorded the slow movement of the Ravel *Sonatine* as a demo. Functionally speaking, the only real difference between this and the previous version of MirrorFugue was that this version featured sound from an acoustic piano, with the keys moving in time. Yet, somehow it felt different. While this particular configuration of information streams, as images, sound, and haptics, was no doubt useful for for “applications of learning and enjoyment”, it felt like there was something more though I couldn’t quite articulate what it was.

Following this intuition, I returned to experiment with different projection configurations, this time attuning myself to the experiential dimension. Early in my master’s research, I had discarded the idea of projecting direction onto the keyboard because it felt too “chaotic”. It still felt chaotic when I moved my keyboard projection from the modified fallboard to the keys. But when I replaced the fallboard with the original reflective one, everything changed. In the fallboard reflection, I saw my virtual hands playing the physical keys, and on the music stand was projected the “reflection” of my face and body. Sitting in front of my virtual self playing the physical piano, All at once, I felt a visceral sense of physical presence of someone who was not there. I felt as if I were my past self. I felt as if I were a ghost who has stepped into a past dream. And there it was, MirrorFugue, no longer a dream but a real, tangible experience.

**Implementation**

**System Configuration**

MirrorFugue’s hardware setup consists of a Yamaha Disklavier with a short throw projector mounted above the piano bench 7.5 ft from the ground, which projects on both the keyboard and vertical surfaces of the
piano without occlusion from a person seated normally on the bench. To create the illusion of a virtual reflection playing the piano, two 720p videos of a pianist’s hands and upper body are respectively projected onto the piano’s keyboard and music stand. Video of the hands is calibrated to align with the physical keyboard and is naturally reflected on the fallboard (the vertical surface in front of the keys). For the upper body display, a 39”x11” piece of 1/4” plywood treated with projection paint is placed on the music stand.

**Software**

A custom Java program controls MirrorFugue for demonstrations and exhibits. For demonstrations, selected performances may be triggered from the keys of a wireless numpad. For exhibits, the program supports looping playlists that are either predefined or randomized. I also prototyped a new version of MirrorFugue in JavaScript that allows playback at various speeds, fast-forward, and rewind. The new program runs from the Chrome web browser, thanks to an extension of the Web MIDI specification published in June 2015.

**Content**

Data for a MirrorFugue performance comprises one MIDI sequence and two videos. MIDI output from the Disklavier is fed to a computer through a MIDI-to-USB cable and is recorded along with an audio stream from the computer’s own microphone. Two 1080p digital camcorders are used, for the pianist’s hands and upper body, respectively. The one for the hands is mounted 7.5 feet above the keyboard, and the one for the face and body is mounted on a tripod at the end of the piano. To record the pianist’s
Reflecting Music Through Movement

face, the music stand of the piano must be folded down or removed. A black fabric is hung behind the pianist for a clean background. Two lights are required for the most basic setup: one from above to evenly illuminate the keyboard, another from the side to create a distinction between the pianist’s body and black background. This basic setup takes advantage of ambient overhead light from the room.

The videos are manually synchronized and edited using Adobe AfterEffects. The final videos for the face and hands are exported after adjustments of framing, alignment, and color correction are made using reference templates. The audio track of one of the videos is then used to synchronize and edit the MIDI track using Reaper.

Performances from 15 pianists were recorded for MirrorFugue. This roster includes 8 professionals, among them three Steinway Artists and several well-known names such as Vijay Iyer, Ryuichi Sakamoto, Jon Cleary and Allen Toussaint. Other recorded players include two children, one piano teacher, and two professors from the MIT Media Lab (Marvin Minsky and Joe Paradiso), as well as myself playing my own classical repertoire. These recordings showcase a range of musical styles and levels of expertise.

![Figure 5.10: New Orleans exhibit: (left) Allen Toussaint, (right) Jon Cleary](image)

**Selected Observations**

Since summer of 2012, the latest version of MirrorFugue has been demonstrated to hundreds of visitors to the MIT Media Lab. Visitors included the full spectrum of musical experience, from complete non-musicians to Herbie Hancock, as well as groups of children and youths.

Additionally, a version of MirrorFugue featuring four pianists from New Orleans (Allen Toussaint, Jon Cleary, Ron Markham, and Nick Sanders) was shown for the Prospect 3 Art Festival at a gallery in the heart of the French Quarter. The exhibit was free and open to the public for ten days.
From both the demonstrations at lab and the New Orleans exhibit, I gathered a set of “natural” user interactions on MirrorFugue, where no specific instructions were given to the visitors except the suggestion that they sit down at the piano bench. These observations illustrate the potential of MirrorFugue as a musical “microworld” to help users of varying levels understand important aspects of playing music through self-initiated explorations.

**Complete Novice: Feeling It**

The most striking interaction on MirrorFugue occurred on the last day of the New Orleans exhibit with a woman who was the self-professed “greatest fan” of pianist Jon Cleary. She informed me that it was her second time visiting MirrorFugue that week, and I invited her to sit down at the piano, when she began to “play along” with her idol. The woman had no musical experience and did not try at all to copy the exact notes. Instead, she mirrored Jon Cleary’s overall movements, appearing to pretended that she was him.

In New Orleans music—as in jazz, which grew from the same tradition—the “rhythmic feel” is the most important part of playing. It is not enough to play the notes or to improvise with the correct idiomatic patterns. “Feel” is a often an elusive quality. It must be understood through the body, and it is impossible to master by brute force. Trying too hard causes playing to become “stiff” and “unnatural”.

When the woman played along with Jon Cleary on MirrorFugue, her body was perfectly relaxed. She appeared to channel the music so naturally that many people, including expert musicians, believed her to be

![Image](image-url)
a “seasoned pro” when I showed them the video with the sound turned off.

This scenario demonstrates a very different approach for beginners to learn to play music from what is typically advocated by pop-ed culture. Instead of first learning the notes and then adding expression, a more effective and enjoyable way may be to first “feel” the music in the body and then gradually add more and more notes, without ever losing the “feel”.

**Children: Improvising to Rhythm**

The principle of “get the feel, then add notes” has also been observed among children. One of the children recorded on MirrorFugue was my advisor’s 8-year-old daughter, Alisa Ishii, who at that time had just begun to play the piano. One day, Alisa was at the lab with her younger sister Lena, age 6, and asked to see her MirrorFugue recording. Both sisters sat down at the piano to listen to virtual Alisa’s performance, and real-life Alisa started to play along with herself from two months in the past.

Alisa’s behavior of retracing her own steps and reflecting on her own progress was interesting in itself and will be discussed later, but the example of “get the feel” was in Lena’s interaction. At that point, Lena had not yet begun to play the piano and did not know her way around the keyboard, but she was inspired by the Alisa duet and wanted to join. After watching the two Alisas, Lena picked one key and played it on a higher octave at a precise rhythm to accentuate Alisa’s performance.

Lena’s playing shows that even for very simple pieces—that consist of only one note!—it is still possible to first “get the feel” by synchronizing
to the underlying pulse. Feeling pulse is a crucial founding for playing more advanced material later on.

**Youth: Getting inside the Sound**

A beautiful example of MirrorFugue as a setting for informal teaching and learning occurred when the composer and Media Lab professor Tod Machover visited with one of his teenage daughters. Both sat down at the piano, and I played for them a performance of the *Gnosienne No. 1* by Erik Satie.

Tod’s daughter plays the piano well and immediately began to imitate the right hand-melody. Tod, on the other hand, invented his own small motifs. More importantly, he paid close attention to the evocative atmosphere of the piece, and whatever he added accentuated the mysterious mood of the piece. In time, Tod’s daughter began to follow Tod’s example to play more “inside the sound world”. Throughout this interaction, no words were exchanged between the two. All communication happened through playing, glances, and other body-language.

Perhaps this interaction is less of a demonstration of MirrorFugue than Tod’s own capacity as a musical mentor. However, it does illustrate the importance of “getting inside the sound”. The fact that no explanation was needed may point to MirrorFugue’s contribution as learning environment—as if Tod’s playing drew attention to features of the virtual performance, and the multimodality of MirrorFugue—audio, image,
haptics—gave more “hooks” for Tod’s daughter to get “inside” the music.

Experts: Technical Details

While novices and amateur players used MirrorFugue to “get the feel” by synchronizing to rhythm and to step inside the sound world, more advanced pianists emphasized the value of MirrorFugue to help convey finer details of physical technique.

A classical concert pianist was fascinated by watching how Ryuichi Sakamoto used “non-standard fingering” to play his own piece and explained that MirrorFugue would be “super useful” for her to more quickly learn the complex fingerings of new pieces.

Fingering is also important for improvised music. Jon Cleary visited the MirrorFugue exhibit and examined Allen Toussaint’s playing for insights on his characteristic riffs. A piano teacher and amateur jazz pianist was able to learn about jazz pianist Vijay Iyer’s voicing by watching his hands on MirrorFugue. Another advanced pianist put his feet on the moving pedals to learn the pedaling technique of different masters.

Based on these observations, it may be tempting for proponents of the pop-ed approach to argue that the technical aspects of playing are the most important because that is what several of the professional pianists who have seen MirrorFugue focused on. However, I prefer a different view in light of my other observations on MirrorFugue. First, the expert’s approach to fingering is different from pop-ed views. Fingering is not just a sequence of numbers but understood more in terms of movements of the hands, which includes when to shift positions, anticipation of large leaps, and the ability to control the weight of the hands to produce various colorings of tone.

More importantly, I believe that the professionals paid more attention to technique because they already understand “feel”. They do not emphasize...
Reflecting Music Through Movement

it because they no longer to consciously think about it, much like native speakers of a language do not actively think about the rhythm of their speech and their intonation. To “speak” music as fluently as a native language, understanding the “feel” and getting inside the sound is a prerequisite to learning more complex physical technique. Without the ability to feel music, technical complexities risk sounding mechanical.

 Potential Interactions

The observed interactions on MirrorFugue have all taken place with previously recorded performances. Beyond these observations, MirrorFugue could also become the basis for a larger space of interactions related to music learning and creation. Full implementation of these interactions would have required tackling technical challenges beyond the scope of this thesis, such as low-latency remote communication. Instead, these scenarios were explored in a series of “vision” videos described below.

 Remote Communication

The same interactions on MirrorFugue with recorded performances could occur when the virtual pianist is streamed over distance. For example, a concert of a famous pianist could be broadcasted to MirrorFugue pianos from all over the world, right into people’s living rooms. Audiences could then sit at the piano, experience the concert from the performer’s point of view, and even play along.

MirrorFugue with two-way remote communication could be used for lessons at a distance, where students and teachers could better observe and demonstrate physical aspects of playing such as posture and fingering. Remote rehearsals or duets could also take place on MirrorFugue, where physical cues could help with synchronization.

Figure 5.14: Remote performance scenario: pianist plays a concert, which is streamed over distance, enabling a listener to play along

Collaborative Composition/Performance
Reflecting Music Through Movement

With the ability to record new performances in real-time, MirrorFugue could become a “sketchbook” for composers, performers, and improvisers. A composer could record ideas as snippets of performances on MirrorFugue, which enables them to capture not only the notes and the sound of their music but also how it is played. The composer could also record demonstrations to go along with the pieces they write to help their performers learn to play the piece. Multiple people could even collaboratively work on a piece together, over distance and asynchronously, by exchanging “sketches” on MirrorFugue.

Conversation with my Reflections

Much of music consists of multiple layers or lines. Canons and fugues typical of baroque music are classic examples of multi-layered pieces for the keyboard. Music of other styles also feature multiple layers with different roles such as melody and bass line. In an ensemble, the timbre of different instruments help clarify the different lines, and in live performance lines are personified by different musicians. In solo piano music, every line is played on the same instrument by the same performer, and it can be difficult to fully appreciate the intricacies of the sound.

MirrorFugue could be used to clarify the layers of a piano composition by personifying each voice with a virtual pianist. The same pianist could be duplicated multiple times, or a different person could represent each voice. Each virtual pianist could project a different character, adding depth to the performance especially when voices exhibit different characters (e.g. a stately bass line with a soaring melody).

Expanding a performance into multiple layers can also serve as a useful practice tool, allowing the student to take over and play along one layer at a time. Practicing this way enables the student to experience the whole piece while working on it portions at a time. This may allow the student to gain a deeper understanding of how the piece fits together.

Figure 5.16: A performance with many reflections, each playing one musical line


Duet with the Past

Recordings on MirrorFugue capture moments from the past. Just as Alisa Ishii played a duet with herself from two months earlier, we could imagine interacting with our reflections from farther in the past. On one hand, such archives become a useful record to reflect on our own progress. We retain a better memory of how we used to play (and use to “think”) and thus become better mentors to help those behind us.

On the other hand, the ability to “go back in time” also creates new, artistic experiences. For instance, what it might feel like to see our reflection from when we were much younger, or to play a duet with our grandparents when they were our age.

Figure 5.17: Scenario across time: a girl learns piano as a child, 20 years later, she returns to share a moment with her past self

User Study

The value of MirrorFugue, both for learning and artistic expression, is largely based on the hypothesis that users at the bench could perceive the interface as an immersive whole and can viscerally feel the presence of the virtual pianist. I conducted a preliminary evaluation to investigate this hypothesis by gathering user feedback.

Through the study, I also wished to gain insight into user perception of what makes the interface viscerally engaging. Specifically, I wanted to understand the contribution of each visual channel (moving keys, projections of hands and body) and how varying their arrangement affects the experience. I devised five interface configurations (see Table 5.1) to present to study participants based on three parameters:

- Music generated from player piano’s moving keys or as sound only
- Projection of pianist or no projection
- Projection of keys in original or alternate configuration
Method

A convenience sample of 15 subjects (8 male, 7 female; Aged 23-49(28)) with varying piano experience watched recordings of five different performances by the same pianist on the five interfaces. Each recording was approximately one minute long, and the order of interfaces and recordings was randomized. After watching all recordings, subjects ranked the interfaces from 1-5 with 1 being most preferred and completed a questionnaire on their preferences. Differences in ratings between interface configurations were evaluated using Friedman’s matched group analysis of variance test with the Nemenyi multiple comparison test.

Results

The average ranking of each interface configuration is summarized in Table 5.1. The classic MirrorFugue (I) had the most preferred average ranking and was significantly more preferred than the configurations that lacked either the projection or moving keys (III, IV, V). The alternate interface (II) had the second most preferred average ranking and was significantly more preferred than the two configurations without the moving keys. The alternate interface without moving keys (IV) had the worst preference ranking.

<table>
<thead>
<tr>
<th>Configurations</th>
<th>Classic (I)</th>
<th>Alternate (II)</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moving keys</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Hand projection</td>
<td>on keys</td>
<td>vertical</td>
<td>none</td>
<td>vertical</td>
<td>on keys</td>
</tr>
<tr>
<td>Average rankings</td>
<td>1.4 a, b, c</td>
<td>2.3 d, e</td>
<td>3.3 a</td>
<td>4.3 b, d</td>
<td>3.7 c, c</td>
</tr>
<tr>
<td>Std. deviations</td>
<td>0.7</td>
<td>1.2</td>
<td>1.2</td>
<td>0.9</td>
<td>1.0</td>
</tr>
</tbody>
</table>

1Friedman’s test and Nemenyi follow up
Values with similar superscripts are significantly different.

Table 5.1: Summary of MirrorFugue Study

Immersion and Presence

11 of the 15 participants felt that the original interface enhanced the presence of the recorded pianist. They spoke of the interface’s immersive experience, and their language indicated their perception of the projection and moving keys as an integrated whole. Some participants described the player as “virtually in the piano”, and others felt “as if the player were
sitting besides me playing for me”. A few even said that the interface “tricked my brain completely into thinking that I was playing.”

**Alternate Interface Arrangement:**
The alternate configuration was four people’s top preference, and they tended to speak of the interface in terms of its separate information streams. Each had a different reason for preferring the hands projection not on the keyboard. One was an experienced pianist and liked how the vertical projection gave a clear view of the hands, explaining that “watching and learning are blurred for me as a former player.” Another experienced pianist had difficulty reconciling the 2D projection “playing” the 3D keys in the original interface but said that he could imagine preferring it once he “gets used to the dimensionality.” Others felt that the hands on the keys was distracting and preferred seeing them more peripherally.

When asked which interface they would use if they were to learn to play, 12 out of 15 chose some variant of the alternate configuration (II: 9, IV: 3). These participants felt that it was easier to follow the vertically projected hands because “it is not blocked by your hands.” Some described direct projection on keys as “too cluttered” and “distracting” for a tutorial interface.

**Role of Moving Keys and Projection:**
When asked about the effect of the player piano’s moving keys, 12 of the 15 participants mentioned that it made the performance seem more real, live and present. For example one participant said that it “gives real energy to the music and made the performance seem more immediate and in the moment.”

14 of the participants mentioned that the projection makes the performance more human and emotionally engaging. For example, one participant said that seeing the movement and expression of the performer was “very much like a traditional performance but directly in front of you.” Some people reported paying more attention to the “nuance of facial expressions that enhanced the character of the piece”. Others paid more attention to the hand movements to better understand the “flow of rhythm and phrasing.”

Though almost all participants enjoyed seeing the moving keys and the projection, the interfaces with only projection or moving keys had the lowest preference averages. This suggests that the positive effects of the
player piano and projection are enhanced when the two appear in combination. Some users explicitly clarified that they found the moving keys “ghostly and unnerving” on their own but really enjoyed it with projection. Others noted that projection on the keys without the keys moving felt “awkward”.

**Takeaways**

After one minute of watching a piano performance on MirrorFugue classic, more than two-thirds of the study participants described feeling the impression of a present pianist. This result supports the claim of MirrorFugue’s interface’s immersive, engaging qualities. Below are some general takeaways from the study:

1. **Combining modalities creates a richer experience** — The projection and moving keys each served a different role and complemented each other in evoking the pianist.

2. **Configuration matters** — Different arrangements of the same information streams yield vastly different experiences for the spectator. In our interface, the coincidence of projecting hands with moving keys created the impression of causality\(^{103}\), which added to the illusion of a present virtual pianist. Though the alternate configuration featured the same video, its arrangement made it more suitable as a useful reference.

3. **A convincing illusion does not require perfection** — When projecting directly on the piano keyboard, the image is only visible on the white keys. Despite not being able to see fingers on the black keys, people still saw the virtual hands as “playing” the physical piano. Similarly, though the lower torso, arms and wrists of the virtual pianist cannot be seen, people still saw the virtual pianist as being there.

This chapter summarized my work on MirrorFugue. I first reviewed a range of work in both research and the arts for strategies on how to represent human presence, which inspired my design. I then recounted how MirrorFugue came about and presented the implementation of the latest prototype. Through real-life observations and envisioned scenarios, I described how MirrorFugue's role as an immersive environment for learning and for creative music expression. In an evaluation, I showed how most spectators do feel the presence of the virtual pianist, perceiving the multiple information streams as an integrated experience.
MirrorFugue reconsiders the musical performance not just as abstract ideas from the composer rendered accurately in time but also as the expressive gesture of a performer who lives, breathes and embodies the music. Whether the performer is live or conjured technologically, the music becomes a visceral communicative experience between the performer and the audience, an experience that is deeply and fundamentally human.

Figure 5.18: Very first sketch of MirrorFugue, fall 2008
Chapter 6: Andante: Microworld of the Keyboard

To really understand how to make music, it is important to understand its building blocks. These building blocks, such as rhythm, melody, and harmony, are fundamental to the work (and play!) of composers, performer, and improvisers alike, yet they are rarely included in the early stages of music learning. When they are taught, it is usually in the context of music theory classes, where the goal is more about learning the correct names and symbols (e.g. ii, V, I) than about making music. When learning music theory, students are often made to identify and manipulate patterns as exercises, out of context and without any expression. Thus the knowledge of musical patterns remains disconnected with the knowledge of playing and expressing music.

Andante was envisioned as a means to “make concrete” the building blocks of music, through the visual metaphor of miniature figures walking on the piano keyboard. The figures make visible both the structure and expression of musical phrases, and present musical movement in terms of the body and in terms of the physical actions to play the instrument.

This chapter first recounts how Andante came to be. I then detail the project’s implementation and describe prototypes that demonstrate the
space of possible content and interactions. Finally, I describe a case study of using Andante in children’s piano lesson and suggest new features for based on these experiences.

**Origins**

After MirrorFugue had reached maturity, I became interested in how the gesture of music could be conveyed in less literal ways with my player piano platform. Inspired by the artwork of Toshio Iwai and Ryuichi Sakamoto for player pianos\textsuperscript{103}, I started experimenting with interactive motion graphics for musical interactions. One example is the “Canon Machine” from a weekend in the summer of 2012, which launched a virtual ball (canon-ball!) with every key strike that would then fall back onto the key and replay it. If you pressed the middle pedal, whatever you played would continue to bounce on the keys, creating a loop. As simple as the rules were, it was a lot of fun to play and mesmerizing to watch for amateurs and professionals alike. The animations literally drew out the gesture across a phrase as an arc of falling balls, and people especially loved watching the virtual balls play the physical keys.

The Canon Machine got people who never played to sit down and improvise their own pieces. The synchronization of the digital and physical, of image and sound drew their attention into the music, freeing their creative explorations from self-consciousness. Pianist friends similarly enjoyed playing with the Canon Machine. They felt that interactions with the machine got them out of their comfort zone in thinking musically and that the visuals helped them anticipate when listening.

Continuing my explorations, I enrolled in a class on animation at Harvard, where I learned the building blocks of the art, as well as techniques both traditional and digital. One of the fundamentals of
animation is the “walk cycle”\textsuperscript{106}. It turns out that it doesn’t take much to animate a walking character, only about a dozen frames that could be repeated on loop. But these frames is enough to express a lot about a character—their gender, age, personality, mood. In real life, we can know a lot about someone from how they walk. We feel it, without any explicit, conscious thought. Attuning to movement works the same in how we perceive animated characters. They don’t even need to look realistic or particularly human for us to feel what they appear to feel.

![Various walks from Learn How to Draw Animated Cartoons by Preston Blair](image)

Around the same time in my piano studies, I was on a personal mission to improve my sense of rhythm. The secret to rhythm is that you have to feel it. It is not enough to understand it intellectually, to count the beats from reading the score. In jazz, there is a concept called the “groove”, which unifies a piece and propels it forward\textsuperscript{107}. Even in classical music, pedagogue Abby Whiteside writes of a “basic rhythm” found in every piece that coordinates the various muscle groups to play the piece\textsuperscript{108}. It occurred to me that the walk cycle might be a way to convey rhythm and phrase on the piano. I was not the first to come up with the idea. There’s that episode of Tom and Jerry where Tom the concert pianist performs Liszt’s famous Hungarian Rhapsody and Jerry runs around on the piano\textsuperscript{109}. And in the movie *Big*, there’s that famous scene where they play the floor piano at the toy store by leaping around on the keys\textsuperscript{110}. Andante brings these ideas to a real, live piano. I envisioned a system where animated figures walk and dance along the keyboard, physically striking a key with each step. With Andante, I wished to create a learning platform that not only conveys important dimensions of music often ignored in mainstream pedagogy but does it in a way that is fun and engaging. Andante’s figures give visual form to musical phrases, encouraging thinking beyond individual key strikes by connecting them in lines of motion. As walks inherently communicate affect, the figures
convey at once both the notes and the expression of a musical line through where and how they step.

**Implementation**

Andante is built on the same piano and projection platform as MirrorFugue. To display on the fallboard, the reflective keyboard cover is replaced with a projection surface made from plywood treated with projection paint, the same material as the MirrorFugue upper body display.

Two versions of the software were built to drive playback and animation. In the original implementation, all animations were drawn by hand using a light-board for precise control of character’s movements. Frame sequences were organized based on type of step (e.g. whole step between white keys). A Java program controls playback by reading MIDI sequences recorded from a human player and selects the appropriate frame sequences to display for each note played on the Disklavier. I also prototyped a new version of Andante in JavaScript that procedurally generates character animations with life-like movement from human input of musical phrases.

Andante could serve as a microworld in which learners familiarize themselves with elements of music through self-guided, open-ended interactions. I now describe the types of musical content that could be learned from Andante and the interactions it allows.

**Content**

**Music as Movement**

A single walking figure can visualize the motion of one musical line. Rhythm can be understood in terms of how and how fast the figure
walks, runs, or skips. Motifs and phrases can also be clearly visible from figures’ movement. For melodies with large leaps, figures visually reinforce the “effort” needed to make the jump, which can help students with expressive phasing. More frequent leaps in a melody are often better understood when heard as two distinct voices, which can be depicted as two distinct figures. For voices with breaks (rests) during or between phrases, figures can fade in and out or decrease in brightness to maintain presence based on the structure of the line.

To demonstrate the basic idea of Andante, I built a set of examples where characters traverse the keyboard playing scales to explore the visualization of linear motion. An essential component of the musical vocabulary, scales are used to construct melodies and are often included as daily practice for students of all levels, though often treated as a mindless exercises devoid of musicality.

I prototyped the C major, pentatonic and chromatic scales, ascending and descending with adjustable tempo. To show variations in scalar exercises, animations were created with multiple figures playing the C major scale in contrary motion and harmonized in thirds.

I also experimented with changing the characters’ appearance and gait, animating a fat man, a tip-toe sneak, and an ostrich for the basic C major scale. Initial observations from viewers seem to suggest differences in how figures look and move noticeably altered the way a scale is perceived, even with MIDI playback lacking in dynamics of the initial prototype. Andante can encourage students to play scales with more shape and expression, imitating the movement of various characters. Particular characters can also be assigned to scales as a way of helping students remember by association.

Musical Structure

Multiple figures can break down more complex compositions into separate voices. I prototyped a visualization for Canon BWV 1073 by J.S.
Bach, a perpetual canon that can be infinitely looped as its ending wraps back to the opening. This visualization assigns a differently colored figure to each of the canon's four voices and allows the user to turn each character on and off.

A student can learn a piece with more attention on compositional structure, inspecting the shape of each voice, interaction between pairs of voices, and how all voices fit together. Isolating subsets of voices could reveal compositional techniques such as imitation, echo, and inversion through the characters' motion. The student can learn the piece one voice at a time from each figure and practice while accompanied by the rest of the characters for a global perspective early in the learning process.

The movement of multiple figures is also one way to understand harmony. While we often think of harmonies vertically, a crucial aspect of harmony is the horizontal movement of component parts. The Canon example highlights horizontal motion as each figure is assigned a different color. Alternatively, figures in the same animation could be differently colored according to underlying chord changes to show harmonic progression (e.g. red for tonic, yellow for dominant).

Another way to represent harmonies in the Andante universe is through cartoon buildings, which emphasize the vertical stacking of notes. Each harmony house consists of three "columns", which each play a note of the chord, and a top with an identifying shape. House for different harmonies can also be associated with different colors, adding another mnemonic device. I prototyped a simple piece consisting of three houses representing the root, the dominant, and the subdominant chords of the left hand.

Abstract Symbols Made Concrete

Andante can also serve as an intermediary between between the physical world of the piano keyboard and the symbolic world of standard notation. As preparation, learners could be told that the figure's head corresponds to pitch and the feet to rhythm. Different stems can then be added onto the head to show rhythmic notation. Staff lines can also be overlaid above the figure, and the position of the head can be shifted up and down based on the note played by the feet. As symbolic notation is introduced, the presence of the figure can help remind learners of the continuity and the expressive qualities of musical lines.
**Interactions**

Most basically, Andante figures can support learning by imitation. Learners may rest their hands on the keyboard to shadow the movement of the keys and gradually play along, connecting audio with the haptic and the visual. They may also practice the melody an octave higher or lower along with the figure. This may help learners identify and correct mistakes where their playing does not blend seamlessly with the figure’s playing.

By adding user input, feedback could be provided on expressive playing. For instance, a user could play a phrase, which could then be repeated by a procedurally generated virtual character. Nuances in timing, dynamics and articulation could be reflected in its movement to visually reinforce expressive dimensions of the phrase. We could also render the same phrase with different characters, adjusting emotional parameters such as “sadness” or “excitement” to experiment with different ways of phrasing. This way, a relative beginner could explore the expressive dimensions of a melody and then learn to play by imitating the character.

Andante figures can also establish a steady rhythm, especially useful in bass accompaniments. In fact, the “walking bass line” is a common pattern in both Baroque and Jazz styles, which evokes the rhythm of alternating steps. To demonstrate this interaction, I prototyped a boogie woogie vamp, an iconic example of walking bass in blues piano.

Though this prototype uses a fairly neutral figure, the distinctive character of various rhythmic styles, such as samba and swing, can be accentuated with characters with corresponding appearance and movement. Playing along with these characters can teach the feeling of rhythms that are difficult to convey only as abstract concepts. A student can also play against a character, focusing on melodic improvisation while offloading the accompaniment. The animation acts as the rhythm section of an ensemble, helping the player get into a groove.

Andante could also become the basis of a composition tool, which builds pieces through layers represented by figures. A composer could record one layer at a time, keeping track of layers and their interactions through the visualization. Representing voices as characters could also help the composer create a narrative throughout the piece.
User Study

In fall of 2014, a local children’s piano teacher, Pablo Puentes, contacted me and offered to pilot Andante with his students. Over the course of the next year, Pablo and I met regularly to discuss how to adapt Andante for children’s lessons and practice. Unlike many music teachers, who often treat learning music as learning notes, Pablo tries to show his students how to play music and how to play with music beyond the notes. He has extensively explored alternative methods of music pedagogy and holds a certification in Dalcroze Eurhythmics, a method of instilling early musicality in children that connects musical structure and expression with movements of the body. Here, I describe the experiments conducted with Pablo and his students. Chapter 7 will discuss in more detail the theories of Dalcroze as well as their connection with more general theories of learning and psychology.

Two studies were conducted to observe Andante in the context of a lesson and a practice session with eight children between the ages of 7 and 13. Each session was captured with two cameras, one for the student’s face, the other for the hands and the projections on the piano. Videos were reviewed and a transcript was made of the main interactions.

We created a graphical user interface running on a computer adjacent to the piano where students may select portions of a piece to play as an auditory reference. The interface enables playback of hands separately and together. For students who already know how to read, projected highlights indicate which portion of the score is selected for playback.

Due to the small sample size of children and short duration of the sessions, our observations remain anecdotal. Still, our findings suggest promising directions for future work, both for Andante and for the design and evaluation of other music learning systems for children. I now describe the setup of each study followed by notable results.

Study I: Lesson

This study compared two lessons, one where Andante was used as a reference from the start, and another that began with only the score. Two students of comparable piano experience were selected (see table 1). As material, we chose an excerpt from the Bach canon from the earlier prototype, showing only two of its voices across four measures. Playing a canon is a challenge because it requires understanding the interplay between the two overlapping voices.
Both sessions began with about 5 minutes of pre-lesson exercises, where the concept of a canon was introduced via clapping exercises. The piece was then presented in a lesson planned to be about 15 minutes long.

<table>
<thead>
<tr>
<th>ID</th>
<th>Age</th>
<th>Gender</th>
<th>Total years playing</th>
<th>Years with collaborator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>F</td>
<td>5</td>
<td>1.5</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>F</td>
<td>7</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 6.1: Study I subjects

Andante was introduced midway through student-2’s lesson to help alleviate visible stress. As a result, student-2’s lesson consisted of 12 minutes with only the score and 14 additional minutes using Andante. The sheet music was kept on the music stand during both lessons, and the teacher guided each student through the piece in one to two-measure increments. An additional experimenter in the room triggered playback on Andante based on voice commands of the instructor. Sections of the piece selected for playback were highlighted on the score using the two colors of the Andante figures.

**Study II: Practice**

A second study was conducted to compare how students use different reference technologies during practice. Alongside Andante, we implemented a version of piano roll notation commonly used by existing learning systems, where “falling blocks” were projected onto the fallboard to indicate note strikes and releases. To be comparable to Andante, our version of Falling Blocks also appeared to play the physical piano, and voices were colored corresponding to Andante figures.

<table>
<thead>
<tr>
<th>ID</th>
<th>Age</th>
<th>Gender</th>
<th>Total years playing</th>
<th>Years with collaborator</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>8</td>
<td>M</td>
<td>2.5</td>
<td>1.5</td>
</tr>
<tr>
<td>A2</td>
<td>7</td>
<td>M</td>
<td>2.5</td>
<td>1.5</td>
</tr>
<tr>
<td>B1</td>
<td>9</td>
<td>F</td>
<td>3.5</td>
<td>1</td>
</tr>
<tr>
<td>B2</td>
<td>9</td>
<td>F</td>
<td>3.5</td>
<td>1</td>
</tr>
<tr>
<td>C1</td>
<td>11</td>
<td>F</td>
<td>3.5</td>
<td>2</td>
</tr>
<tr>
<td>C2</td>
<td>11</td>
<td>M</td>
<td>2.5</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 6.2: Study II subjects

Six students were selected for this study grouped into three pairs of two (see Table 6.2). Pieces were chosen based on the level of the students. The youngest group learned a modified version of Allegro in G from the
### Table 6.1: Study II Summary

<table>
<thead>
<tr>
<th>ID</th>
<th>System Used</th>
<th>Part Practiced</th>
<th>Total system starts</th>
<th>Total repeats</th>
<th>Average repeats per start</th>
<th>Corrected mistakes by ear?</th>
<th>Practiced from memory?</th>
<th>Times RH practiced</th>
<th>Times LH practiced</th>
<th>Times HT practiced</th>
<th>Performance Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Andantino</td>
<td>Allegro_A</td>
<td>7</td>
<td>25</td>
<td>3.57</td>
<td>Yes</td>
<td>Yes</td>
<td>8</td>
<td>10</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Blocks</td>
<td>Allegro_B</td>
<td>7</td>
<td>39</td>
<td>5.57</td>
<td></td>
<td></td>
<td>23</td>
<td>7</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>A2</td>
<td>Andantino</td>
<td>Allegro_A</td>
<td>7</td>
<td>48</td>
<td>6.66</td>
<td>Yes</td>
<td>Yes</td>
<td>8</td>
<td>8</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>B1</td>
<td>Andantino</td>
<td>Canon_A</td>
<td>3</td>
<td>11</td>
<td>3.67</td>
<td>Not always, preferred to read score</td>
<td></td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Blocks</td>
<td>Canon_B</td>
<td>3</td>
<td>15</td>
<td>5</td>
<td>Not always</td>
<td></td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Blocks</td>
<td>Canon_All</td>
<td>3</td>
<td>19</td>
<td>6.33</td>
<td>Sometimes used score for hard spots</td>
<td></td>
<td>8</td>
<td>5</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>B2</td>
<td>Blocks</td>
<td>Canon_A</td>
<td>3</td>
<td>14</td>
<td>4.67</td>
<td></td>
<td></td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Blocks</td>
<td>Canon_B</td>
<td>3</td>
<td>9</td>
<td>3</td>
<td></td>
<td></td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Andantino</td>
<td>Canon_All</td>
<td>3</td>
<td>16</td>
<td>5.33</td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>C1</td>
<td>Andantino</td>
<td>Canon_A</td>
<td>4</td>
<td></td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Blocks</td>
<td>Canon_A</td>
<td>4</td>
<td></td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Blocks</td>
<td>Canon_All</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>C2</td>
<td>Blocks</td>
<td>Canon_A</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>Yes, except one hard spot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Blocks</td>
<td>Canon_B</td>
<td>3</td>
<td>7</td>
<td>2.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>

Suzuki curriculum, where the left hand pattern was simplified into block chords to make the harmony more explicit. The other two groups learned the same Bach canon from the first study.

A graphical user interface on a laptop computer next to the piano allowed students to select portions of each piece, two measures or longer, to be played back using either Andante or Falling Blocks. Students could select whether to play back only the left hand, right hand, or both. As in the first study, the score was placed on a music stand, which was highlighted to display the selected measures. Pressing the space bar starts and stops looped playback of the selection.

Sessions began with a 10 minute overview, where Pablo introduced the piece and taught the student how to use the computer interface. Students were then given 25–30 minutes to practice with Andante and Blocks, respectively. Students decided what to play using the system and what to practice. They were also free to choose when to listen or read from the score. Pablo stayed in the room and took notes on the students' behavior. He also interjected brief instructions if a student appeared stuck.

Following each practice session, Pablo conducted a small survey with each student about the technologies.
Results and Discussion

Table 6.3 details how students used each system to practice. I now discuss three recurring themes observed across the two studies.

Promoting Listening and Imitating

The multi-modal, enactive nature of both Andante and blocks seems to promote deeper listening. 6 of the 8 total students (all except 2A and 2B) expressed more awareness of the structure and expression of sound when using our interactive technologies. This awareness was reflected in their body language as well as in their playing. In the second study, 5 of the 6 students (all except 2A) rarely looked at the score, relying instead on the memory of what they have heard. When listening, 1A, 3A, as well as 2B liked to put their hands on the keyboard to feel the movement of the keys. When playing, student-1 followed closely the tempo and articulation of the Andante figures even though none were indicated in the score. Interestingly, 1A’s playing also became more detached to follow the unintentional articulation of the falling blocks. In the survey post Study 2, all students indicated that they would like to be taught with both Andante and Blocks.

These observations suggest that the combination of audio, visuals, and haptics may help engage the attention of children when learning music. We also observe that imitation may be a good strategy to convey several dimensions of music at once (e.g. notes, timing, articulation, structure). Said otherwise, redundancy (achieved either through multi-modality or imitation) serves as an organizing principle that brings about new insights without conscious effort on the part of the student. Since students imitate all dimensions of what they hear, designers should be careful to avoid haphazard, non-musical choices such as the detached notes of the falling blocks.
The Score and Preoccupation with Correctness

Three students’ experiences support our conjecture that playing by ear rather than reading the score is a more effective and enjoyable way to learn. When the score was first introduced in student-2’s lesson, she immediately appeared nervous (“What if I mess up?). Even though the teacher showed each segment of the piece by playing it on the piano, her attention was completely focused on the score. Later, she struggled to play with both hands together and appeared more tense with each successive mistake until the introduction of Andante.

Similar attitudes were observed for students in Study II. For example, student-2A’s attention was totally fixed on the score during the entire practice session. When attempting to put the hands together, she focused so much on reading that she did not notice the mistakes that she made. Even though student-3B generally listened carefully and learned by imitation, there was one point where he read the score to practice one transition that was not covered by Andante and Blocks. While reading the score, 3B’s playing became detached, and he lost his usual richness in tone quality, which returned after one more listen to Andante.

These observations demonstrate how the score may become a “safety line” that students cling to when they encounter difficulties. However, clinging to the score may not result in easier learning. In fact, these anecdotes show how reliance on interpreting symbols inadvertently leads to prioritize surface correctness at the expense of listening and expressive playing.

Andante and Emotional Engagement

In the first study, the introduction of Andante visibly lightened the mood of student-2. She asked curious questions about the figures (“What are their names?”, “Did they eat dinner?”, “Do they need water?”). She then practiced each voice in a duet with a figure playing the other voice. Students 1, 1B, and 3A also displayed visible delight at the figures’ appearance. 1B could not help laughing whenever the figures appeared. He would dance along with their movements and practiced along with the figures on a higher octave.

In Study II’s post-practice survey, all except student-2B indicated that Andante is the easiest for memorization compared to Blocks and the score. Students were also asked to close their eyes when an audio snippet was played for them. All except student-1A imagined Andante figures. These experiences suggest that the whimsy of Andante figures appeal to
children’s imagination and help motivate practice. More longitudinal studies could be beneficial to observe how an evolving narrative around Andante figures may continue to engage children in the long term.

**Design Extensions**

Three main features may help improve Andante’s impact as a didactic tool based on what took place during the studies.

First, we could include indication for finger usage. Because the current system lacked the information, the teacher had to intervene on several occasions to show students how to play a tricky passage. To emulate the teacher’s hands, we could also combine projections of a pianist’s hands on the keyboard with Andante figures.

Second, we may add a speed control for the playback of reference materials. We noticed that the default playback speed for both Andante and Blocks was too fast, and all students except 3A imitated the speed of the demonstrations by default.

Finally, Andante may be improved by including a mechanism reminding students to listen (and not reinforce incorrect playing). The head of the figure could be used to give feedback to student. For example, if students consistently make a mistake, the face could turn to them with a funny expression and proceed to show the correct phrase. Introducing an element of comedy in error correction could help ease students’ frustration and motivate them to try again.

Allowing a tool to be more didactic, however, does not always make for better learning experiences. We all know how annoying artificial help agents can be (smiley faces, unwanted for advice), and clearly, we would not want Andante to turn into a set of miniature explainers, rather than a reflection of our play!

![Figure 6.14: Hand-animated Andante walk sequence](image)
III: Music Learning and Beyond

Chapter 7: Dalcroze, Design, and Mental Development

The next three chapters look beyond my projects to examine learning music beyond the piano and learning beyond music. This chapter addresses both of these topics by way of Dalcroze ideas on musical development.

Émile Jaques-Dalcroze was a Swiss composer and music educator who developed a method to instill early musicality in children through a series of full-body activities and games, which he called Eurhythmics. Dalcroze Eurhythmics was not specifically designed to teach any instrument. Rather, the goal was for children to acquire as “sense” for the building blocks (e.g. rhythm, melody, harmony), form, and expression of music, as preparation for later studies in performance and composition.

Dalcroze Eurhythmics is similar in spirit to a number of other methods developed around the same time by music educators such as Willems, Orff, Kodály, and Suzuki, as part of the Progressive Education movement. This thesis focuses on the ideas of Dalcroze partially because I had the most exposure to it than the other methods, via my collaboration with Dalcroze-certified piano teacher Pablo Puentes. Of all the methods, Dalcroze is also the most compatible with my own projects, given its focus on feeling rhythm in the body. Dalcroze’s views on the genesis of the musical mind also has the most interesting parallels with the underpinnings of Constructivist education philosophy.

This chapter first presents an overview of Dalcroze’s philosophy on musical development. It then presents a framework that adapts Dalcroze’s ideas for learning musical instruments. Finally, I connect Dalcrozeian ideas...
with the ideas of Piaget, Bruner, Papert, and Minsky, drawing parallels between music learning and learning in general.

**Genesis of the Musical Mind**

Dalcroze described 3 progressive stages of musical understanding, which he called the Instinctive, the Conscious, and the Intellectual. His method of Eurhythmics cultivates all three stages, each building on the previous.

**Instinctive Stage**

Students of Eurhythmics begin by developing an Instinctive sense for music, which involves connecting different features of sound with movements of the body. A teacher would play music, and students would engage in various full-body activities based on the music. To understand *tempo*, students might step to the beat, and may walk or run depending on the speed. Different rhythms might correspond to different types of steps, such as skips to convey the feeling of dotted rhythms. To understand melodic shapes, students might go up and down physical steps that correspond to ascending and descending patterns of sound. Other activities might include physical props, such as tossing a ball (or a silk scarf) to understand “falling” motifs in music.

These full-body activities begin to establish a body-based foundation for musical “thinking”, sensitizing students’ ears for the building blocks of music. In Eurhythmics, musical building blocks are introduced very early on but never in the abstract. Rather, they are always presented in terms of human movement, which is inherently expressive.

**Conscious Stage**

After an Instinctive understanding is solidly developed, Dalcroze Eurhythmics then proceeds to the Conscious stage, where students begin to create their own patterns of music. Even though students may not yet know how to play any musical instrument, they could still create and experiment with patterns of sounds with only the movements of their bodies. For instance, anyone can clap, tap their feet, or use vocalizations to express rhythms. To experiment with melodies, people could sing or hum. Students could also “conduct” the playing of another person on an instrument, expressing desired properties of sound through their gestures.
To get started in the Conscious stage, the student might begin to imitate patterns given by the teacher—through clapping or humming. To make the pattern their own, students might be asked to try out different variations. For instance, they could discover what it is like to clap a pattern twice as fast (or twice as slow), to clap the pattern backwards, or to break a big pattern into smaller pieces. Eurhythms classes usually occur in groups, and students could engage in collaborative activities. For example a group might hold a beat while a soloist plays variations.

In the Conscious stage, students gradually gain the ability to imagine sound, to invent their own patterns, and to express what they imagine through the movements of their body. This stage builds on the body-syntonic knowledge of the Instinctive stage, both the musical building blocks and how they are expressed.

**Intellectual Stage**

Only after reaching proficiency in both the Instinctive and the Conscious stages are symbols to standard musical notation introduced in the Intellectual stage. When introduced, all symbols are presented in terms of existing knowledge. For instance, the symbols for different types of notes representing different rhythms are connected to their corresponding full-body exercises from the Instinctive stage. To connect to the Conscious stage, students not only try to enact patterns given to them via symbols but also try to translate into symbols patterns of their own invention. In a group setting, students could also take turns writing out instructions for their peers to perform.

Introducing symbols after the Instinctive and the Conscious stages ensures that students understand that music is ultimately about making expressive statements. Notation is a way to capture and transmit music, much like how the alphabet captures and transmits expressive thoughts. Even though writing doesn’t capture every expressive nuance, we still know that the symbols stand for speech and read sentences with a natural sense of flow and meaning. Likewise, in the case of music, though its writing system is more complex, it is still possible to learn to use it in the service of expressing ideas and feelings. Learning to do so requires not only an understanding of musical building blocks but also their connection to expression through the body.
Design Framework

Dalcroze’s Eurhythmics and his framework of musical development encapsulate many ideas about music learning at the root of my projects. I should note, however, that MirrorFugue and the core ideas behind Andante were not directly inspired by Dalcroze but rather by my own experiences learning to play the piano. My exposure to Dalcroze came from a year-long collaboration with Dalcroze-certified children’s piano teacher Pablo Puentes. By the time I had met Pablo, my work on MirrorFugue was already all but complete. However, the collaboration with Pablo did help me gain new insights into the art of music learning, through many hours of very inspiring discussions. It also resulted in several new features for Andante, such as ways of introducing symbolic notation and harmony.

The fact that Pablo and I resonated so well despite our differing backgrounds demonstrates that effective ways to learn music share many core ideas. As a resource for other designers, I worked with Pablo to distill his version of the Dalcroze method into a five-step iterative cycle, or framework. This framework adapts Dalcrozian philosophy for lessons in how to play a musical instrument, and applies equally for piano and other instruments. For each step of the framework, I describe how a lesson might take place as well as how features of MirrorFugue and/or Andante could play a supporting role. Examples of MirrorFugue and Andante demonstrate how the framework corresponds to actual designs.

1. Listen

The student is first exposed to a piece through the ear rather than the score. During a lesson, the teacher plays the piece for the student on the instrument. An audio recording is also made available for the student to reference during the lesson and practice.

Both MirrorFugue and Andante provide an easy way for students to listen to reference material. Both interfaces could offer different versions of any piece, played by various people or figures with varied expressions. Both could allow students to select and listen to subsections or individual lines within a piece.

2. Internalize

Before going to the piano, the student must first internalize the piece through the body. The melody is introduced several measures at a time.
based on motifs and phrases. It is first separated into the components of pitch and rhythm, which are taught and practiced through various exercises. Exercises to train pitch include singing the tones (without a definite rhythm) and a technique known as body solfège, where pitches are mapped onto parts of the body. Exercises for rhythm include clapping and speaking. In the speaking exercise, the words in a phrase (ex: “mom’s cooking lemon cake for tomorrow”) are mapped to different subdivisions of the beat (figure 2). After the pitch and rhythm are successfully learned in isolation, they are then combined to form the full melody.

![Figure 6.2: Pablo’s speaking exercise for rhythmic subdivisions](image)

MirrorFugue and Andante could both play an important role in this stage by presenting music in a way that relates to the student’s body. MirrorFugue enables the student to resonate with the rhythms, phrases, and emotions channeled by the virtual pianist while Andante enhances the awareness of musical structure.

3. **Extend**

Only when students have internalized the melody (demonstrated by singing and clapping) do they go to find the notes on the instrument. This ensures that natural expression from the body is extended to playing on the piano. Attention should be devoted not to reading notes but to the expression in the sound.

Both MirrorFugue and Andante could also play a key role in this step, by allowing students to learn through imitation. MirrorFugue makes clear the physical technique of playing while Andante could amplify the expression of rhythms and melodies.

4. **Analyze**

If the student is still learning how to read, the score is introduced at this point, associated to what the student already knows how to play. At this stage, harmony is also presented. In the typical learning method, harmony is generally ignored for beginnings, and especially for children. In Pablo’s approach, harmonic grammar is introduced early in the process. Images
and spatial (or figurative) “re-presentations” are privileged over symbolic notation, as in the following “mapping game”: “home” stands for root position, “school” stands for the dominant, and the “forest” is the subdominant. Such body-syntonic mapping technique serve as a precursor to understanding standard notation because the spatio-temporal relations at play can be understood in terms of familiar metaphors—and consolidated in action!

Andante is most useful in this step. It could be helpful to introduce notation, by modifying the head of the figure and superimposing staff lines. The feature of graphical “houses” for different harmonies was actually designed based on Pablo’s mapping game.

5. Improvise

According to both Dalcroze and Suzuki, learning music should mimic the way children learn their mother tongue. This means not only repeating what is written but playing with the material in new combinations. Pablo plays improvisatory games with his students, for example, taking turns with the student to invent new rhythms by mixing patterns from a piece. Student should also learn to incorporate improvisation during practice to avoid mindless repetition.

Both MirrorFugue and Andante could become the setting for improvisatory games to help students play with music. For instance, the virtual pianist or Andante figures could hold a pattern on which students could improvise, or setup a game of “call and response”.

Figure 7.1: Sketch correlating Andante features with stages of a Dalcrozan learning process
Connection with Constructionism

There are many parallels between Dalcroze Eurhythmics and the core tenets of Constructivist and Constructionist theories of education. Here, I discuss three themes, connecting the dots between music learning and learning at large. For all three themes, I also show how examples from music contribute important nuances to established ideas of general cognitive development.

Concrete vs. Abstract

Dalcroze’s stages of musical understanding share striking similarities with Jean Piaget’s stages of children’s cognitive development. Piaget was a Swiss psychologist who pioneered the field of developmental psychology, starting with his detailed descriptions of how his own three children came to make sense of the world over time. Piaget was particularly interested in the genesis of the scientific mind, and his own observations and experiments led to his formulation of Constructivist education philosophy. Constructivism states that knowledge is actively “constructed” by each individual in their mind, not “transmitted” ready-made.

Interestingly, both Dalcroze and Piaget hailed from Geneva. Though Dalcroze was active 30 years earlier than Piaget, it is unclear whether his work influenced Piaget or whether the two ever crossed paths. In any case, both described a progression from the concrete to the abstract, beginning in human’s innate sensory-motor capabilities and ending in the development of abstract, symbolic thought. However, Piaget and Dalcroze differed in how they conceived of the relationship between stages.

A Rationalist at heart, Piaget saw formal, abstract thought as the final, most mature stage of cognitive development. For him, the stages progressed one to the next, toward the increasing “sophistication” of “adult” thinking. Though Piaget’s greatest scientific contributions arguably lay in his descriptions of concrete modes of thought (e.g. sensorimotor and pre-operational), his biases toward abstract thought could be seen in his belief that adults who heavily rely on concrete thinking have not fully reached the end of cognitive development.

In contrast, for Dalcroze each stage of musical understanding arise as a new layer that does not eclipse previous stages. In fact, the scaffolding of the previous stages is integral to the development of new kinds of musical thinking. Learning symbolic notation and musical analysis does not mean
that instinctive understanding is a “less advanced” stage to be left behind. In fact, music-making without connection to the Instinctive and the Conscious becomes an empty exercise.

This aspect of Dalcrozian thought more closely relates to American psychologist Jerome Bruner’s discoveries from his modifications on Piaget’s original experiments\textsuperscript{120}. Bruner showed that earlier cognitive stages don’t simply disappear but instead layer and build “on top” of existing infrastructures.

Dalcroze could also be seen as a precursor to the ideas of Seymour Papert on the importance of connecting to the concrete. Papert’s Constructionism inherits heavily from Piaget’s Constructivism but places more emphasis on constructing in the physical world as a means to construct knowledge in the mind. Papert disagreed with Piaget on the privileged status of abstract, formal modes of thought. Instead, he advocated for learning environments to support a variety of learning styles, including ones more rooted in the concrete\textsuperscript{121}.

According to Papert’s student Uri Wilensky, “abstract” and “concrete” do not describe the global cognitive style of an individual but rather the quality of relationship between an individual and a specific subject\textsuperscript{122}. Wilensky further states that any subject could shift from abstract to concrete if a person develops the right relationship with it. Thus, someone who is fluent in math is not one who has mastered the abstract, rule-based forms of thought but rather one who has built the mental connections to be able to engage with symbols in a concrete way. Dalcroze Eurhythmics demonstrates exactly the same idea for music. Connecting musical symbols to an existing foundation of body-based understanding is one reliable strategy for making the abstract concrete.

**Multiple Representations**

Many theorists associated with Constructivism and Constructionism promote the importance of having many ways of knowing a subject. Papert himself called for learners to build a rich network of connections to existing knowledge. For Wilensky, the difference between a concrete and an abstract way of relating to a subject is dependent on the number of connections one has with the subject. Marvin Minsky once stated that “a thing or idea seems meaningful only when we have several different ways to represent it—different perspectives and different associations”\textsuperscript{123}.
Many educators encounter these ideas and focus on giving students access to as many representations as possible. While the quantity of available representations is certainly important, Dalcrozian music learning tells us to also consider the types of representations. Bruner makes the distinction between three types of knowledge representations: enactive, iconic, and symbolic\textsuperscript{124}. While this framework was originally developed as a reaction to Piaget’s stages, where children think through actions, through images, and through abstract symbols, we could also connect it to Dalcroze’s Instinctive, Conscious, and Intellectual stages of musical understanding. The Instinctive and Intellectual directly parallel Bruner’s Enactive and Symbolic representations. As for the Conscious stage, Dalcroze’s ideas actually expands Bruner’s notion of the Iconic representation (and Piaget’s “thinking in images”) from purely visual images to the audio realm.

While the English word “image” carries strongly visual connotations, the verb “to imagine” is not limited to only imagining pictures. We could also imagine in sound—think of hearing a dear friend’s voice or having a song stuck in your head—and even other senses like touch, smell, and taste. The essence of Dalcroze’s Conscious stage is the development of the aural imagination, and a key insight of Eurhythmics is that engaging the body is one way to build up the imagination.

The fact that a learning environment is accessed through a computer screen does not limit it from engaging enactive representations. A classic example is Seymour Papert and Cynthia Solomon’s suggestion to children programming in Logo to imagine themselves moving as the turtle\textsuperscript{125}. Children project themselves in the place of the turtle, and sometimes even begin to physically orient their bodies or walk around the room. In this way, they are using both enactive representations and their imagination. The importance of connecting to the body is reflected in Papert’s notion of body-syntonicity. While body-syntonicity and “multiple representations” are often seen as two separate concepts from Constructionist learning theory, music reminds us that they are in fact two sides of the same coin.

\textit{Make It Your Own}

Dalcroze Eurhythmics encourages students to invent their own musical patterns. This relates to Piaget’s saying that “to understand is to invent” as well as the core Constructionist tenet that the best way to learn something is to “make it your own”\textsuperscript{126}.
In an effort to encourage children’s creativity, Constructionist educators sometimes adopt an “anything goes” attitude where whatever children make is equally valuable and meaningful as long as it is “their own”. While all open-ended explorations could have value for learning, music shows us that there is a difference between “stringing notes together” and “making an expressive statement”.

Anyone could learn to accurately play the notes of a piece, even a machine. In fact, machines could play more accurately than humans. Similarly, any machine could be programmed to generate endless patterns of notes, just like anyone could make their own music by haphazardly putting together patterns. What then is the difference between composition and performances that move us and those that do not?

Some might explain that the difference is “talent”, a mysterious quality possessed by select individuals. Others might dismiss the question all together by citing the subjectivity of musical taste. Eurhythmics suggests that key to this question lies in the word “move”. Great performers and composers can “move us” with their music because they understand the link between music and motion, and the link between motion and emotion. Composers have learned to express emotion in the sense of movement within the music they create, and performers have learned to conjure emotion through their own movements to make sound. Neither is possible without the foundation of the Instinctive stage and the practice of enacting music.

A striking example of the link between enactment and creative musical expression in a software learning environment could be seen in an anecdote recounted by Jeanne Bamberger. Bamberger had designed MusicLogo, a software environment that enabled children to make rhythms by typing numbers. Different numbers was played back by the computer as different durations of beats. Bamberger had brought MusicLogo to a classroom of 8 and 9-year-old children, and one child, Leon, invented an interesting experiment. Leon purposely programmed simple beats (with only two numbers) on the computer and tapped along with his finger at the same time as the computer played the beats. Instead of just passively hearing the computer’s sound, Leon mobilized his body to get a sense of the different rhythms. After feeling out the rhythms, Leon then proceeded to compose his own piece, with his self-imposed limit of two numbers. By first feeling out the two numbers with his body, I believe that Leon was then able to "hear" what he wanted in his mind, and later worked to translate what he “heard” into symbols to create his
Leon’s composition was not just randomly stringing together numbers that “sounded cool” but mimics exactly what real musicians do to invent.

* * *

By way of Dalcroze Eurhythmics, this chapter first introduced a theoretical grounding for this thesis. To contextualize my projects and to generalize beyond them, I presented a design framework for music learning based on the ideas of Dalcroze. Finally, I connected Dalcrozian ideas of music learning to existing theories about learning at large. The next chapter delves more into the connection between music learning and learning at large, looking beyond theory and into practice.
Chapter 8: Mathetic Principles from Music

Seymour Papert once wrote that learning anything leads to two types of knowledge\textsuperscript{128}. One is knowledge of the subject at hand, and the other is knowledge about how to learn. Referring to the latter, Papert coined the word “mathetics”, meaning “the art of learning”, where -\textit{math} is the Greek root for “having to do with learning” (e.g. polymath). This chapter examines the mathetic insights that could be gained from learning to play music.

According to Constructionism, knowledge is not transmitted ready-made but is actively “constructed” by learners in their own minds. Music is no exception, as demonstrated in previous chapters. However, constructing musical knowledge differs significantly from “classic” constructionist activities. “Classic” constructionist learning involves making projects in the physical world (or in the computer), which in turn builds up knowledge structures in the mind. It also stresses the importance of playing with models, or representations, either in the physical world (e.g. gears) or in digital environments (e.g. Logo).

In contrast, playing music constructs no persistent external artifact. What we play at any given moment is an externalization of our mental structures at that time. Unlike external artifacts, mental structures are more elusive. They cannot be directly examined, only inferred, and only a small portion is revealed at each moment. Moreover, while examining models or representations could be helpful for understanding musical structure, the knowledge gained is not the same as what is required for actually playing music, much less playing with expression.

As an example, consider the rhythmic pattern of “2-against-3”. A typical representation, such as explored by the work of Bamberger\textsuperscript{129}, might look something like this:
This representation clarifies what is going on with the two layers of sounds in time and also unveils the mathematical patterns behind this musical structure. While it is certainly useful (and beautiful), this way of knowing is quite different from being able to recognize the “2-against-3” pattern by ear, which is yet different from being able to play it.

Another difference between classic “making” activities and making music is the element of time. Continuing our example of “2-against-3”, even for listening there is a difference between the ability to instantly recognize the rhythmic pattern when it occurs in a song and having to count and reason it through before arriving at the answer. The latter, slower, type of knowing is usually sufficient for making projects or for everyday problem solving, but it is not for playing music, which has the additional constraint of time. That said, it is interesting to point out that experts in any field do tend to have the ability to instantly recognize patterns relevant for what they do while the less advanced need more time to think problems through. As examples, consider how master chess players could instantly “see” patterns in the configuration of pieces on the board and how mathematicians could recognize instantly recognize relationships between numbers\(^{130}\).

Returning once again to our example of “2-against-3”, there’s yet another distinction between just managing to play the pattern (perhaps by meticulously counting) and playing it fluently and fluidly such that the people around you start to bop their heads or tap their feet (if they are not too stressed, tense, or uptight). The latter requires more precise control of the body in time, and the ability to find what jazz players like to call “the groove”\(^{131}\).

My point in drawing out these distinctions is to say that music involves different ways of knowing than other sorts of making\(^{132}\). The ways of knowing that come from creating and examining models could help with recognizing and playing our “2-against-3” pattern through counting, but using the pattern in actual music-making requires constructing different
sorts of “knowledge” in the mind. Regarding knowledge construction, I prefer to think of learning music as “building infrastructure” rather than building “mental models” or “mental representations”. Models and representations imply that knowledge consists of static structures. Infrastructure, on the other hand, maintain the structural aspect of knowledge, but also implies that the structures are actively used, like the roads of a city.

In fact, it could be argued that all knowing is more like building “mental infrastructures” rather than “mental models” or “mental representations”. Music’s peculiarities simply draws more attention to the nature of its knowledge. It is by no means the only domain with “different ways of knowing”, and many other domains involve building similar sorts of infrastructures in the mind. However, neuro-imaging studies have shown that playing music “lights up” more parts of the brain in synton than any other activity. This suggests that the infrastructures built by learning to play music do stand out as more wide-spread and all-encompassing. Thus, aside from all the well-known benefits of learning music—for cognition, health, and enjoyment—we could add that it is particularly mathetically valuable.

The rest of this chapter is divided into two parts. The first part describes three strategies for reaching and maintaining states of mind that are optimal for building mental infrastructures. While the previous chapter on Dalcroze examined activities to build mental infrastructures—as seen from the outside—these sections focus on the subjective experience of learning. The second part details how to build advanced mental infrastructure through the example of learning to play a classical piece. This example demonstrates not only the role of basic mental infrastructures (e.g. listening, using the body, and the imagination) but also how learning music makes use of classic “computational” thinking.

**Conditions for Building the Mind**

Classic Constructionism is about learning through building projects in the world. Emphasis is placed on creating the right external environments that support learners in building projects. For music, the building occurs primarily in the mind. Correspondingly, we could set up the right environment in the mind to support our learning. Doing so requires recognizing one’s own subjective state of mind, but until recently Western scientific thought has historically avoided discussions of subjective experience.
Nevertheless, awareness of subjective mental state plays a central role in the learning of many subjects involving the senses, the body, and creative expression. These strategies for managing states of mind are well known and have been cultivated in the arts, athletics, and Eastern contemplative traditions. Here, I discuss three such strategies, relating my own experiences of learning music to meditative practices and other examples of learning across domains.

**Tuning In**

It is well known that listening is an important part of playing music. In terms of “mental construction”, learning to listen means building infrastructures that attune to various features and layers in music. It involves building up the ability to instantly recognize patterns such as rhythms and harmonies, on one hand, and on the other hand, attuning to the expressive shape and quality of sounds. In order to build these mental structures in the first place, we must first learn to relate to sound in a way that is very different from how most people listen.

To help clarify the distinction, I refer to the musician’s way of relating to sound as “tuning in”. In high school, my piano teacher would constantly tell me to “listen!”, sometimes with frustration. I never knew what to do because I thought that I was listening. Whenever I listened to a piece of music, whether my own or the playing of others, I only paid attention to the notes. I did not know what it was like to “tune in”, neither to the patterns behind the notes, nor to the expressive qualities of what I was playing. And when I played, if I was pronouncing words but not understanding at all their meaning.

![Figure X: Unfortunately, even trained music students don't always know how to tune in](image)
The mental quality of “tuning in” is exactly what meditation trains. A classic meditation exercise involves closing the eyes and tuning in to sounds of the environment. If you have never done it, or even if you have, try it for a moment. Close your eyes, start to focus on your breathing, and then let sounds in your surroundings come to you. Perhaps you hear sounds in the distance, like cars or people’s voices. Or maybe there are sounds close to you, like the hum of your appliances. Maybe you continue to hear your breathing or even your own heartbeat.

Both meditation and music playing also involve tuning in to the body. In meditation, another common exercise is the body-scan, where attention is sent to each part of the body in turn, moving from the feet all the way to the top of the head, to “feel” what is there. What is there may be tension, or relaxation, comfort or discomfort, as well as sensations, such as tingling or warmth.

When playing music, holding tension in the body prevents most people from playing expressively. When the body is tense, playing sounds still and unnatural. Playing expressively also involves feeling emotions in the body and channeling them through our movements. Learning to “tune in” to the body, as in mediation, is therefore a pre-requisite to expressive playing.

Given the similarities between music and meditation for “tuning in”, we could say that the two are “inter-mathetic”. Learning one could then transfer to the other, as long as we recognize the connection. “Tuning in” to the senses is also important for other artistic fields. The art of cooking involves “tuning in” to taste and smells, and the visual arts are about learning to see. Regarding the body, dance, sports, and martial arts all have to do with “tuning in” to physical sensation and movement.

**Sensory-motor Coupling**

In everyday life, our senses and our capability for movement work together to enable us to navigate the physical world, interact with objects, and communicate with people. In music, as well as other expressive activities, involve developing a closer coupling between what we perceive and how we move. For example, in Dalcroze Eurhythmics, children learn to “tune in” to sound and to reflect what they perceive in their physical movements. The connection between perception and movement simultaneously builds and links the mental infrastructures for listening and for moving the body.
Connecting perception and movement takes a lot of focus and leaves no room for mental chatter, making it a sort of meditative activity. It is exactly the same strategy employed by Tim Gallwey, author of the *Inner Game of Tennis*\textsuperscript{139}. On a TV segment, Gallwey successfully taught a 50-year-old overweight woman who had not exercised for twenty-years to proficiently play tennis in under 20 minutes. In the very first activity, Gallwey instructs the woman to move her gaze to follow the trajectory of the ball.

We also find a close coupling of perception and movement in other artistic activities. The very first exercise in *Drawing on the Right Side of the Brain* by Betty Edwards involves echoing the slow movements of the eyes around the contour of an object with the movement of a pencil on paper\textsuperscript{140}. As in the case of Gallwey and tennis, Edwards was able to train people who never believed that they could draw to produce life-like drawings using her methods.

A special case of sensory-motor coupling is imitation. Humans, along with other primates\textsuperscript{141}, seem to have special neural mechanisms to imitate the movements of others. Imitation is how infants learn how to speak and how to move around in the world. As a meditative activity, the sensory-motor coupling of imitation could be especially effective for learning in the audio domain. For example, my fluency in French was greatly improved by listening to the radio and trying to emulate as closely as I could the sounds that I heard. I did not pay attention to the vocabulary or the grammar but simply tried to copy the voices on the radio with my own voice, listening to myself and making adjustments as I went along.

On the piano, I would learn new techniques and new ways of expression by playing along with famous performances, such as Glenn Gould’s rendition of Bach. I pretended that the recording were sounds in my mind, and my body would try to find the movements on the piano to render what I heard.

**Making Variations**

Both “tuning in” and “sensory-motor coupling” describe ways to keep the mind fully engaged so that thoughts do not wander. A third strategy for keeping the mind fully engaged was mentioned in Chapter 1 but deserves another mention here, in the context of mathetics. In Chapter 1, I discussed the fallacy of “typical” music practice as hours of tedious repetition and invoked the ski slope analogy of the mind to explain the
importance of introducing variations to avoid mindless repetition. Variations not only keeps the mind engaged but also creates more flexible and connected mental infrastructures, which for music results in more organic and expressive playing.

When making variations, maintaining the right state of mind is more important than the appearance of change from the outside. When I am practicing a musical pattern, the goal is not just to arbitrarily add complexity to the pattern for the sake of adding variation. Rather, the goal is the keep the mind fully engaged. If there is enough for the mind to attend to, I don’t necessarily need to change what I am playing. But as soon as I notice myself becoming bored, I then introduce a small change to what I am playing to keep the mind alert and interested. When playing, it is also important to continue to “tune in”, both to the sounds from the instrument and to the state of the body.

The combination of “tuning in” to the body and adding variation is also found in the practice of yoga. Vinyasa yoga is based around repeating sequences of movements called “flows”, such as the Sun Salutation. Even though movements are repeated, there are infinite details in the state of the body and quality of movement within which to direct attention. Any portion of the flow could be adjusted if it is too difficult or too easy. In yoga as in music performance, it is important to continue to breathe and monitor the body to prevent tension from creeping in.

Knowing what variations to add involves learning to “tune in” to the imagination. In the case of music, we could “hear” sounds in the mind. One thing I often do to practice improvisation is to hold a simple rhythmic pattern with the left hand and try to create variations in the right. To know what to play in the right hand, I “tune in” to sounds I hear in my mind and try to re-create it on the piano. If what is in my mind is too difficult for me to play, I simplify it or play only a portion. Whenever I play something, I always compare it with what I had in mind, and make adjustments until it matches the mental image. When doing this exercise, it is important to not try to hard. Trying too hard results in tension in the body, which blocks the creative imagination.

Douglas Hofstadter once wrote that variation is the essence of creativity. Beethoven was famous for his variations on simple themes, which he composed through improvising on the piano. In fact, variation at the heart of how any piece of music unfolds, regardless of genre. We also find the art of variation in other creative pursuits, such as cooking,
Reflecting Music Through Movement

fashion design, and visual arts. Even in subjects without such a close coupling of perception and movement as in music and yoga, the art of variation involves listening to the imagination.

**Building and Debugging the Mind**

The previous sections showed how, when learning to play music, we could familiarize ourselves with attaining and maintaining certain states of mind that are ideal for constructing mental infrastructures. Such states of mind have parallels in other domains, such as meditation, arts, and athletics. Learning to play music is one of the many entry-points for learning to manage states of mind, which has high mathetic value across the many other domains discussed.

The following sections demonstrate another category of mathetic insights that could be learned through music, that is not always so prevalent in other artistic, athletic, and contemplative domains. These are strategies of thought usually associated with problem solving and computational thinking: divide and conquer, making a sketch, reusing what you know, and debugging.

To demonstrate these principles in action, I will give the example of how to learn a classical piece on the piano. This is not the only instance in music learning where these principles occur, but I chose it because it is a familiar problem for anyone who has had some experience taking classical music lessons. It also gives us an opportunity to circle back and re-address the problems presented in Chapter 1 in light of what we have since discussed about building mental infrastructures and managing states of mind.

**Divide and Conquer + Making a Sketch**

One of Polya’s heuristics of problem solving is to try solving a simpler problem. This could involve breaking the problem into smaller parts and solving them one by one. Another approach to simplifying, according to Seymour Papert, is to create a “sketch”\(^{145}\). The sketch that may not address the entire situation but creates a “quick and dirty” way of getting to know the problem or project. Both approaches come into play in our example of learning a piece.

Recall that followers of pop-ed music learning prioritized the notes of a piece, and tried to “get them into the fingers” as soon as possible. A better way of starting to learn the piece is to apply the strategies of dividing and
simplifying. One way of dividing is to recognize how the piece could be split into distinct sections. Another way is to divide by layer, not just right hand part versus left hand part but delineating the individual musical line. A line could be split between both hand, or one hand may play multiple lines. Another approach is to simplify the piece by function. For example, only play the melody, or only play through the harmonic changes.

All of these methods of practice are equally valuable and should all be explored to gain different perspectives about the piece. Figuring out how to divide a piece requires “tuning in”. While the written score is available for reference, the ultimate judge of how sections, melodies, and lines are structured remains with the ear.

**Reusing What you Know**

Once we have built a mental image of how our piece should sound, we begin to fill in more details. An important principle in this phase is to reuse material and not approach everything as entirely new sequences of notes. This principle is also important in programming, where students are instructed to write code in a modular way. Commonly used material is grouped under functions, which could easily be reused.

We also could approach brand new material with this principle in mind. While we cannot “write functions” that neatly encapsulate our physical abilities to play, we could learn to recognize material that we have previously encountered, such as scales and chord progressions. Simply recognizing the higher level building blocks of a piece speeds up our learning.

Another part of reusing material is to learn new passages in such a way that could be reusable for the future. Once we recognize the underlying pattern behind a specific passage, say a scale, we practice it by introducing variations. This way, we do not only learn the specific notes of our pattern but also build more flexible, well-connected infrastructures for the pattern. The variations do not need to be complex. For the less advanced, it is enough to try playing it at different octaves, different speeds, and with different feeling. Making variations also helps us to discover our own interpretation of the piece.

**Debugging**

When we are writing code and we encounter a problem, we do not just repeatedly run our program, assuming that it would fix itself. Instead, we
stop to think through what might be causing the problem, and once we have a hypothesis, we design and test solutions until the problem goes away.

We could follow a similar approach when there is a glitch in our playing. When I say “glitch”, I do not just mean wrong note. A glitch is whenever what we play does not match the mental image of how we would like our piece to sound. It could be uneven playing, shaky rhythm, or expression not coming through. A sure signal to stop and debug is when repeating a passage several times results in no improvement. It means that we are not understanding something about our playing and should change what we are doing.

Recognizing glitches requires the ability to “tune in” and the ability to form a detailed mental picture of how a piece should sound (though it always helps to record ourselves and listen afterwards). Once we recognize a spot to debug, we reason through what is causing the problem. It could be an awkward fingering, or perhaps we lack the technique to play the passage, or we may not understand exactly what is going on between the voices. Once we have a hypothesis, we try to design exercises (or games!) to target the problem. When playing through our games, it is important to tune in to our playing and to make sure that tension does not creep into our bodies. Adding variation to our games is another way of keeping the mind engaged, especially when practicing a technical problem.

**Meta-Mathetics: Learning Across Domains**

The example of putting together a classical piece shows how learning to play music could be another way through which to gain exposure to computational or “systems” thinking. Alternatively, people who already understand the tenets of systems thinking could apply these strategies toward learning to play music.

When applying “systems thinking” for music, it is crucial to understand that the systems we are building and managing are within ourselves. Understanding how to build infrastructure of the mind relating perception, movement, imagination, and emotion is immensely valuable not only for music and for learning but for life. Knowing how to manage states of mind and to “debug” our own thoughts and emotions ultimately helps us live more fulfilled and flourishing lives.
While classic construction projects in the world remain an excellent way to learn in a variety of subjects, the learning examples given in this chapter suggest that we also need to pay attention to building structures in the mind. To do so, it would help to have a better vocabulary with which to precisely talk about structures of the mind.

In *The Society of Mind*, Marvin Minsky proposes one way to precisely model mental structures. Though Minsky's ideas were originally developed for the purpose of Artificial Intelligence, they are equally relevant for human learning. Interestingly, Minsky worked out for himself how to improvise fugues on the piano as a way to learn about the workings of the mind. Minsky's experience certainly points to the mathetic value of learning to play music!

![Figure 9.2: Mental infrastructure likely to result from learning in the pop-ed approach](image)

While the models presented in *The Society of Mind* could be used to describe mental infrastructure for mathetic purposes, its abstract, symbolic nature may not be the best tool for everyone. The above diagrams were my attempt to model different sorts of musical minds based on Minsky's ideas. While creating these models was a helpful for
me to articulate my ideas about music and the mind, I found it difficult to convey these models to others. The difficulty came from the lack of grounding of the abstract description in their own concrete experience.

Perhaps a more accessible way to speak about structures of the mind is through metaphor. As Lakoff and Johnson demonstrated in their linguistic studies, abstract concepts are often understood in terms of the more concrete. In fact, the idea of Constructionism, of “building structures” in the mind is already a metaphor, and as Seymour Papert acknowledges, constructing is not the only metaphor for learning. Another metaphor that Papert proposed is one of cultivating a garden.

My favorite mathetic metaphor is one of building invisible cities of the mind (see appendix B). A city consists of not only buildings but also roads (e.g. structures and infrastructure). We could think of playing music, or any sort of performance as giving a tour of a part of the city. We could either plan an exact route or explore as we go along (as long as roads exist!). The city metaphor also accounts for states of mind. We could think of our emotions as weather patterns. If there is too much fog, parts of our city may not be accessible. Having too many thoughts at once might be related to a traffic jam. Finally, we could think of learning different subjects as building different cities. Each city could be structured differently, but once we learned how to build one city, we could use what
we know to more rapidly build new cities. We could also build roads in between our cities to connect our entire mental realm.

In today’s fast-paced, interconnected world, it is increasingly important to excel in the art of learning, not within individual subjects but across diverse disciplines. While many words exist to describe working across, between, and even outside traditional disciplines (e.g. cross-disciplinary, trans-disciplinary, even anti-disciplinary), no good word exists in English to describe the corresponding sort of learning.

Learning across disciplines is not just about learning many individual things, as implied by the word polymath. Rather, the key is finding “higher-order knacks”, as Marvin puts it, that help us connect and organize many domains of knowledge. This chapter attempted to describe some of these “higher-order knacks” that could be gained from learning to play music. To describe these “higher-order knacks”, I propose the word “metamathetics”. Metamathetics evokes the concept of “metaphor”, which literally means “to transfer” in Greek.

It could be argued that all mathetics is metamathetics, but the “meta” makes more explicit the connections across domains. Adding the prefix also conveniently gives us the bonus word “metamath” to describe one who learns by seeing higher level patterns and by transferring existing knowledge across domains. In reality, a polymath is very likely to also be a metamath. But while “polymath” emphasizes only the multiplicity of one’s knowledge, leaving a mystery of how the knowledge came to be, “metamath” emphasizes the connections.

This chapter focused on uncovering some of the metamathetic connections between music and other domains. The next chapter will present a vision for metamathetics of the future and the role of interactive technologies.
Chapter 9: As We May Play: A Vision and Variations

According to my advisor Hiroshi Ishii, research in computing can be categorized as to whether it is in pursuit of technology, application or vision. Technology-centric research seeks to invent new technologies or optimize existing ones\textsuperscript{147}. Examples range from designing a new computer chip to devising improved algorithms to speed up data transfer. The second category builds applications using existing technologies to fulfill some needs of users. Almost all existing music learning technologies as reviewed in Chapter 2 fall under this category.

In contrast, vision-based research looks beyond currently existing technologies and current common applications to imagine a world of the future. Rather than incrementally pushing forward existing trends, vision-based research imagines new applications and inspires new directions for technological development. Hiroshi’s own visions include Tangible Bits and Radical Atoms. As discussed in Chapter 3, Tangible Bits imagines a world where digital technologies break free from the forms factors of traditional computing (e.g. screen, mouse, keyboard) to combine seamlessly with the physical world. Radical atoms envisions an even more distant future, where physical atoms become just as dynamic and programmable as bits.

Vision based-research is the least common in the world, often due to practice constraints of industry or to researchers’ desire to see immediate, real-world results. However, it is just as necessary and important as research focused on technologies and applications. Furthermore, vision-based research is also the most enduring. Technologies become obsolete every one or two years. New applications emerge every five to ten years to edge out the old. But vision has the ability to power human inquiry for over 100 years. In Chapter 3, we saw the vision of Charles Babbage about the programmability of calculating machines, and later Alan Turing
whose vision became the bedrock for our modern-day, general-purpose computers.

**Humanist Computing**

Many of the greatest visions for the computer have been profoundly humanist. In 1945, Vannevar Bush penned the essay “As We May Think”, outlining how machines could become tools to augment human intelligence\(^{148}\). Inspired by this essay, pioneers of Human Computer Interaction (HCI) like Ivan Sutherland and Douglas Englebart envisioned radically new applications for the computer in the 1950s and 60s, including drawing, word-processing, and video-conferencing that have now become staples of today’s computers\(^{149}\).

Although the vision of Artificial Intelligence is often seen to be techno-centric, its conception, at least according to Marvin Minsky, is also fundamentally humanist. Minsky sought to build intelligent machines to liberate humans from necessary activities of life that are dangerous or dull. And in the process of building these machines, we also gain a deeper understanding of the human mind.

Alongside the visions of HCI and AI is Seymour Papert’s vision for the computer in education, which also seeks to augment the human mind by way of learning but without longterm reliance on the machine. I call these three visions as the 3 pillars of “Humanist Computing”. Not one approach is superior to the others. All are necessary, and the efforts within each informs the others.

My own work has inherited from HCI questions of how to represent, interact with, and collaborate around information. Particularly from Hiroshi Ishii’s HCI vision, I have inherited an approach to design that aims to seamlessly blend the physical and digital worlds and to engage many sensory modalities. Though my work appears the farthest from AI, Marvin Minsky has inspired me to to probe the intricacies of the human mind, and I have followed his footsteps in using music as a means to gain insight about the mind.

Of the three pillars of Humanist Computing, my vision for is most similar to the vision of Papert. Interestingly, Papert was by no means the only pioneer of using the computer to support human learning. However, while others worked to adapt the computer into models of existing classroom “applications” such as lessons and drills, Papert envisioned
technology as the vehicle for a radically different way to learn—for learners themselves to construct their own knowledge.

It is important to note that the way of learning espoused by Papert is not new, nor was he the first to observe it, characterize it, and give it a name. That honor goes to Jean Piaget, who opened new grounds in the study of the mind through his careful observations of children and experiments on how children think. And learning by “constructing” was originally a concept from Piaget, which he also did not invent. Children, and adults who have maintained a “beginner’s mind”, have always and still learn in this way.

On one hand, the technologies that I have created could enable more people to learn music more effectively and enjoyably. On the other hand, I seek to disseminate ways of “thinking” and “knowing” from music, which could also be applied to learning in general. “Thinking” like a musician involves learning about musical building blocks, but it is also not only about making patterns in the mathematical sense. Through my projects, I have tried to convey what it “feels” like to play music and to play with music, which has to do not only with mathematical patterns but also with the body, emotions, imagination, and movement through time.

This way of learning music is not “new” and is often how the most “talented” approach music “naturally” without being able to articulate what it is they do. However, articulating these methods and designing interactive experiences based on them (as Piaget and Papert did) empowers more people with better ways to learn. And as discussed in Chapter 8, these approaches from music are not only relevant for music learning itself but could become a powerful source of insights for learning everything else.

One question that remains is how to bring the ideas explored in this thesis into the world. The “enabling technologies” on which I have implemented the technology-side of my vision were chosen for their ability to emulate the experiential dimensions of my designs, rather than for the ability to cheaply and easily scale. Unfortunately, digitally controlled player pianos remain quite expensive at the moment, and thus my projects are still a ways from being integrated into daily life.

The good news is that in the world of technology, there is a trend for devices to become less expensive and more widespread over time. Once upon a time, it was unimaginable for a computer to be in every home because they were so large and so expensive. Just because something has
limited availability today does not mean it will forever remain so. And in fact, the vision of Radical Atoms describes a trend for actuators to become cheaper and more embedded into everyday objects. In fact, we could even say that the player piano is one of the earliest examples of Radical Atoms!

**FÜGO: A Platform for Playing with Music**

The final pages of this chapter shares some thoughts of how my projects might scale into an everyday product in the future. I assume a future where actuators and displays will be cheap and widespread, and where sound synthesis will be advanced enough so that we could have a portable electronic player piano augmented with pixels. I will call it FÜGO. A series of sketches will illustrate how ideas of MirrorFugue and Andante might extend to the FÜGO platform, as well as how the platform could become a part of a wider ecosystem of both hardware and software.

Each FÜGO unit consists of a digital player-piano keyboard (with moving keys), speakers to output sound, a vertical “fallboard” screen, a camera facing down onto the keyboard, and of course an embedded computer. Like the glass screens of many laptops and smartphones, the FÜGO screen is reflective when off, acting as a natural mirror like the surfaces of lacquered pianos. The surface of each key could also have display capabilities.

Each FÜGO consists only of two octaves, to decrease the price-point of initial investment. Multiple units could be connected to form a larger keyboard, and each single unit folds up into a portable case. Additional
display could also be connected to the system, such as one in the location of the music stand.

We could also imagine the display and keyboard portions of each unit being able to detach. Several displays, when detached, could fit on an acoustic grand piano and replace the normal fallboard.

**MirrorFugue**

Both MirrorFugue and Andante could be ported to FUGO. The full version of MirrorFugue would make use of the additional music stand display surface, but recordings of only the hands could still be played back on the basic FUGO. Original recordings could be made either on FUGO or on acoustic pianos.

Additionally, MirrorFugue would integrate with a software app. The app could be directly accessed through FUGO or from a tablet device, which could serve as a remote control. The app could allow the user to select performances from different artists on FUGO. Each user could also have their own profile, where they post their own recordings, which might
serve as a personal progress log. The could also save their favorite recordings from other users or artists.

The app might also have a social component, where learners could find and connect with peers and mentors. The connection might be asynchronous, in the form of following and commenting, or learners could use the app as a way to schedule lessons or jam sessions over distance.

The app could also serve as a learning resource for different pieces of music, offering MirrorFugue recordings of performances, lessons, exercises, as well as the score and other analytical representations of the piece.

Figure 9.6: (Left) library of recordings, (center) user profile, (right) materials for pieces

**Andante**

Similarly, Andante could be ported onto FUGO and combined with an app. Through the app, the user might select different characters with which to play back musical phrases. There could also be a set of mini-games through which players learn about different components of music.

Figure 9.7: Andante and accompanying app

**More Microworlds**

Beyond MirrorFugue and Andante, FUGO could become a platform for countless other piano-learning applications, games, or “microworlds”. Exploring each world might illuminate some dimension of music through
Reflecting Music Through Movement

a concrete metaphor. We could concretize the rise and fall of phrases with physics simulations or the “surges” of dynamics through a metaphor of waves.

Best of all, FUGO could also enable users to share their ways of thinking about music by programming their own visualizations or activities. This way, creative expression through music could become an entry point for students to learn about how to program the computer. By combining music learning with programming allows learners to draw powerful parallels between the two domains.

![Figure 9.8: Many more microwords on FUGO and the ability to program your own](image)

**Beyond Music**

The final portion of this chapter extends key concepts from my piano projects into learning for other subjects and envisions future interfaces. One concept is the creation of “sandboxes” in the physical world, which combines the “microworld” concept of Seymour Papert with principles of Tangible Interfaces to allow a vocabulary of self-guided interactions that combine the physical and the digital. The second concept is about establishing the connection between enactive, iconic, and symbolic representations, such as explored through Andante. Finally, both MirrorFugue and Andante provide opportunities for the user to “shift perspectives”. MirrorFugue enables the user to project oneself into the point of view of the virtual pianist while Andante enables children to “miniaturize” themselves, by going into the perspective of the walking figures. Three projects prototyped on the shape displays of the Tangible Media Group explored each of these concepts, respectively, for learning beyond the piano in the future world of Radical Atoms.

**Physical/Digital Sandbox: Kinéphone**

The first project, is a “sandbox” for constructive music-making with everyday objects. Each pin of the shape display could be augmented with
“caps” filled with sound-making everyday objects such as beads and bells. Pins could also act as stands for larger sound-making objects such as xylophone blocks and chimes, where other pins act as strikers. These instruments could be played either directly or by programming.

Kinéphone was designed to explore music-making outside of the constraints of piano technique and the patterns of classical melody and harmony. It gives the opportunity for learners to explore not only sounds from physical objects but also instrument and interaction design.

![Image: Instruments prototyped on Kinéphone](image)

**Figure 9.9: Instruments prototyped on Kinéphone**

**Embody & Symbolic Representations: xForm**

xForm is based on a platform similar to FUGO but with a pin-based actuated shape display instead of a piano keyboard. Each pin is augmented with display capabilities on all its surfaces, and a vertical screen extends from the back edge of the shape display. To detect user input, pins have the ability to sense touch, and the entire system could detect the user’s mid-air hand gestures.

On one hand, the shape display allows the user to “sculpt” shapes via direct manipulation and mid-air gesture techniques. A user may also stage animations using direct manipulation. For example, a user could define the trajectory of a ball by moving it along a path, which which “programs” the shape display’s pins to push the ball along the same trajectory. Both sets of interactions have previous been explored by research projects in the group\(^2\). However, the mathematics and the code of how the shape display moves have been hidden from the user, following the convention of HCI, which differentiate “users” of the computer from “programmers”.

To adapt these interactions for a learning context, the symbolic commands associated with shape change or movements on the shape display could be made available on the screen whenever the user
physically manipulates the shape display. These commands could then be altered by tweaking their parameters, which in turn affects the physical state of the shape display. In a simple example, a user might create a “dome” shape by cupping the hands, which would trigger the command on the screen. The user might then alter the center or radius of the dome. In the process, users learn about the connection between embodied geometry and their corresponding symbolic specifications. For more playful learning, users could program their own physical games.

I began designing and implementing an API of these commands were explored along with two undergraduates during the fall of 2015. This project is still ongoing.

Shifting Perspectives: MacroScope

MacroScope is an application for exploring architectural spaces. Its hardware consists of the shape display, the vertical screen, and a Virtual Reality (VR) headset swivel-mounted on a stationary stand, which I call the “scope”.

The shape display physically renders a portion of a space while the vertical screen acts as an extension of the physical display. The physical enables haptic interactions while the virtual shows more details, such as the location of furniture or simulations of characters. Looking into the VR headset reveals a virtual “double” of the space. However, unlike typical VR, which is mostly limited to the visual and audio modalities, what is seen through the “scope” could also be felt by the hands.

The “scope” allows the shifting of perspectives from a “bird’s-eye view” of the model to a view from inside the model. I refer to these respectively as “maquette view” and “avatar view”. Maquette view is the default view of physical models from above. To enter into “avatar view”, the user
could place their hands anywhere inside the shape display model and “walk” their fingers. The “scope” then shifts to reveal the perspective as if the user were standing where they have placed their fingers in the model. They could then “look around” inside the model by turning the “scope”.

MacroScope was built as a demo for the 2015 fall member meeting at the lab. It is still an ongoing project, and will soon be re-implemented using a new generation of shape display and VR platforms.

***

According to Mitch Resnick, that too much of education technology has been the equivalent of stereos and CDs instead of “computational pianos”\textsuperscript{153}. The same is true of interactive systems in general, and many interactive systems that do engage the user’s abilities tend to place high demands on cognitive skills and neglect the opportunity to build bodily skill\textsuperscript{154}. As computation expands more and more into the physical world, the metaphor of the “computational piano” becomes increasingly relevant as a foundation for interaction design. By extension, a better understanding of playing the piano, beyond the usual assumptions, could then inform and enrich the design of interactive systems.
Conclusions

This thesis has examined the questions of how to design technologies to support music learning and how music learning could bring new insights to learning at large. My research has taken a “contrapuntal” approach, weaving together my experiences from learning to play the piano and my background in computer science and human computer interaction with a keen interest in the workings of the mind. By combining disparate points of view and connecting diverse domains, it has contributed the following:

- A new perspective to the problems of music learning:
  - Teased apart widespread, misguided assumptions about learning music, which I termed “Pop-Ed Music Culture” following Papert’s concept of “Pop-Ed Culture”
  - Identified a gap in the space of existing music learning technologies in how systems could support learning to play an instrument beyond the notes, and how to play with music on an instrument to a high level
  - Offered a theory on how neglect of the body in technologies for music dissemination contributed to the rise of pop-ed music culture

- Designs of new interactive technologies that demonstrate the role of the body, imitation, and empathy in musical understanding:
  - Enabled users to empathetically feel the performance of a virtual pianist in their own bodies in MirrorFugue
  - Conveyed the link between movement of the body and musical structure in Andante
  - Put forth a theoretical framework for both projects rooted in Dalcroze Eurhythmics

- Bridges between music learning and learning at large
  - Connected a theoretical foundation of music learning to general education theories
Reflecting Music Through Movement

- Derived mathetic principles from music and gave examples of how they occur in other domains
- Articulated a view of learning as building infrastructure in the mind

By way of a conclusion, I offer some final reflections on a central topic of this thesis: the relationship between the mind and the body.

There are some widespread misconceptions about the nature of the body deeply rooted within Western cultural beliefs. Western philosophy has traditionally drawn a sharp line separating the mind and the body. The human mind was seen as the source of our unique ability for abstract reasoning, which elevated us above all other creatures, who could only act and react through their bodies. The human body, in contrast, was seen as the source of our base desires, our “animal nature”, and was at once fetishized and feared. To this day, the mind and the body are often held in a false opposition, exemplified by a common belief that athletes, and others who spend time to develop physical skill, are less “smart” than those who focus on purely cerebral pursuits.

These views extend to the way that our society educates children. While young children blithely explore the world through their bodies, and learn through physical play, they’re taught to sit still and study as they get older, to focus on their minds—as if body and mind were mutually exclusive. Even Piaget viewed the emergence of abstract, rational thought as the end-goal of childhood cognitive development. Perhaps as a result of this cultural disposition, people often adopt an attitude of “learned helplessness” toward their bodies as they reach adulthood, believing that they just “can’t dance”, “can’t play sports”, or “can’t be good with their hands”.

To reconcile mind and body, we can look to present-day scientific knowledge about the brain. Speaking in terms of the brain, we can clearly point to a separation between the brain organ and say that everything else constitutes our “body”. But if we speak in terms of the “mind”, which Minsky once summarized as “what the brain does”, we have to take into account that a large part of what the brain does is to enable interaction with the world. In fact, neuroscientists have theorized that the brain evolved expressly to allow creatures—both humans and other animals—to move around in the world and to deal with constantly changing environments.
The body is how the brain interfaces with the world, and it is not only the brain but our “nervous system” that enables all cognition. Even though the brain, or central nervous system, is a distinct organ, the peripheral nervous system extends into all parts of the body. Therefore, we must expand the notion of “the mind” to encompass the relationship between the brain and the body, as well as the interaction of both with the physical world. This is the core idea underlying the field of embodied cognition.\textsuperscript{158}

Interestingly, Marvin Minsky’s theory of mind articulated in his book \textit{The Society of Mind} includes the body’s capabilities in its models of cognitive processes. The classic example of a child playing with blocks example is not about abstract thought but about how minds could interact with object in the physical world. It models a Builder agent, which in turn employs sub-agents such as Add-Block, which has its own sub-processes to Find and Place blocks. Find and Place encapsulate lower-level sensory and motor processes, treating them as black boxes but acknowledging their place in the overall system.

Sensory and motor capabilities themselves could be broken down into hierarchical systems, giving us a more precise language to describe functions of the “mind-body”. Our sensory perceptions (e.g. sight and hearing, etc.) emerge from the activities of neurons that respond to stimuli such as light and sound. For vision, groups of neurons perform “lower-level” functions such as edge and color detection. Results then feed into progressively “higher-level” processes, eventually enabling the mind to recognize objects, faces, and places.\textsuperscript{159}

For our motor capabilities, while neurons are responsible for actuating individual muscles, muscles usually work in coordinated groups. For example, locating and moving one specific muscle of the face takes much concentration, but smiling, which uses many different muscles, happens effortlessly when we feel happy. Some activities involve many different groups of muscles. These are often referred to as “full body” activities in everyday speech. A basic example is walking, which involves many muscles in the feet, legs, hips, and the arms. Motor capabilities could also involve the interaction between the body and an external object, such as riding a bike. Often, full-body activities become easier once a rhythm emerges to coordinate the actuation of different muscles groups. For example, playing a regular beat enables Parkinson’s patients to walk normally, without their characteristic limp.\textsuperscript{160}
synchronize motor activity in the bodies of different individuals, such as in dancing and music playing.

The example of smiling also illustrates the relationship between motion and emotion, between our body and our feelings. It is difficult to “construct” a convincing smile by individually locating all the muscles involved, but when we imagine a friend, the conjured emotion instantaneously actuates all the different muscles to bring a smile to our face. Conversely, studies have shown that even constructing a fake smile could simulate feelings of happiness. Neuroscientist Antonio Damasio proposed that the body serves as a map onto which emotions are projected. In support of this idea, a cross-cultural study of 700 people found that common emotions elicit similar topographic signatures across the body.

According to Minsky, emotions alter which processes of the mind an individual has access to at any given time. For instance, Anger selects processes that enable a creature to react with more speed and strength but suppresses processes for careful planning. Learning to recognize how different emotions are felt in different parts of the body could then give people insights into their own affective states and how these states alter their cognitive processes. Damasio goes even further to emphasize that the ability to recognize feelings in the body plays an indispensable role in rational decision making. His theory is supported by numerous examples of clinical patients with damage to a specific part of their brain which took away their ability to experience feelings in their bodies. Even though their knowledge, attention, memory, language, and even ability for abstract thinking were otherwise flawless, these patients became severely impaired in practical reasoning and decision making.

Finally, the body even plays a role in high-level abstract thinking. Based on a comprehensive study of language, Lakoff and Johnson argue that much of human knowledge is understood based on existing knowledge by way of metaphor, and that abstract concepts are always framed in terms of the more concrete. For example, in English, the elusive concept of “time” is commonly understood in terms of the more familiar concept of “money”. This is evident, when we speak of “spending”, “wasting”, and “saving” time. If all knowledge builds metaphorically from existing knowledge, all conceptual knowledge could then be traced back to our original embodied experiences of existing in the physical world. In fact, the very notion of “high-level” and “low-level” maps an abstract concept
to a physical metaphor that is ultimately understood in terms of our body, and its relationship with the world. A connection with the body exists even for the most advanced theoretical physics and mathematics, as evidenced by Albert Einstein’s descriptions of how he arrived at what ultimately became his Theory of Relativity by feeling particular muscular sensations in his own body.6

I gathered these ideas about the mind-body not just as a cabinet of curiosities about cognition. Rather, it is all in pursuit of my original goal of deriving effective strategies to learn. To really understand how we learn, it is essential to examine not only the intellectual aspect of our mind but also how it relates to our body and even to our emotions. But only collecting cross-disciplinary research anecdotes under the (now trendy) banner of “embodied cognition” is insufficient. We need more concrete and more specific examples from the actual processes of learning complex subjects.

This is where music comes in. Music is a unique activity that is widely acknowledged to encompasses the physical, emotional, and intellectual dimensions of the mind-body. It is also unique in that even though most people do not profess to be musicians, they have an intuitive understanding and appreciation for music. This makes music an invaluable source of concrete and accessible examples to better understand the processes of learning and the workings of the mind-body. Better understanding of the interplay between different processes of the mind-body not only helps designers create better tools for music learning but also informs the design of learning technologies at large, especially the design of richer interactive experiences with the computer.
Appendix

A. Meditations from the Practice Room

In the fall of 2013, I started a small journal archiving the snippets of insights on various facets of piano playing. Intended as a book to flip through when feeling uninspired, I wrote entries as concise reminders with a paragraph of elaboration nudging myself toward the right space. This I had affectionately entitled Meditations from the Practice Room.

Ironically, I had never meditated in my life before, at least not in the traditional sense of sitting down and “training” the mind in the style of Buddhist monks and other contemplatives. It wasn’t until fall 2014 when, catalyzed by the Awareness class, I finally began a daily meditation practice. For two months straight, I opened each morning with ten to fifteen minutes of meditation via the Headspace app. Through the process, I delighted in discoveries of profound parallels between traditional meditation and my own musical practice, one enriching the other. The two practices provided different arenas to explore similar dimensions of the human experience alternating as shifts in perspective and reminders of nuance.

These reflections attempt to tease out the parallels between meditation and my piano practice. I discuss five themes—Listening, Breathing, Posture, Tone, and Resonance—which focus on the mind-body
connection in music playing and relate directly to themes I have encountered in my recent practice of traditional meditation.

Over the past five years, my musical practice has fundamentally changed how I approach learning, how I communicate, as well as how I consider personal well-being, particularly emotional health. I have noticed that people who have not known me for so long assume that I have always been this way, or see progress as an inevitability from years of study. To dispel this perspective, I begin by detailing my formerly mindless practice habits.

I began learning to play the piano when I was 4 years old, years before I spoke a word of English. I didn’t remember it as boring, but then again, I didn’t keep a regular practice schedule. Serious practice came at age 13, when I started a strict regimen of one to two hours a day, but I didn’t embark on my current practice until age 22 when I began a musical apprenticeship under pianist and composer Donal Fox. Prior, I had followed the typical classical tradition, learning to play pieces of increasing difficulty from written notation. This I did in the typical way, first programming notes into the fingers slowly, then repeating to train muscle memory, only adding expression after a certain baseline fluency.

Like most children on whom music lessons was inflicted, I was often bored. While many chose to end their misery after a handful of years, I saw boredom as a necessary cross to bear toward mastery. Taught by a Chinese upbringing to believe in the power of “hard work”, I often forced myself by sheer will to repeat pieces again and again, cementing muscle memory. In this way, I worked myself up to a fairly advanced repertoire in high school which I played to win statewide competitions that “officially” certified my skill. However, I could rarely play when spontaneous occasions arose as pieces quickly faded from memory. With each new piece, I started right from scratch, like Sisyphus once again rolling his boulder uphill.

These days, boredom during practice has become a foreign feeling. Without other obligations, I can easily practice through most of a day, continuously discovering new details, solving new problems, and inventing new games for myself. Though I still feel I am far from mastery, I feel quite comfortable sitting down at a moment’s notice to play from memory, improvise, or demonstrate various building blocks. Though I spend longer hours on the piano than ever before, I no longer feel that I
"work hard", in the sense of the oft-celebrated reliance on will power—I play the piano.

* * *

Beyond various dimensions of musical skill I have acquired over the past five years, the most fundamental shift was the one of mindset—from work to play. Before, my approach to music was motivated more by goals of playing correctly rather than the process of learning to play. I treated practice as work, narrowing my study to the notes of the page with a single-minded seriousness inherited from my former approach to school learning.

I have since realized that deeper learning means going beyond the information given to explore possibilities and personal connections to the material. These explorations require not just the intellectual mode but physical and emotional connection with the material, creating layered webs of understanding far beyond correct and consistent recitation.

While this distinction in mindsets may be concisely described, embodying it in practice require exercises that establish new patterns of behavior over time. I share some of my own exercises which guided me into my new mindset, presenting them under 5 themes that build from one to the next. Though my list is a convenient summary of ideas, it is important to remember that real understanding within a person exists more as pathways than aphorisms and that the real process, though trending toward improvement, is far from a linear, monotonic function. For each theme, I offer a description and a Meditation, distilled from my exercises.

1. Listening

As music is an art of sound, listening is the most fundamental skill. Although eardrums vibrate indiscriminately to surrounding frequencies, hearing is not the same as listening, where the mind tunes in to perceptual qualities of what we hear. Very often, amateur pianists are taught to attend to pitch, dynamics, and articulation of notes only to the extent that they are “correct” according to the written score, resulting in a shallow listening that neglects the quality of sound. Rather than listening for correctness, we can expand our perception to how the sound feels. Is it harsh or round? Is it bland or colorful? Does it sing? Does it speak? What does it say?
**Meditation:** Strike a key and hold it, letting it ring until you can no longer hear it. Try it again with the same key, striking it a bit differently. Keep doing this a few times with the same key or different keys, striking each with a different intention, varying the loudness, softness, and color. It's ok if the strike lacked the control to match its intention. Hold it and listen to it still. Notice changing qualities in the sound, but don't try to put it into words. Just perceive. After a while, try striking two keys at once and hearing how their waves collide. Are the sounds consonant or dissonant? Can you hear the beats between the frequencies?

2. **Breathing**

When listening, it is important not to try too hard, not to feel frustrated if you don’t hear something you think you should, and not to force yourself to reach for it. When playing music, working hard introduces strain in the body, which blocks the flow of sound. We tense up and forget to breathe. When listening and when playing, it is important to maintain a long, even breath. For me, practicing scales is a great way to practice breathing. The cognitive load of musical content is relatively low, which frees attention for the body.

**Meditation:** I started with the C major scale, played up and down across four octaves, but this can be done with any scale. Play the scale slowing with one or both hands, maintaining evenness of notes and evenness of breath. The breath does not have to be perfectly aligned with the length of the scale. Try to take deep breaths, allowing your body to expand and contract in a continuous movement.

3. **Posture**

When working on the computer, I always slouch, even in the most ergonomic chair. It’s as if all the energy is flowing toward my head, leaving the rest of my body deflated. I can try to straighten my back, but as soon as I forget to consciously prop myself up, I slouch back down. I used to slouch quite a bit when playing the piano and tolerated lower back pain as an unavoidable consequence of the piano bench. Now, I sit for hours on the same bench without a second thought, good posture, zero pain. An awareness firmly rooted in the body is all but engrained as instinct though on occasion, when preoccupied with a technical problem, I do still find myself “going into my head”, where my playing reduces to twiddling of the fingers and my body deflates.
**Meditation:** Start with the scale again, with attention on the breath. Feel where your torso is and where it wants to be. You may find that your breath will carry you into a straight posture. If it does not, do not force yourself. Continue playing, breathing, and going where your body wants to go. Play the scale a few times stronger or softer. Notice that your body follows, sitting up straighter with stronger playing, and bending over to follow softer sounds. Feel where the body wants to go, how it wants to move. Do not hold your body a certain way only because you feel you should.

4. **Tone**

We play music with the entire body, not just the peripheral parts directly responsible for actuating notes. However, many amateurs view the piano as a keyboard of pressure sensitive buttons to be played only with the fingers, whose velocity correspond to loudness. On an acoustic piano, increasing velocity yields a louder note but with increasing stridence. Pianists cultivate a keen sense of touch, using the weight of the fingers, hands, arm, and body for rich palette of sound. To play with weight touch, the body must not be tense. When practicing, I listen for harshness of sound and make adjustments to diffuse tension in my body.

**Meditation:** Play a scale not only as single notes but as thirds and octaves, which require the weight of the hand and arm. Fingerwork alone can produce single notes, but third and octaves are too strenuous to be played only with the fingers. Start with a scale in thirds, playing each third slowly and evenly, focusing on how the weight shifts within the hand. Keep the fingertips glued to the surface of the keys, with the fingers channeling weight into each key to produce sound. The fingers must develop strength to support weight from the body and may feel sore after a while, but you should not feel sharp pains or tension. Switch between thirds, octaves, and broken octaves, paying close attention to the quality of sound and feeling of weight from the body into the keys. Continue to breathe.

5. **Resonance**

Early in my exploration of tone and touch, my teacher had alluded to a phenomenon where the pianist feels sound resonating in various parts of the body: head, heart, gut, groin. I had previously encountered this idea in singing, where the body is the instrument, but as the piano is distinctly not part of the body, I wrote off this idea as an artifact of the artistic imagination. This fall, I have finally started to experience the difference
between resonance in the head, the chest, and the gut. Whether or not it is currently scientifically measurable, I can hear distinct differences in my sound on recordings. By default I’m very much in my head. After a heated conversation, I can always feel a familiar buzz circling my skull. Currently, I am still in the process of learning how to access and retain resonance lower in the body.

**Meditation:** This exercise tries to concentrate feeling of the sound in the chest and in the gut. Start with a simple chord spread between the hands. I usually play a triad in the right hand and an octave in the left. Play it with a full, weighted touch, allowing it to ring for a few seconds. Listen to the quality of sound and pay close attention to where you feel it in the body. Repeat the chord with adjustments each time. You may not tune into the feeling immediately, but do not cease to listen and breathe as you repeat. Chords around the middle register of the piano resonate in my chest. Lower chords I feel more in my gut. Once you get the feeling, play a scale or a piece while trying to retain the feeling.

* * *

I have observed many common concerns across meditation and my musical practice, such as attention to breath and posture. The quality of listening advocated by meditation correspond closely with how musicians start to listen, and the type of focused but relaxed attention described for meditation is exactly the sort of attention required for music performance. Even the locations in the body where music resonates relate to points where certain traditions of meditation focus energy.

During meditation, we take time aside to train the mind, familiarizing ourselves with modes of being that we gradually integrate into life at large. Similarly, mindsets from my musical practice has pervaded life away from the piano bench. In fact, the spectrum between my formerly goal-oriented approach to music and my current approach reflect certain patterns in how people approach life in general. Thus, I see music as a powerful sandbox in which to practice experiencing modes of being with broader life relevance.

Through music, I learned many crucial life lessons counter to prevailing societal beliefs. These reflections address three in particular. First is the difference between understanding something in theory, being able to recognize it, and embodying it in behavior. While understanding in theory may be reached quickly, a deeper understanding is the internalization, which require practice over time to build pathways. Second, the
conventional wisdom of “no pain no gain” does not always hold true. The brute force associated with “working hard” often introduces tensions and limits explorations required for deep understanding. Finally, understanding is not only of the mind. Cognition, not just of music, involve integrated processes of the physical, emotional, and intellectual.

Even from my limited exposure to traditional meditation, I feel that these lessons are very much consistent with its teachings. I also feel that music has the potential to complement a traditional meditation practice. Music is a cognitive process that does not resort to symbolic, representational language, which has allowed me to observe the mind without mapping thoughts to words. Playing music, especially improvising, yields an output from which we can measure and reflect on changing states of the mind and body in real time. Music is also a form of communication with others, where a jam session is akin to a group meditation that fills the room with palpable energy.

* * *

I am convinced more than ever that the goal of meditation and my musical practice are firmly aligned—both seek deeper understanding of the human experience that extends beyond the practice into life itself. In fact, Science, Arts, and various Contemplative Traditions are all ultimately about knowing, both the world and ourselves. They differ mainly in methodology and primary focus. Traditionally, Science has looked out while Contemplation has looked in. The Arts have existed at the interface, historically aligned with Religion, with recent movements to ally with Science and Technology.

Art is both a search for Truth and a means of expression through its process and practice. However, its tremendous potential as a widespread practice akin to meditation is both undervalued and underexplored in our current Western cultural context. This is partially due to the societal myth of Talent, that the arts is not worth pursuing if one is not born with the right “talent”. Furthermore, artistic Truth, due to its inherent subjectivity, is too often dismissed as frivolous, lacking the rigor of quantitative fields. This is exacerbated by the alienation felt by the general public, due to the recent over-intellectualization of certain genres of Art, that erect layers of esoteric theories as gatekeepers to widespread understanding. For much of the artistic world, an ego-centered strategy game has eclipsed the search for Truth as the ultimate goal.
We stand at the cusp of a Renaissance, just as broader conversations bridging the worlds of Art, Science and Contemplative Traditions are beginning to bloom. Many traditions have made great strides in the eternal quest for deeper understanding of the human experience. The time is ripe to connect disparate bodies of knowledge as well as methods of knowing across these diverse cultures. The reconciliation of these insights lay a fertile ground for advances in fields most closely related to the human, such as learning and health. I am thrilled that the Media Lab is taking a lead in catalyzing this convergence. I am grateful to be a part of the movement and look forward to more opportunities for contribution.
B. Invisible Cities of the Creative Mind

How does the creative mind come to be?

We all start out small, with only basic sensory-motor skills and a grand curiosity, but unable to do so much. When talking about development, certain experts like to cite 10000 hours of “hard work”. Asian parents, like mine, certainly insist on that, and a name-brand education. But many more people insist that creativity is just a matter of talent.

While hard work, education, and natural ability certainly all contribute to mental development, they abstract away all the details of how we come to be who we are. That how requires a deeper understanding of how our minds work, not just from objective, scientific explanations but from a subjective “user’s perspective”.

Three key ideas particularly influenced how I see the mind as a user. One is the ancient technique of expanding the memory by constructing vast palaces in the mind, filling them with objects, and picturing oneself walking through to recall.

Two is an idea from contemporary neuroscience, of the plastic mind as a snow-covered slope. Genetics and environment may dictate the initial terrain and weather conditions, but you shape the land by skiing down the slope every time you think or do something, leaving a groove that deepens with repetition.
Last but not least is Marvin Minsky’s theory from his pioneering thoughts on Artificial Intelligence of the human mind as a society of interconnected agents, each responsible for doing a task. Each task could in turn be broken down into a tree of sub-tasks such that every skill we acquire is not monolithic but involves mastery of many lower-level skills.

These ideas do not literally describe how our brains work on a molecular level. Rather, they are models, based in metaphor, where one thing is seen in terms of another. Metaphors are fundamental to how we understand everything we encounter. You could even say that understanding is really just assimilating new knowledge in terms of what we already know. For example, in English time is understood in the transactional framework of money we save, spend, or waste. I’d like to share with you my own metaphor of the mind, drawing from all three of the key ideas I just described.

I’d like to think of the mind as a world of interconnected, invisible cities, where ideas are buildings, connections are roads, and we could even think of agents as inhabitants. Just as we see infinite variation in cities of the world, cities of the mind vary endlessly in form, organization, and populace.
As I speak to you, I'm giving you a tour of my cities, walking through neighborhoods, pointing out details. But for you to learn something deeply, to go beyond sightseeing, you have to create a city in your own mind, building structures and roads to inhabit.

Now, when we inhabit this physical world of ours, we tend to easily fall into certain habits. A French urbanist once did a study where a girl in Paris tracked her whereabouts for an entire year, and she pretty much stayed in a triangle between her house, her school, and her piano teacher’s house\textsuperscript{165}.

Well, that could have been me in high school! Back then, I had figured out how to make straight A’s at school and win piano competitions—by being a bit of a goody two shoes and memorizing copious amounts of
Reflecting Music Through Movement

information (historic facts, math formulas, Bach fugues), using the traditional, tried and true methods of repeating a lot to myself, doing a ton of practice problems. You can call it “hard work” in the sense that it was hard... to stay focused... and not be bored... But how could I have been doing it wrong if I got A’s and won trophies!

Luckily, I discovered a better way when I learned to program the computer. You might be wondering where exactly is the connection between building code for a compiler and building cities of the mind. But I realized that beneath the surface symbols of a computer program, there are intricate structures that must exist in the mind before we build them into the computer.

I like to think of them as little factories. You start by building small factories for simple tasks—counting, drawing circles, building them piece by piece from manuals and examples. You gradually learn to build more, bigger factories, connecting them in a supply chain. In so, you learn to solve bigger problems by breaking them down.

Our factories pretty much never work the first time around. But that’s ok! You just have to go in and figure out where the problem is, a process we call debugging. It helps to have well organized facilities, but sweeping the floor alone doesn’t find the bug. You have to understand the inner workings of the program, trace how the logic flows, check how the parts...
connect. The more that you build (and debug, especially debug), the better builder you become. Pretty soon, I had a bustling city of code in my mind.

I had another city somewhere else in my mind, this one for music. On the surface, its architecture appears far more elaborate than my mini silicon valley, with opulent palaces built according to specifications of classical blueprints. But upon closer inspection, you find that the buildings lacked structure and detail. Roads too were not so well maintained, many passages blocked by the debris of time.

A sound city, as I learned from my mentor Donal Fox, is not all façade and surface brilliance. We start with a solid foundation, of rhythm. And we feel what it's like to be the buildings as we put them together piece by piece. When something doesn't quite fit, we don't just repeatedly bash it
with a blunt hammer, hoping to smooth away the mistake. We must go into the building and inspect it to determine the root cause.

As performers, we build cities to give tours, so we need to elegantly maneuver streets and structures in real-time, channeling the feeling of place to transport our audience. The more we build and maintain, the more we have ready to share, even at moment’s notice.

It might occur to you that my descriptions of programming and playing the piano sound awfully similar. Well, that’s no accident. I don’t mean to say the two are exactly the same activity. Rather I’ve discovered that the process of putting together a computer program is strikingly similar to the process of putting together a concert programme.

But I didn’t always think this way. When I first learned to play the piano, I worked quite hard, but stayed on the surface. Learning programming gave me a valuable toolset, but that knowledge stayed within its own domain. It was only when I revisited the world of music in the past few years, along with explorations in Human Computer Interaction, did I start to bridge my cities of piano and programming. Seeing connections between disparate domains has deepened my understanding of both, profoundly influencing the sorts of ideas I invent and implement.

But enough about me! Think about something you know well, a set of skills in a domain you love. These aren’t just skills in an of themselves, but an arsenal of tools and metaphors to apply to each new thing you encounter. In other words, everything you really know is a thriving city in the mind with resources to help develop new territory. And when our diverse mental cities start coming together, we start to see emerge solutions to previously unsolvable problems.
Reflecting Music Through Movement

So how does the creative mind come to be? It’s constructed! We all construct our creative minds by accumulating knowledge and experience over time, which we continually build into exquisite invisible cities.

This metaphor of the mind as invisible cities gives us a framework to reconsider how we learn and how we can reshape institutions for deeper, more creative learning. A learning rooted in our personal metaphors, that’s more than just surface correctness, with the ability to observe and act on unlikely connections between and beyond the disciplines.

At this point on my tour, I’ve led you to an open edge of my city, with structures and connections still waiting to be built. And here I will leave you with one more open question: How can we help each other build the cities of our creative minds?
C. Reflections on Marvin Minsky and Music

I only came to know Marvin through music—or through the piano, to be precise. In undergrad, I had known of Marvin, had heard of him as a larger than life figure of myth and mystery. This was around the time that his book The Emotion Machine was published. I had attended a lecture at the Media Lab as a sophomore and returned starry eyed, with an autograph.

I first really met Marvin at the 25th anniversary of the Media Lab, in my sophomore year of grad school. During an open house, he strolled right up to my demo of MirrorFugue and started playing. I would later learn from his wife Gloria that Marvin was a “piano detective”. If there were a piano anywhere in a building, he was sure to find it. Unknowingly, I had set up a sure way to make Marvin’s acquaintance.

Through my friends at the lab, I had heard of Marvin’s ability to improvise contrapuntal music but just assumed that it was the stuff of myth. It’s hard enough to improvise a single melody, much less two or more at the same time. But he really did! Right in front of me! I even got to add a note here and there.
Later, through his daughter Margaret, I had the pleasure of joining a series of musical salons at the Minsky residence. Through these salons, Marvin kindly agreed to be recorded for MirrorFugue. That was all before his health began deteriorating. Then I didn’t see too much of him for a couple of years.

In Fall of 2014, I saw Marvin again at the symposium on Beethoven’s Improvisations organized by Margaret. Not long after the event, I had heard of an AI meeting at the Minsky residence and decided to attend. Somehow, I had gotten the date wrong and showed up at the house right before a meeting of the Brookline Women’s Commission organized by Gloria. Since I had made it all the way there, I sat down at the piano and began to play a set of variations by Beethoven. At the end of the final variation, Marvin came up to the piano and doubled the main melody in a higher octave. When I finished the piece, he sat down and once again began to improvise. But he couldn’t play for long. The Brookline
Women’s Commission had assembled in the meantime and were waiting to start their meeting.

When Gloria came in to tell Marvin he had to stop playing, Marvin looked really, really sad. This was how I came up with the idea to come back every week to play the piano for and with Marvin. I would play my Beethoven or my Bach and when I ran out of pieces I knew, I’d sightread whatever was on the piano, which was more Beethoven and more Bach. Sometimes, Marvin would come and improvise something, and a couple of times near the end, we would each be on a piano and would take turns making up phrases.

Marvin and I didn’t really talk too much. When we did, it was usually about childish things—squirrels, ice cream sandwiches, wind-up houses that powered everything inside. Sometimes, I’d share some insights about music that I discovered that week, and he would smile knowingly in return. Once, I asked Marvin how he got started with the piano, and he said that he grew up with a player piano that he would “program” by punching holes at the blank beginnings of piano rolls. You could even correct your mistakes with sticky tape. It occurred to me that the piano was how Marvin got his start programming before modern computers existed. I wonder how much it’s all related—getting machines to think and feel and programming the piano to play Bach and Beethoven.

* * *

I don’t claim to speak for Marvin, not about music nor about anything else. I had intended to spend more time at the Minskies while writing this thesis to pick his brain about ideas that I had, but I guess I can’t anymore...

I’d like to think that he would have enjoyed hearing about these ideas—thoughts that certainly owe a great deal to his own thinking. Yes, Marvin was famous as the founding father of Artificial Intelligence, for inventing the confocal microscope, and for his thoughts on computation, but there was another side of Marvin, another voice in his fugue—perhaps the *cantus firmus*—that always wondered about the nature of these marvelous machines we call human beings, machines that can move, that can remember, that can feel.

I feel that Marvin’s intimate relationship with music had something to do with this line of thought. Many mathematically inclined individuals appreciate music, especially the music of Bach and especially the clever
Reflecting Music Through Movement

games he played. They would analyze the structure of canons and fugues, coming up with unifying principles connecting math and music. Sometimes they learn how to play, but their playing stays on the level of intellectual acrobatics. (Nothing wrong with that. It's a fine way to appreciate Bach, and on some level music is math). It's easy to chunk Marvin with the other mathematicians, but his musical inquiry was fundamentally different.

First of all, Marvin’s favorite composer was actually Beethoven, not Bach (though I’m sure Bach comes a close second). That says something. You cannot truly appreciate Beethoven if you cannot embrace emotion. When I was a teenager, I stayed away from Beethoven for that very reason, like many of those technically and mathematically inclined who fear or even vilify emotion. But not Marvin! He embraced emotions by really trying to understand them, and not just intellectually. In his younger days, Marvin took it upon himself to learn Beethoven’s Appassionata sonata and according to his son Henry played it quite well, with true feeling. I wasn’t there to hear it, but I believe it. Even though Marvin lost much of his technical dexterity in his old age, he still played with warmth in his touch. And you can’t play the piano with a warm touch if you don’t have warmth in your heart. I can’t ask for certain anymore, but I wonder whether Marvin went to the Appassionata as a way to understand human emotion, not just by passive listening, nor by pure intellectual analysis, but by embodying the music.

Beethoven was also known for his motivic development and his architecture. If you take apart any piece of his, any individual building block is not much more than trivial, just triads, a diminished 7th here and there. But Beethoven’s genius was in how all the small units, each in itself not so “intelligent”, came together. The 5th symphony is famously built up from just two ambiguous thirds in the “da da da dum” motif.

Also, Beethoven wasn’t known for his fugues but he wrote beautiful ones. In his youth, Beethoven was apparently criticized by his teacher Hayden for being bad at contrapuntal writing, so the fugues at the end of many of his later great works (9th symphony, the end of the Diabelli variations, the Opus 111 piano sonata) were his demonstration—his building of the tower—that he had mastered his art.

Perhaps Marvin learning to improvise fugues was his way of learning how to learn. You could “understand” a fugue through its rules. You could segment the voices, label the harmonies, and study the subject in all its
very mathematical transformations. And I’m sure Marvin did do all of that, but he also learned how to embody the processes to create fugues.

How was he able to do that? By building up smaller modules—inventing small motifs, transposing them into every key, then figuring out how to connect the smaller modules. To the musically naive, Marvin’s ability appears as the product of genius because they were not there to see all the small steps it took and how they all added up.

Marvin himself never claimed to be a genius musician. In fact, he never claimed to be any sort of genius. It was always others who ascribed that quality to him. And even though he rushed to play on any piano that he found, he never played to show off. When I asked him about improvising fugues, he quickly corrected me that what he played weren’t real fugues, only two voices, sometimes another half if he was lucky. Marvin played in an eternal quest to better understand the mysteries of the human mind—especially how we learn and how we are able to feel. These insights then fueled his mission to create machines that could learn and machines that could feel.

My relationship with music is similar. For me, learning to embody music is also a vehicle to better understand the human mind, especially learning and feeling. And ultimately, I think this was what brought Marvin and me together through the piano even though my research has little to do with
AI. But perhaps it is more related than I had imagined myself. Rather than creating machines that could learn and feel, I wish to create machines to help us learn and machines that make us feel. To attempt both requires unraveling the mysteries of the human mind and heart.
Notes


Unfortunately, it does not include the anecdote I recounted, which I had heard from Seymour Papert’s former student, Ken Kahn, creator of the visual, animated programming language ToonTalk: Ken Kahn, “ToonTalk™—An Animated Programming Environment for Children,” Journal of Visual Languages & Computing 7, no. 2 (June 1, 1996): 197–217, doi:10.1006/jvlc.1996.0011.


4 Émile Jaques-Dalcroze, Rhythm, Music, and Education (Barclay Press, 1967).

5 Zoltán Kodály, Let Us Sing Correctly (Boosey & Hawkes, 1965).

6 Carl Orff, Musik Für Kinder I (Schott, 1950).


8 Edgar Willems, L’Oreille Musicale (Éditions “Pro Musica,” 1946).


In general, historical accounts of pianists and composers discussed in this thesis are referenced from Schonberg or David S. Grover, The Piano: Its Story from Zither to Grand, 1st ed edition (London: Robert Hale Ltd, 1976),

17 This is a well-known quote that likely originated orally, as it is difficult to find an exact written source in Minsky’s writings. See Marvin Minsky, The Society of Mind, 64, 193 and 308 for the closest statements.

18 The ski slope analogy of Alvaro Pascual-Leone is found in Norman Doidge M.D and Jim Bond, The Brain That Changes Itself: Stories of Personal Triumph from the Frontiers of Brain Science, MP3 Una edition (Brilliance Audio, 2015), 209-210. The book also describes Pascual-Leone’s experiments on the effectiveness of mentally practicing the piano.


22 Werner, Effortless Mastery.

23 Minsky, The Society of Mind. 80.


Reflecting Music Through Movement


36 Yousician is a Finnish startup founded in 2010 that creates music learning apps for computers and tablets for guitar, piano, bass, and ukelele. For more information: https://get.yousician.com/.

37 JoyTunes is an Israeli startup that creates apps for computers and for tablets for learning to play the piano. For more information: https://www.joytunes.com/.

38 As of August 31, 2016, both Yousician and Piano Maestro by JoyTunes hold an average of 4.5 stars in the iTunes App Store, with 12679 and 4858 ratings, respectively.

39 One of my favorite Youtube resources with tips for classical piano playing is Paul Barton’s channel https://www.youtube.com/user/PaulBartonPiano.


Reflecting Music Through Movement

44 Analyzing the body's movements in music learning:


47 Both Lin and Liu, “An Intelligent Virtual Piano Tutor,” and Rogers et al., “P.I.A.N.O.” claim that although expression is an important part of music playing, it can and should be ignored by absolute beginners.


50 Schnell et al., “Technology and Paradigms to Support the Learning of Music Performance.”


A similar idea is explored in the popular music game Magic Piano by Smule: http://www.smule.com/listen/magic-piano/80


55 For an in depth discuss of what is play, see:


Reflecting Music Through Movement


58 McLuhan, Understanding Media.


62 Schonberg, The Great Pianists. All of the anecdotes about the history of the piano in Chapter 3 are taken from this book.


66 Improvements to the piano are described in Grover, The Piano.


67 When not specifically attributed to other sources, the history of piano music in this chapter came from Schonberg, The Great Pianists. and Grover, The Piano.


72 Gioia, The History of Jazz.


Reflecting Music Through Movement


Nguyen and Canny, “More Than Face-to-Face.”

Okada et al., “Multiparty Videoconferencing at Virtual Social Distance.”

Ishii, Kobayashi, and Grudin, “Integration of Inter-Personal Space and Shared Workspace.”


Lockhart, “A Mathematician’s Lament.”


Preston Blair, Animation: Learn How to Draw Animated Cartoons (Walter T. Foster, 1949).


Whiteside, Indispensables of Piano Playing.


McGill, Sound in Motion.


Jaques-Dalcroze, Rhythm, Music, and Education.
Marie-Laure Bachman, Dalcroze Today: An Education Through and into Music (Clarendon Press, 1991). This DVD produced by the L’Insitut Jaques-Dalcroze shows eurhythmics exercises taking place: Mireille Weber, Association suisse des professeurs de rythmique Jaques-Dalcroze, and Institut Jaques-Dalcroze (Genève), La rythmique, le solfège, un chemin vers la musique de la pratique a la connaissance, une pédagogie du mouvement corporel et musical (Gèneve: Coté face, 2010).

Alternative methods of music pedagogy bear many similarities with the Progressive education movement such as espoused by Maria Montessori and the Reggio Emilia schools.

Willems and Kodály methods emphasize singing, and Orff relates music to speech. Kodály and Orff also provide students with simple instruments to start learning to play. Suzuki emphasizes playing pieces on one’s instrument as soon as possible, such as the famous example of taping the neck of the violin.


Minsky, “Music, Mind, and Meaning.”
Bruner, Toward a Theory of Instruction. 10-12.
Papert, Mindstorms.

I discovered the similarities between recognizing musical patterns and chess patterns through a series of conversations with chess grandmaster Maurice Ashley. Maurice has written a book, which explains his ways of thinking: The Secret to Chess - How Grandmasters Find Amazing Moves and Common Sense in Chess E-Book Bundle: 2 Items, n.d.


Levitin, This Is Your Brain on Music.


Common places to hold tension are in the shoulders, neck, and jaw. In classic meditation, the goal is simply to notice, not necessarily to change what is felt. In other meditative practices, such as yoga, we could actively release tension in the body by “breathing into the spot” or by “choosing to let go”.

For a detailed discussion comparing music practice and meditation, see Appendix A.


The video clip where Tim Gallwey teaches Molly, the middle-aged overweight woman who never exercised to play tennis, is featured prominently in Kay, Doing with Images Makes Symbols.

Gallwey has an entire series of "Inner Game" books on various topics, including: Barry Green and W. Timothy Gallwey, The Inner Game of Music, 1st edition (Garden City, N.Y: Doubleday, 1986).

Betty Edwards, Drawing on the right side of the brain


But it is recommended that those who are curious go take a class!


153 Resnick, Bruckman, and Martin, “Pianos Not Stereos.”


156 See Damasio, Descartes' Error. and Varela, Thompson, and Rosch, The Embodied Mind.


158 Varela, Thompson, and Rosch, The Embodied Mind.


Reflecting Music Through Movement

162 Lakoff and Johnson, Metaphors We Live By.


Reflecting Music Through Movement

Bibliography


Reflecting Music Through Movement


Reflecting Music Through Movement


