Inspect, Embody, Invent: A Design Framework for Music Learning and Beyond

Xiao Xiao and Hiroshi Ishii
MIT Media Lab
75 Amherst St.
Cambridge, MA 02142
[x_x, ishii]@media.mit.edu

ABSTRACT
This paper introduces a new framework to guide the design of interactive music learning systems, focusing on the piano. Taking a Reflective approach, we identify the implicit assumption behind most existing systems—that learning music is learning to play correctly according to the score—and offer an alternative approach. We argue that systems should help cultivate higher levels of musicianship beyond correctness alone for students of all levels. Drawing from both pedagogical literature and the personal experience of learning to play the piano, we identify three skills central to musicianship—listening, embodied understanding, and creative imagination—which we generalize to the Inspect, Embody, Invent framework. To demonstrate how this framework translates to design, we discuss two existing interfaces from our own research—MirrorFugue and Andante—both built on a digitally controlled player piano augmented by in-situ projection. Finally, we discuss the framework's relevance toward bigger themes of embodied interactions and learning beyond the domain of music.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords
Reflective Design; Reflexive Design; Music Learning; Musical Expression; Learning Interfaces; Embodied Interaction; Piano; Auto-ethnography; Digital Arts

INTRODUCTION
Playing a musical instrument has many well-documented benefits for both cognitive function and overall health [28]. Of all the instruments, the piano is one of the most popular around the world. Unfortunately, learning to play an instrument, especially the piano, is widely regarded as difficult, tedious, and often unpleasant. Many interactive systems have been designed to support this process, mostly from a novice's perspective, framing the task as one of finding correct notes, learning correct fingering, and repeating enough to solidify muscle memory [7, 11, 17, 22, 25, 26, 33, 41, 46, 52, 57].

This paper aims to support designers to create music learning systems that foster greater expression. Prior research on musical expression within HCI have studied expert performances [12, 30], but relatively little work has been devoted to the process of acquiring expertise from an experiential perspective.

Following a methodlogy articulated by Sengers et al. as Reflective Design [50], we begin with a review of existing music learning systems to tease out their underlying assumptions. Drawing from both music pedagogy literature as well as auto-ethnography, we then present alternative approaches to learning for musicians of all levels where expression plays a key role. From these approaches, we distill the Inspect, Embody, Invent framework, the core of this paper. To concretize this framework, we discuss two of our existing projects on an acoustic player piano augmented with projection, MirrorFugue [68, 69] and Andante [70]. Finally, we discuss how our framework may generalize to domains beyond music.

This paper offers two potential contributions by connecting a well-established perspective from the art of music performance to the HCI discourse. Firstly, our framework and examples may guide the design of new music learning interfaces. Though our examples center on the piano, ideas presented are applicable across instruments. Additionally, ideas from this paper may also contribute to a deeper understanding of both embodied interaction and the learning process for creative expression. This may inform the design of systems for diverse domains beyond music.

RELATED WORK
Interactive technologies for music learning encompass a broad range of existing work from toys that convey simple musical concepts to integrated tutorial systems. To narrow our scope, our review focuses on piano learning. Learning systems are generally classified by their intended use-case, such as support for lessons (traditional and remote), tools

CHI'16, May 07 - 12, 2016, San Jose, CA, USA

ACM 978-1-4503-3362-7/16/05···$15.00
DOI: http://dx.doi.org/10.1145/2858036.2858577

for practice, and self-learning. In contrast, our review categorizes prior work across usage scenarios based on what specific roles systems have adopted to support learning. In the vast majority of existing work, systems depart little from the traditional role of the instructor, which may be divided into two main tasks: presentation of new material and feedback on student progress.

Presentation of Material

The presentation of new material by a traditional instructor typically involves some physical demonstration on the instrument. Video tutorials, as found on a number of online courses [2], seek to virtually emulate the teacher's demonstrations for distance learning. Computer generated 3D graphics have also been employed for more technical demonstrations, such as fingering of phrases [25, 33].

The teacher may also use the written score as reference, explaining its structure and pointing out important details. In a similar vein, many systems begin with information from the score, which is then translated to visuals or haptics that facilitate the mapping of score to instrument [7, 11, 13, 17, 22, 25, 46, 57]. For visuals, light-up keyboards have long been used for learning [5], and many systems have adopted the piano-roll notation to depict the score, where blocks representing notes fall onto a keyboard to indicate note strikes. This may be shown in a separate display or projected onto a physical piano [13, 26, 46]. Research on haptics have developed wearable devices that stimulate the fingers to passively learn note sequences [17].

Feedback on Progress

Based on listening and observing student playing, a traditional teacher gives feedback, which includes pointing out mistakes as well as description and demonstration of new sounds, new methods of playing, and exercises for practice. Tutorial systems have thus modeled their interactions after this process. The computer would capture some facets of the student's playing, which is then analyzed and presented back to the student [13, 15, 22, 33, 35, 36, 45, 46, 51]. Capture and analysis may focus on sound or physical movements. Systems based on sound typically capture either audio or MIDI from the student, which is then analyzed for accuracy of pitch and rhythm, sometimes articulation and dynamics to a limited degree [13, 22, 45, 46, 48, 51]. Systems based on movement have employed sensors as well as motion capture to give feedback on the lower-level mechnical aspects of playing, such as posture and muscle usage [15, 32, 33, 36].

Feedback from interactive systems may be classified into three types. On the most basic level, technology may act as a neural mirror through which students may gain awareness of their own playing. To this end, audio recording is frequently used by amateurs and professionals alike, and video has been increasingly adopted by teachers and students to inspect performances [60]. Researchers have explored the presentation of various data streams for feedback either during or after playing. For example, data from physical sensors has been sonified as real-time indicators of a student's arm movements while visual renderings of MIDI have been used to study accuracy of timing [13, 15].

As some data streams may be difficult to understand in isolation, a second type of feedback presents data from student playing in comparison with either the written score or with an expert performance. This approach allows students to inspect their own playing for deviations from the score or example. Some systems involve analysis and modeling of audio and gesture to enable score following [48]. Others present data visually, as graphs, charts, or annotated 3D models [33, 36].

Finally, feedback from a system may intervene to point out mistakes and weaknesses in a student's playing [13, 22, 46, 51, 57]. In order to identify mistakes, systems must include some notion of what is correct. To avoid the subjectivity of interpretation, existing systems tend to define correctness based on what is specified by the score, focusing on accuracy of pitch, timing, dynamics, and articulation based on a literal reading of notation.

Motivation

As practicing regularly is commonly acknowledged as challenging [41], another approach centers on motivating the student to play [11, 38, 39]. Systems have been designed to encourage practice with game-like mechanics, where students are rewarded with points for successful performances [26, 41]. To avoid the discouragement from frequent errors, projects have relaxed the criteria for what is deemed correct, increasing strictness as students progress [11, 38]. Using simular technology, interfaces have introduced an element of social playing by allowing non-musician family members to accompany their children learning the piano [39].

THE NOVICE MINDSET

Within the HCI Community, piano learning systems have predominantly focused on supporting novice classical students. From their designs we may trace a set of underlying assumptions about piano playing rooted in the priorities of the novice. We describe three key assumptions that have motivated a large number of existing projects described in the previous section.

Correctness is Paramount

Many see the primary goal of playing music as reproducing a sequence of correct notes with correct timing based on what a composer has specified in the score. For the beginner, there are two major difficulties: deciphering the score into notes on the instrument and the long repetitive process to cement muscle memory. Systems rooted in this idea attempt to present the relationship between notes and instrument in a more legible way as well as offer practice support by pointing out mistakes in the student's playing [7, 13, 22, 46, 51, 57].

Physicality is Mechanics
Playing any musical instrument requires knowing how to act through the body. The prevailing view dictates that a beginner must first acquire basic technical facility to emit sounds with the instrument before considering musicality. Finger usage, considered the most important physical component of piano playing, is the focus of several systems [17, 22, 25]. Interfaces have also been designed to train posture as well as control of various muscle groups used in playing [32, 33].

Expression is Limited
How to play the correct notes and how to use the body is usually treated as independent from expression. Based on an assumption that expression will naturally emerge from a solid technical foundation, some designers of novice learning systems forgo teaching expression altogether [17, 57]. When considered, expression is often seen as an extension of decoding the score. Thus, systems have been made to present dynamics and articulation markings as well as analyze playing for "accuracy of expression" [46, 52].

BEYOND THE NOVICE MINDSET
HCI pioneer Alan Kay once remarked that "the problem with being a beginner is that you get a lot of practice staying a beginner" [21]. The biggest preoccupation for beginners is finding and playing correct notes. Advanced players, however, focus not on the notes but on now to *speak* through them. While the beginner phase is almost universally seen as a frustrating and tedious process [7, 22, 25, 46, 57], advanced musicians derive pleasure from playing. Most interactive systems treat the beginner phase as unavoidable and focus efforts on helping students through it. We argue that designers of interactive systems should focus on fostering core musicianship skills for players of all levels. By looking beyond the novice mindset, we may bypass the usual frustrations of of the novice to find more fun and fulfillment in learning.

As background, we first discuss the aspects of playing most imporant for acquiring musicality. We then describe the concrete process to learn a classical piano piece. Finally, we demonstrate the relevance of these ideas for learners of all levels, with a note on learning outside of the Western Classical tradition.

SKILLS FOR MUSICIANSHIP
To illustrate the contrast between music learning for novices and experts, we paraphrase an analogy used by a number of respected musicians [23, 29, 66]. The novice approach of music learning is compared to classrooms that teach foreign languages through the memorization of rules, rewarding correctness over communication on exams. As a result, students may recite rules and repeat phrases but struggle to express thoughts and feelings through the language.

How people learn to speak their native language gives insights on a better way to learn music. Native language learning begins in an immersive entironment, where people first learn to listen. They gradually attune their ears to the tones and rhythms of the language and gain the ability to distinguish sounds and phrases. At the same time, people learn to create sounds with their body. Through imitation, often unconscious, they practice not only the muscles of the mouth but also the idiomatic facial expressions and gestures of the new language. Most importantly, the speaker always practices expression using the new language through constant attempts to communicate with others in the environment.

Though people learn their mother tongue as children, several courses have adapted a similar process for adults to learn foreign languages, which emphasize listening, speaking, and conversation from the outset even with a limited vocabulary [3, 60]. Though grammar rules are introduced to help the adult's analytical mind, correctness is not enforced at the expense of communication.

Ear, Body, Imagination
The above analogy highlights three aspects of language learning—listening, bodily engagement, and frequent expression—which correspond to the three aspects of music most central to an expert's musician's craft: the ear, the body, and the creative imagination. As with language learning, almost anyone can gain proficiency with playing music if they focus on developing these key skills.

Ear
Expert musicians always stress the importance of developing the ear [4, 29]. Though all students are told to "listen", most untrained ears attend to content by default, focusing on notes rather than the quality of sounds [63]. Learning to listen in the musical sense means cultivating a sensitivity for qualities such as tone, rhythm, and phrasing. True listening transforms a piece from a sequence of correct notes with correct timing into a statement in a tonal language with its own motivic vocabulary, logic, and flow.

Body
Though training the body requires much practice time for any musician, novices and experts differ drastically in their approach. Novices by default target the extremities directly responsible for producing sound [64]. On the piano this translates to a preoccupation with correct fingering. Other concerns of the body, such as the arms, wrists, and posture are later introduced and trained in isolation. The expert musician, however, feels music in the whole body. Abby Whiteside describes this phenomenon as feeling the *basic rhythm* of a piece which then coordinates the various muscle groups responsible for manipulation of the instrument [64]. Though the keyboard is the interaction locus, pianists must not think of playing as simply pressing keys but as music from the body flowing through the arms, wrists, fingers into the keys to emanate sound. As Whiteside notes, rhythm coordinates the body, but it must

be internalized as feeling rather than intellectualized counting. The student must learn to *speak* through the body, inflecting and coloring the shape and timing of phrases.

Attendance of live performance is highly encouraged among serious music students, for learning to use the body relies heavily on imitation. Studies in mirror neurons have shown that the human brain possesses specialized hardware for imitation [31], which enables people to observe and copy complex, coordinated movements difficult to explain analytically. Through imitation, students may absorb not only technique but also how music is felt in the body of another player.

Imagination

For expressive playing, a student must not only develop the ear and the body but also the imagination. In musical parlance, this is also referred to as "developing the ear" [4, 63], showing the close link between perception and imagination. To avoid the confusion of terms, we will distinguish the imagination as the *inner ear*.

Training the inner ear means growing the ability to hear music without the instantiation of physical sounds. Composers hone this skill for their craft, exemplified by Beethoven's ability to compose even while deaf. Expert performers must also develop their inner ear for expressive playing. In *The Art of Piano Playing,* Heinrich Neuhaus instructs classical pianists to develop a crystal-clear *artistic image* of how a piece ought to sound with full expression, which then guides the body's movement to physically render the piece [34].

Even though classical performers no longer improvise original material[1], practicing written pieces still requires an active imagination. To discover the artistic image of a new piece, the expert practices by replaying passages in different ways (varying speed, dynamics, phrasing, etc.) while carefully listening to craft an original interpretation. Studies have shown that creative variation during practice boosts both their effectiveness and enjoyment [23].

PROCESS TO LEARN A PIECE

To understand how these skills translate to the practical task of learning pieces, we draw from both pedagotigal literature and personal experience. In the classical tradition, any musician—from the novice to the virtuoso—always learns from the score written by a composer. While the amateur translates instructions from the score directly into motions on the instrument, the expert first converts the score into an *artistic image* (or *aural image*) in the mind, which then guides the body's movement. When training the body, practice methods should reinforce performance habits [23]. Students must always aim to play with ease and expression, never slipping into auto-pilot mode where the mind wanders elsewhere. In practice as in performance, students should listen carefully and make physical adjustments based on what they hear.

First Person Perspective

For a glimpse into the process of learning a classical piece, we turn to the first-person perspective of the first author. Employed by David Sudnow in the now classic *Ways of the Hand* [55], first-person methodologies are gaining increasing traction within the field of HCI. In particular, first-person approaches have been applied to themes of embodied cognition and aesthetic experience to derive insights often overlooked by traditional third-person research practices [16, 43, 46].

The first-author studied piano for 18 years in the typical classical approach before taking an interest in alternative methods of learning music. For the past six years, she has studied intensively with a world class pianist and composer. She has documented her progress in journal entries, which include take-aways from weekly lessons and notes from daily practice. To reflect her personal perspective, the following section is presented in the first person voice.

Reference, Reduction, Reconstruction

When I begin to learn a new piece, the first step is always to familiarize myself through *reference* material. Generally, I try to listen to recordings from different artists. It also helps to watch videos of performances. Through these materials, I begin to form a high-level impression of how the piece might sound and feel. To get into the sound, I sometimes sing along with recordings, move my hands as if conducting, or even dance to the music.

Training myself to play the piece involves two main processes, which I will call *reduction* and *reconstruction. Reduction* means creating a sketch. I would first map out its structure, delineating sections and seeing their logical flow. In each section, I would find and play the melody by itself, which also helps with finding its internal structure and shape. In this phase, it is useful to understand some theory, which helps with identifying structure. However, it's important not to get caught up in intellectualizing for its own sake at the expense of playing musically. For the performer, theory is just a tool to help us with greater expression.

The *reduction* phase also involves pinpointing the difficulties of the piece. These could be due to complexity in content (e.g. unusual harmonies, dense counterpoint), in physical techniques that need more training (e.g. fast passages, octaves), or even in expressive interpretation (e.g. rapid changes in dyanmics). A good goal for the *reduction* stage is to be able to play every section of the piece only as melody, with full dynamics (and perhaps a simple harmony if you are able). In a way, it's like creating your own jazz lead sheet.

[1] It is worth noting that the current predominant culture of classical performance has not always been the norm throughout history. Until well into the 19th century, improvisation played an integral role in what we now know as "classical" music, especially for the piano [49].

After an initial sketch has been made, then comes the *reconstruction* phase, where I gradually begin to fill it in. There are a few strategies I typically use. For musically complex spots, I start with an outline, then gradually add more details as each version becomes fluent. Starting with the melody, I first add a bassline, then more of the harmony. If the material is too complex to play well, I simplify it. Sometimes, I ignore rhythm in the accompaniment and play it only as block chords. Singing along with the melody also helps to maintain musicality.

For technical difficulties, the trick is to invent exercises for yourself in order to keep your mind engaged. An easy way to do this is through octave displacement, playing a passage at different parts of the piano. I can also change the rhythm of the passage, or break it apart into smaller units and play the units in a different order. In the end, it doesn't matter too much what type of exercises you invent, as long as you maintain a flow state, where you are neither frustrated nor bored [6].

Finally, I can save a lot of practice time by grouping similar passages. Many pieces are built up from motifs, which are then repeated and varied. Recognizing them means what I have already practiced may be applied to speed up learning new material. This strategy also works from piece to piece. Let's say one piece uses a lot of octaves. If we encounter another piece that also has octaves, we should use our existing knowledge of the technique to help us learn the new piece faster. It's really just a switch in mindset and may seem simple, but even in practice, I often have trouble thinking in this way because of my own long training in the typical brute-force approach.

Though I segment practice into three distinct stages, the reality is never such a strict forward progression. It's very helpful to keep *referencing* recordings and videos (and to attend performances) even in more advanced stages of learning a piece. It's also helpful to go back to *reduction* every so often to not lose track of the underlying structure.

LEARNING FOR ALL LEVELS

Focus on listening, embodied understanding, and creative imagination is not only reserved for the already advanced. Several established methods for beginners focus on developing exactly this set of skills. For example, the Suzuki method advocates learning by ear [56], and Dalcroze Eurhythmics presents music as full-body exercises that connect physical movements with sound [20]. Dalcroze practitioner Claire-Lise Dutoit summarizes the principles of eurhythmics as cultivating "the ability to hear", to "understand and express music in movement", and "the call made on the pupil to improvise and develop freely his own ideas" [9]. These guiding principles also underlie several other teaching systems such as the Kodály, Willems, and Orff methods [53, 37, 1]. In all these approaches, symbolic notation is introduced only after the feeling of music is firmly grasped by the ear and the body. While these methods were originally conceived for children, they have also been effectively adapted for adult beginners.

Role of the Teacher

Teaching music for all levels should share the same priorities; instruction for the less advanced only differs in the level of guidance from the teacher. A good teacher may use demonstrations to tune the student's ear to different qualities of sound and to show how sounds manifest in bodily movements. The teacher may also break down a complex piece into more manageable parts, distill a knotty passage into its essence, and introduce intermediary exercises to guide a student through difficulties. Ultimately a good teacher also instills self-sufficiency—teaching students how to teach themselves.

Music Across Cultures

With few exceptions, discussion of music within HCI has primarily centered on the Western Classical tradition. While our ideas of piano learning readily apply across instruments in Western Classical music, it is important also to look beyond our own cultural defaults [50]. The perspective espoused by this paper may go against the "typical" view of music learning within HCI, but the same ideas are the accepted norm in many cultures around the world.

Historian Craig Wright observes that notation made the West an "odd culture out", for its adoption of a symbolic and quantitative representation. Other cultures described music differently in writing if they did so at all [67]. The Japanese, for example, record gestures instead of specific notes in *shakuhachi* scores. In cultures without a precise quantitative notation, primacy of the ear, the body, and the imagination take on greater importance during learning. Often, as in the case of Brazilian Samba Schools, learning to play is a social activity where imitation and interpersonal communication play a crucial role [40]. Samba derives in part from African drumming traditions, where a robust rhythm felt in the body animates performances coordinates ensemble playing. African traditions played a key role in the birth of American Jazz, which has in turn influenced much of popular music. In these traditions, development of the ear, an embodied understanding, and the creative imagination pervade all stages of the learning process [14, 19].

FROM DOMAIN KNOWLEDGE TO DESIGN

The previous sections offered a condensed overview of the knowledge and experiences that have shaped our thinking. From this background, we now turn to consider the design of interactive systems, introducing a design framework and discussing two example projects. The framework distills our desired learning process into three key components—Inspect, Embody, and Invent. We summarize each and put forth general guidelines for design. Two projects from our own prior research show how actual systems may support the desired learning process.

The same background has shaped both the framework and the two projects, though the framework had not yet been formalized when the projects were designed and deployed. To correlate the framework with concrete designs, we reinterpret interactions on the two projects in light of Inspect, Embody, Invent. Though it may not qualify as a formal evaluation of the framework, this exercise enriches the connection between domain knowedge, design theory, and working prototypes.

THE FRAMEWORK: INSPECT, EMBODY, INVENT

To open opportunities for the design of novel music learning systems, we summarize the main takeaways of the previous sections, which we term Inspect, Embody, and Invent.

Inspect

refers to the necessity to convey an understanding of repertoir beyond surface correctness:

- Learning should tune the ear to the quality of sounds, such as tone, phrasing, and rhythm.
- Students should learn to recognize underlying structure beyond individual notes (e.g. broad harmonic shifts, cadences, and motifs).
- It's important to present musical structures and qualities not just as intellectualized symbols. The eye should support the ear, not take its place.

Embody

refers to the crucial role played by the body when learning to play music:

- Music should be felt through the whole body, with a basic rhythm corrdinating the body's movements.
- Imitation is an effective way to learn to use the body for both technique and expression.
- During practice, students should always feel the music so that it flows with ease and expression, even if they have to simplify a piece to its essence.

Invent

refers to the development of the creative imagination.

- Students should not only train the ear for physical sounds but should also learn to "hear" music with the *inner ear*.
- It is helpful to invent exercises for difficult passages to decrease frustration and to engage the mind for more effective practice.
- Like learning to speak a language, the ultimate goal of playing music is to *speak* through the music.

EXAMPLE SYSTEMS

To bridge theory with actual designs of interactive systems, we describe two existing projects from our own research. Both projects augment a Yamaha Disklavier player piano with projection [71]. The first, MirrorFugue, is inspired by reflections on the lacquered surface of a grand piano and simulates the presence of a virtual pianist playing the physically moving keys [68, 69]. The second, Andante, visually echoes the bodily sensation of musical phrases through the silhouettes of miniature figures walking and dancing on the keys [70]. Both are already documented in their respective publications. We describe here only their technical implementation and musical content in order to ground the discussion of interactions in light of Inspect, Embody, Invent.

Figure 1. Learning through imitating a virtual performer on MirrorFugue

MirrorFugue

Setup

To create the illusion of a virtual reflection playing the piano, 720p video of a pianist's hands and upper body is 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 ¼" plywood treated with projection paint is placed on the music stand. Both videos are beamed from a a short throw projector mounted above the piano bench 7' from the ground. This setup ensures that a person seated normally at the piano does not occlude the projected images.

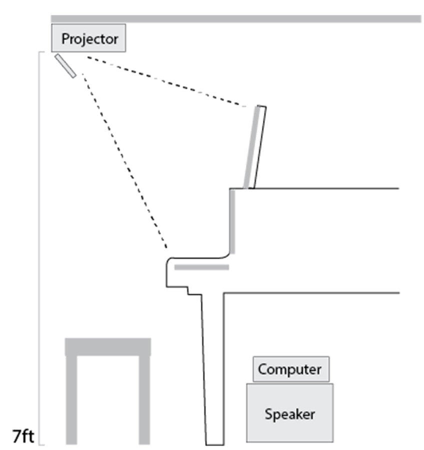

Figure 2. System diagram showing projection setup for both MirrorFugue and Andante

Data for a MirrorFugue performance comprises one MIDI sequence and two videos, one to capture the piano keyboard and the other to capture the pianist's head and shoulders. 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. Video and MIDI data are then manually edited using their accompanying audio for synchronization. The causal link between the life-sized projected video, the physical movement of the keys, and the acoustic sound from the piano creates a striking illustion of the pianist's physical presence. The effect is especially vivid for someone seated at the piano bench.

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. We are currently developing a new version of the program 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 [65].

Content

Performances from 15 pianists were recorded for MirrorFugue. This roster includes 8 professionals, among them three Steinway Artists and several well-known names (e.g. Allen Toussaint, Vijay Iyer, and Ryuichi Sakamoto) [54]. 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 the first-author herself. These recordings showcase a range of musical styles and levels of expertise.

Many aspects of playing the piano may be absorbed by listening and watching a MirrorFugue performance more effectively than watching a video on a screen [68, 69]. From the hands on the keyboard, a student may observe fingering as well as shifts in hand position and weight balance. From the upper body, a student may understand the relationship between breath and phrase, how rhythm is felt, and how a performer may personify characters and moods. Additionally, a student may pick up elements of personal style, delving into multiple interpretations of the same piece or emulating the technique of artists whose hand and body type is most similar to theirs.

Andante

Setup

Andante retains the same projector setup as MirrorFugue. To display on the fallboard, the reflective keyboard cover is replaced with a projection surface made form the same material as the MirrorFugue upper body display.

For the original implementation of Andante, all animations were drawn by hand using a lightboard for precise control of character's movements. Frame sequences were organzed based on type of step (e.g. whole step between white keys). A Java program controls playback by reading MIDI recorded from a human player and selecting the appropriate frame sequences to display for each note played by the Disklavier. We are currently working on a new version of Andante in JavaScript that procedurally generates character animations with lifelike movement from human input of musical phrases.

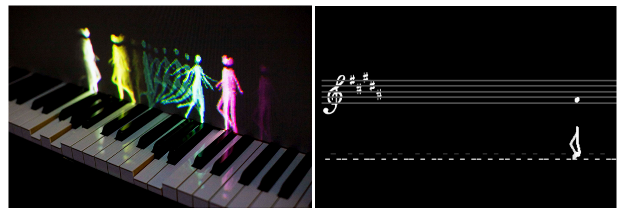

Figure 3. (left) Rendering polyphonic music on Andante. (right) A screenshot from the new JavaScript program that shows symbolic notation with Andante.

Content

To show a range of content for Andante, the original implementation featured three example applications: scales played by a variety of characters, a blues baseline to accompany improvisers, and a Bach canon that maps polyphony into a different figure for each voice. Unlike MirrorFugue, which seeks to reflect reality, Andante abstracts and miniaturizes the figure, which personifies the shape and structure of musical lines. Since people easily emphathize with characters in animated films from their expressive form and movement [42], Andante's figures invite the user to project themselves into a virtual microcosm.

The cartoon nature of the figures makes Andante especially well-suited for teaching children. The characters' appearance and gait echoes and amplifies the shape of musical phrases and may help attune the ear to qualities of sound. We may also use Andante to introduce symbolic notation, taking advantage of the resemblance between a figure's head and how notes are written. The head may be shown with variations in stem and filling to indicate the type of note, and it may be raised or lowered on a staff to reflect what keys are played by the figure. This maintains a connection to a body-based intuition of music when introducing symbolic notation.

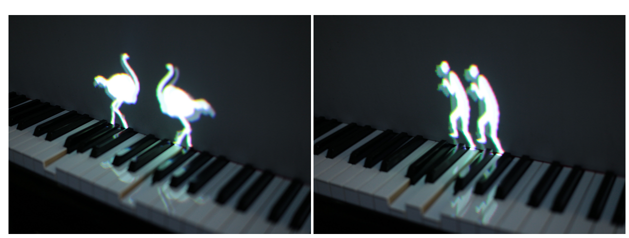

Figure 4. Different characters on Andante

INTERACTIONS

MirrorFugue has been shown to 12 concert-level pianists and 2 professional teachers. In-depth interviews were conducted with 5 of these expert musicians while quotes and general impressions were recorded from the rest. We also observed interactions of MirrorFugue “in the wild” at a public exhibition in New Orleans over two week in December 2015. The exhibit featured recordings from 4 local pianists and allowed visitors to sit at the piano and interact with the instrument.

Andante has been shown to 5 professional performers, with 3 of whom we conducted in-depth interviews. Andante was also integrated into lesson of 30-50 minutes with 8 children between the ages of 7 and 12, following 15-20 hours of design discussions with a Dalcroze certified piano teacher over the course of 9 months.

Based on comments from the expert interviews as well as observations from the MirrorFugue exhibit and Andante lessons, we derived three main ways to engage with both projects: listening, imitating, and playing a duet. These are not rigid modes of interaction, but instead allow a user to fluidly go from one mode to another. These interactions may be linked to the Inspect, Embody, and Invent stages of learning, which we describe in more detail along with illustrative scenarios.

Figure 5. (left) A child improvises with her own reflection. (right) A visitor imitates blues pianist Jon Cleary.

Inspect

Most basically, a user may sit at the piano to review recordings on both MirrorFugue and Andante. All the pianists who saw MirrorFugue expressed that it was an excellent tool to understand fingering as well as musical expression. At the exhibit, we observed that children would rest their hands on the moving keys as they listened to the performance. After the Andante lessons, the teacher observed that his students listened more carefully to pieces play with Andante and retained a better memory of what they had heard.

We hypothesize that both MirrorFugue and Andante may help students develop their ear as visual elements may draw attention to nuances in quality of sound. Both systems also relate the perception of sound to other dimensions of musical understanding. MirrorFugue presents sound in relation to actions of the performer while Andante brings forth the symbolic, structural aspects of a piece. Additionally, students may also capture their own playing on the piano and review it on either system. On MirrorFugue, students may become more aware of their own body during playing, observing technique, posture, as well as expressive gesture. A recording mapped to Andante figures may help focus attention to the shape of phrases and the feeling of rhythms. Both systems may teach students to assess the quality of their own playing.

Embody

MirrorFugue and Andante both advocate learning through imitation. Both support playback at slower speeds as sound is generated on the piano from MIDI instructions. Through copying the physical movement of playing, expression is absorbed as an integral part of the overall musical statement.

A striking example of this interaction took place at the MirrorFugue exhibit, where a fan of one of the recorded pianists sat down in front of the piano and began to imitate her virtual idol. Though she has no training in piano, she was able to mirror almost exactly the overall movements of the virtual body and hands. This demonstrates how MirrorFugue may allow students to grasp the larger gestures of a piece before preoccupation with specific notes.

During the Andante lessons, the difference between quarter notes and eighth notes were introduced as “walk” vs “run”, which allowed children to grasp patterns more quickly than from the score alone. Andante may help students internalize rhythms before learning specific notes on the piano. This may speed up learning by breaking down the process, decreasing cognitive load at each step.

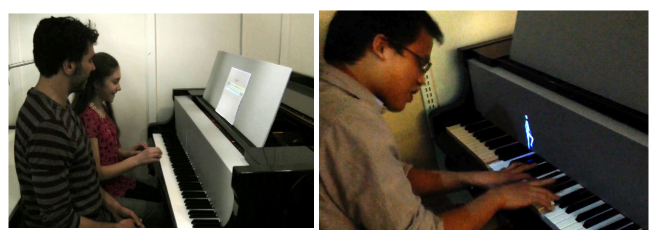

Figure 6. (left) A lesson using Andante. (right) Playing the blues with a virtual figure walking a bass line.

Invent

More advanced pianists who sat at the bench of MirrorFugue and Andante almost always began to play with the recordings without prompt. They were often highly motivated by the experience of playing a duet with famous virtual pianists. We also observed younger users, in this case children below the age of 10, who couldn’t help “accompanying” the pianist in MirrorFugue by playing along.

These interactions demonstrate how students may play virtual duets with Andante and MirrorFugue as a part of their training. This may be useful when the student needs to repeat a phrase several times to secure it in the ear and the body. In such cases, MirrorFugue may provide a background that varies as the student repeats, to help focus the student’s ear. As the student becomes more fluent with the material, MirrorFugue and Andante may provide a base against which the student may improvise with the material. For example, the system might provide a harmonic progression, allowing room for the student to invent melodies.

BEYOND MUSIC

To close, we reflect on how insights from the music learning process may inform the understanding of learning and embodiment in general. We first connect ideas presented in this paper to existing theories from the congnitive sciences. We then consider the broader relevance of Inspect, Embody, Invent by applying it to diverse domains beyond music.

Connection to Cognition

The prevailing view of music learning for novices and alternative approaches advocated by this paper echo contrasting views of human cognition. Descended from the likes of Plato and Descartes, the dominant view in Western philosophy has long held that humans live in an objective, external world. Thus, cognition is the process of forming and manipulating mental representations of that world [62]. Following this tradition, Cognitivists (as termed by Varela et al.), particularly in the field of Artificial Intelligence, liken the human mind to a computer, where representations are abstract symbols, which the brain manipulates independent of the body, the physical world, and any deeper meaning.

Underlying assumptions of typical classical piano learning align closely with a Cognitivist worldview. Playing music is seen as the process of programming the body to perform instructions encoded in abstract symbols to produce sound. Practice requires repetition only because human muscles (unlike machines) must repeat to remember. After instructions are programmed into muscle memory, the player then adds expression, like applying formatting to a text document.

In contrast, the approach to learning music summarized by our framework correspond to ideas of embodied cognition, a topic of growing interest among HCI researchers in recent years [8, 16, 18, 27, 43, 48, 58]. Embodied cognition argues that the body's ability to perceive and act within an environment plays a crucial role in how humans understand the world [19, 62]. Learning to play music can thus be seen as the process to acquire the perception (listening) and motor skills (technique) to act within the environment of the instrument.

Compelling arguments across disciplines place embodiment at the root of human cognition. In his work with children, Piaget described a progression from pure sensory-motor experience to symbolic logical thinking in human cognitive development [21]. Within linguistics, Lakoff and Johnson argue that abstract concepts are understood in terms of the more concrete, ultimately scaffolding upon basic sensory-motor experience [24]. The effectiveness of music learning methods like Dalcroze Eurythmics [44], which begin by cultivating a body-based musical intuition, also support theories of embodied cognition.

Generalizing the Framework

The Inspect, Embody, Invent framework is relevant for diverse skills that include a physical and creative component. Examples include dance, sports, martial arts, cooking, crafts, as well as plastic arts such as drawing and sculpture.

In the generalized framework, **Inspect** calls for systems to support the development of the senses. For example, learning to draw requires learning to see [10], dance requires enhanced kinesthesia, and cooking requires a heightened sensitivity to taste. To support the development of perception, reference materials as well as feedback methods provided by a system should emphasize the qualitative and the sensorial more than quantitative, intellectualized representations.

Embody calls for systems to support and enhance the role of the body during learning. Video is already a popular medium for people around to world to learn new skills, demonstrating the efficacy of imitation. Novel interfaces might consider how to go beyond the simple video in supporting imitation. Systems might also explore strategies to coordinate movements of the entire body.

Finally, **Invent** calls for systems to encourage and support users to creatively engage with what they have learned. This includes not only making personal statements through the medium (e.g. inventing new dishes or dance moves) but also creative experimentation during practice to acquire basic skills. Systems may provide examples to inspire the user as well as create environments to motivate creative output.

Though Inspect, Embody, Invent is most immediately applicable for domains involving physical skill, it may also have broader relevance for more abstract subjects. A compelling example is Seymour Papert's work with the LOGO programming language to teach children mathematics [40]. Papert introduced math concepts through a programmable *turtle* with both a virtual and a physical manifestation. Programming LOGO enabled children to solve problems not by blind symbol manipulation but by projecting themselves into the perspective of the *turtle* (Embody). Papert also observed that effective learning environments invite participants in with sensorially rich discoveries (Inspect) and foster learning through varied enactments that allow room for expression (Invent).

CONCLUSION

Our ultimate goal in writing this paper is to help designers create more effective and enjoyable music learning systems that promote a deeper understanding of musicality. To this end, we first uncovered the assumptions within the prevailing view of music learning in the Western classical tradition, tracing how these assumptions have shaped the design of existing music learning systems. Next, we presented an alternative view of music learning drawing from domain literature and first person experience. Finally, we distilled key points of the alternative view into a new framework—Inspect, Embody, Invent—and used it to reinterpret two existing projects on the piano. This paper also tried to connect the dots between the process of music learning to broader domains for design, relating ideas from music learning to existing theories of embodiment and cognition and applying our framework to the learning of other skills beyond music.

We recognize that our general approach may be seen as unorthodox against the de facto rules of HCI research. Thus, we included references to methodologies and precedents (e.g. Reflexive design and auto-ethnography) as necessary to justify our course of research. Conversely, we observe that traditional HCI methodologies with their focus on optimizing quantifiable metrics risk blinding researchers to the richness and nuance of artistic practices. We hope that this work may contribute to a broader discussion on appropriate research methods to design for the arts on their own terms.

CODA

The field of HCI is always seeking new metaphors and frameworks to guide the design of new interactions with the computer. For us, music has been a rich source of inspiration, with the piano an especially interesting example. Unlike other musical instruments, which feature a direct connection between the human body and the sound-producing object, the piano keyboard is an artificial interface used to operate a back-end machine. It thus has striking parallels with the early form-factors of the computer—a machine accessed through commands on the keyboard. (Interestingly, another name for the first typewriters was "literary piano" [61]).

The piano keyboard interface allows a single person to play more complex music than on any other instrument. Consequently, learning to play all too often prioritizes a symbolic understanding over embodiment and emotional expression, drawing an interesting parallel with the many problems of using the traditional computer. However, experts on the piano do learn to play the machine with the command of the entire body, authentic emotional expression, and a fluent grasp of sophisticated musical structures. We believe that deeper understanding of all these facets of of musical performance may yield valuable insights for designing more engaging, more *human* interactions with the computer.

This paper has attempted to shed some light on these topics, providing a primer for deeper music learning. There is still much unexplored territory, and we have offered pointers to some promising areas of future research for musical applications and beyond. As Alan Kay called for all aspiring technologists to delve into the liberal arts for an "endless well of inspiration" [21], we encourage researchers to delve more deeply into the endlessly fascinating world of music.

ACKNOWLEDGEMENTS

This paper was motivated by learnings from several years of piano lessons with Donal Fox. Conversations with Pablo Puente, Hanoi Hantrakul, Edith Ackerman, and Ken Perlin also contributed much to our thinking. We are grateful to Yamaha for their generous loan of a Disklavier piano over the past 4 years and to the Tangible Media Group for continual support. Thanks to the anonymous reviewers for their constructive feedback.

REFERENCES

1. American Off-Schulwerk Association. What is Orff Schulwerk? http://aosa.org/about/what- is- orff-schulwerk/. Accessed: 2015–09–25.
2. Berklee. Berklee Online. https: //online.berklee.edu/school/courses/music_education. Accessed: 2015–08–25.
3. Berlitz. The Berlitz Method. http://www.berlitz.us/berlitz- method/. Accessed: 2015–09–25.
4. Ran Blake. 2010. Primacy of the Ear. Lulu.com.
5. Casio. Lighted Keys. http://www.casio-usa.com/ products/Digital_Pianos_%26_Keyboards/Lighted_Keys/. Accessed: 2015–09–20.
6. Mihaly Csikszentmihalyi. 1990. Flow: the Psychology of Optimal Experience. Harper Perennial.
7. Roger B. Dannenberg, M. Sanchez, A. Joseph, P. Capell, R Joseph, and R Saul. 1990. A Computer-Based Multi-Media Tutor for Beginning Piano Students. Interface - Journal of New Music Research 19 (1990), 155–173. Issue 2-3.
8. Paul Dourish. 2004. Where the Action: The Foundations of Embodied Interaction. The MIT Press.
9. Claire-Lise Dutoit. 1965. Music Movement Therapy. The Riverside Press.
10. Betty Edwards. 1989. Drawing on the Right Side of the Brain. Tarcher.
11. Y. Fukuya, Takegawa Y., and H. Yanagi. 2013. A Piano Learning Support System Considering Motivation. (2013).
12. Rolf Inge Godoy and Marc Leman (Eds.). 2010. Musical Gestures: Sound, Movement, and Meaning. Routeledge.
13. W. Goebl and G. Widmer. 2006. Unobtrusive Practice Tools for Pianists. In Proceedings of the 9th International Conference on Music Perception and Cognition.

14. Lucy Green. 2002. How Popular Musicians Learn: A Way Ahead for Music Education. Ashgate Publishing, Ltd.
15. A. Hadjakos, E. Aitenbichler, and M. Mühlhäuser. 2008. The Elbow Piano: Signification of Piano Playing Movements. In Proceedings of NIME ’08.
16. Kristina Höök. 2010. Transferring Qualities from Horseback Riding to Design. In Proc. NordiCHI ’10.
17. Kevin Huang, Ellen Yi-Luen Do, and Thad Starner. 2008. PianoTouch: A Wearable Haptic Piano Instruction System for Passive Learning of Piano Skills. In Proc. ISWC ’08. IEEE Computer Society, Washington, DC, USA, 41–44.
18. Caroline Hummels, Kees C. Overbeeke, and Sietske Klooster. 2007. Move to Get Moved: A Search for Methods, Tools and Knowledge to Design for Expressive and Rich Movement-based Interaction. Personal Ubiquitous Comput. 11, 8 (Dec. 2007).
19. Vijay S. Iyer. 1998. Microstructures of Feel, Macrostructures of Sound: Embodied Cognition in West African and Africa-American Musics. Ph.D. Dissertation. UC Berkeley.
20. Emile Jacques-Dalcroze. 1967. Rhythm, Music, and Education. The Riverside Press.
21. Alan Kay. Doing with Images Makes Symbols. https://archive.org/details/AlanKeyD1987. Accessed: 2015–09–25.
22. T. Kitamura and M. Miura. 2006. Constructing a Support System for Self-Learning Playing the Piano at the Beginning Stage. In Proceedings of the 9th International Conference on Music Perception and Cognition.
23. Gerald Klickstein. 2009. The Musician’s Way: A Guide to Practice, Performance, and Wellness. Oxford University Press.
24. George Lakoff and Mark Johnson. 2003. Metaphors We Live By. University of Chicago Press.
25. Chih-Chun Lin and Damon Shing-Min Liu. 2006. An Intelligent Virtual Piano Tutor. In Proc. VRCIA ’06.
26. Synthesia LLC. Synthesia. http://www.synthesiagame.com/. Accessed: 2015–09–20.
27. Lian Loke and Toni Robertson. 2013. Moving and Making Strange: An Embodied Approach to Movement-based Interaction Design. ACM Trans. Comput.-Hum. Interact. 20, 1 (April 2013).
28. Michael Matthews. 18 Benefits of Playing a Musical Instrument. http: //www.effectivemusicteaching.com/articles/directors/ 18- benefits- of- playing- a- musical- instrument/. Accessed: 2015–09–25.
29. David McGill. 2007. Sound in Motion: A Performer’s Guide to Greater Musical Expression. Indiana University Press.
30. Andrew P. McPherson, Adrian Gierakowski, and Adam M. Stark. 2013. The Space Between the Notes: Adding Expressive Pitch Control to the Piano Keyboard. In Proc. CHI ’13 (CHI ’13).
31. P. Molenberghs, R. Cunnington, and J. Mattingley. Is the Mirror Neuron System Involved in Imitation? A short review and meta-analysis. Neuroscience and Biobehavioral Reviews 33, 975–980. Issue 7.
32. R. Montes, M. Bedmar, and M. Sol Martin. 1993. EMG Biofeedback of the Abductor Pollicis Brevis in Piano Performance. Biofeedback and Self-Regulation 18, 67–77. Issue 2.
33. J. Mora, W.-S. Lee, G. Comeau, S. Shiromohammadi, and A. El Saddik. 2006. Assisted Piano Pedagogy Through 3D Visualization of Piano Playing. (2006).
34. Heinrich Neuhaus. 1995. The Art of Piano Playing. Hollowbrook Pub.
35. K. Ng and P. Nesi. 2008. i-Maestro: Technology-Enhanced Learning and Teaching for Music. In Proceedings of NIME ’08.
36. Kia C. Ng, Tillman Weyde, Oliver Larkin, Kerstin Neubarth, Thijs Koerselman, and Bee Ong. 2007. 3D Augmented Mirror: A Multimodal Interface for String Instrument Learning and Teaching with Gesture Support. In Proceedings of the 9th International Conference on Multimodal Interfaces (ICMI ’07). ACM, New York, NY, USA, 339–345.
37. The Swiss Website of Willems Music Education Teachers. Willems Music Teaching System. http://www.willems.ch/en/education.htm. Accessed: 2015–09–25.
38. C. Oshima, Y. Miyazawa, and K. Nishimoto. 2002. Color-in Piano: A Piano that Allows a Performer to Concentrate on Musical Expression. In Proceedings of the 7th International Conference on Music Perception and Cognition.
39. Chika Oshima, Kazushi Nishimoto, and Norihiro Hagita. 2007. A Piano Duo Support System for Parents to Lead Children to Practice Musical Performances. ACM Trans. Multimedia Comput. Commun. Appl. 3, 2, Article 9 (May 2007).
40. Seymour Papert. 1993. Windstorms: Children, Computers, and Powerful Ideas. Basic Books.
41. Graham Percival, Ye Wang, and George Tzanetakis. 2007. Effective Use of Multimedia for Computer-assisted Musical Instrument Tutoring. In Proc. EMME ’07.

42. Ken Perlin and Athomas Goldberg. 1996. Improv: A System for Scripting Interactive Actors in Virtual Worlds. Computer Graphics 29, 3 (1996).
43. Sebastiaan Pijnappel and Florian Mueller. 2013. 4 Design Themes for Skateboarding. In Proc. CHI '13.
44. P. Puentes and P. Godall. 2012. Evaluation of the Dalcroze Rhythmic Approach in First Steos of Music Education. Master project research. Barcelona: UAB.
45. Kathleen Riley. 2007. New Views of Piano Playing Through Motion Analysis and MIDI Technology. Journal of Technology in Music Learning 4 (2007), 17–30. Issue 1.
46. Katja Rogers, Amrei Röhlig, Matthias Weing, Jan Gugenheimer, Bastian Könings, Melina Klepsch, Florian Schaub, Enrico Rukzio, Tina Seufert, and Michael Weber. 2014. P.I.A.N.O.: Faster Piano Learning with Interactive Projection. In Proc. ITS '14 (ITS '14). ACM, New York, NY, USA, 149–158.
47. Thecla Schiphorst. 2011. Self-evidence: Applying Somatic Connoisseurship to Experience Design. In CHI '11 Extended Abstracts.
48. N. Schnell, F. Bevilacqua, D. Schwarz, N. Rasamimanana, and F. Guédy. 2006. Technology and Paradigms to Support the Learning of Music Performance. (2006).
49. Harold C. Schonberg. 1987. The Great Pianists: From Mozart to the Present. Simon and Schuster.
50. Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph 'Jofish' Kaye. 2005. Reflective Design. In Proc. Critical Computing '05 (CC '05). ACM, New York, NY, USA, 49–58.
51. S. Shiromohammadi, A. Khanafar, and G. Comeau. 2006. Midiator A Tool for Analyzing Students' Piano Performance. Revue de recherche en Education Musicale 24 (2006).
52. Stephen W. Smoliar, John A. Waterworth, and Peter R. Kellock. 1995. pianoFORTE: A System for Piano Education Beyond Notation Literacy. In Proc. ACM Multimedia '95.
53. International Kodály Society. The Kodály Concept. http://www.iks.hu/index.php/home1/the- kodaly-concept. Accessed: 2015–09–25.
54. Steinway & Sons. Steinway Artists. http://eu.steinway.com/en/concert-artist/artists-alphabet/. Accessed: 2016–01–07.
55. David Sudnow. 1993. Ways of the Hand. The MIT Press.
56. Shinichi Suzuki. 1974. The Suzuki Concept: An Introduction to a Successful Method for Early Music Education. Diablo Press.
57. Yoshinari Takegawa, Tsutomu Terada, and Masahiko Tsukamoto. 2012. A Piano Learning Support System Considering Rhythm. In Proc. ICMC '12.
58. Jakob Tholander and Carolina Johansson. 2010. Design Qualities for Whole Body Interaction: Learning from Golf, Skateboarding and BodyBugging. In Proc. NordiCHI '10.
59. Michel Thomas. The Michel Thomas Method. http://www.michelthomas.com/. Accessed: 2015–09–25.
60. Mary Toy. 2003. The Video Camera: A Teacher's Best Friend. The Piano Adventures Teacher (2003).
61. *Type Writing Machine*. Scientific American. Vol. 1017. Issue 1.
62. F.J. Varela, E. T. Thompson, and E. Rosch. 1992. The Embodied Mind: Cognitive Science and Human Experience. The MIT Press.
63. Kenny Werner. 1996. Effortless Mastery: Liberating the Master Musician Within. James Aeberold.
64. Abby Whiteside. 1955. Indispensables of Piano Playing. Charles Scribner's Sons.
65. Chris Wilson, Google, and Jussi Kalliokoski. Web MIDI API. https://webaudio.github.io/web- midi- api/. Accessed: 2015–06–15.
66. Victor Wooten. Music As a Language. http://http://ed.ted.com/lessons/victor- wooten- music- as- a-language. Accessed: 2014–02–04.
67. Craig Wright. 2013. City, Chant and the Topography of Early Music. Harvard University Press, Chapter Quantification in Medieval Paris and How it Changed Western Music.
68. Xiao Xiao and Hiroshi Ishii. 2011. MirrorFugue: Communicating Hand Gesture in Remote Piano Collaboration. In Proc. TEI '11.
69. Xiao Xiao, Anna Pereira, and Hiroshi Ishii. 2013a. MirrorFugue III: Conjuring the Recorded Pianist. In Proc. NIME '13.
70. Xiao Xiao, Basheer Tome, and Hiroshi Ishii. 2014. Andante:. In Proc. NIME '14.
71. Yamaha. Disklavier Pianos. http://usa.yamaha.com/products/musical- instruments/keyboards/disklaviers/. Accessed: 2016–01–07.