P.I.A.N.O.: Faster Piano Learning with Interactive Projection

Katja Rogers[1], Amrei Röhlig[1], Matthias Weing[1], Jan Gugenheimer[1], Bastian Könings[1], Melina Klepsch[2], Florian Schaub[3], Enrico Rukzio[1], Tina Seufert[2], Michael Weber[1]

[1]Institute of Media Informatics, [2]Institute of Psychology and Education
Ulm University, Germany, [firstname.lastname]@uni-ulm.de
[3]School of Computer Science, Carnegie Mellon University, Pittsburgh, fschaub@cs.cmu.edu

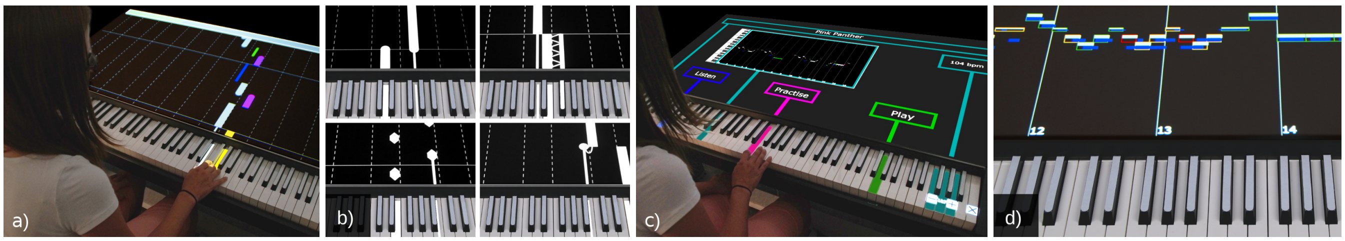

Figure 1. In P.I.A.N.O., (a) music notation is projected onto the piano to facilitate direct mapping of notes to respective piano keys. Correct fingering is supported with color highlights. (b) The basic piano roll notation is extended to support articulations (legato, staccato) and note ornaments (trill, grace notes). (c) The learning process is supported by three learning modes (*listen*, *practice*, *play*). (d) After *practicing* a song, the *play* mode provides detailed feedback on the achieved skill level.

ABSTRACT

Learning to play the piano is a prolonged challenge for novices. It requires them to learn sheet music notation and its mapping to respective piano keys, together with articulation details. Smooth playing further requires correct finger postures. The result is a slow learning progress, often causing frustration and strain. To overcome these issues, we propose P.I.A.N.O., a piano learning system with interactive projection that facilitates a fast learning process. Note information in form of an enhanced piano roll notation is directly projected onto the instrument and allows mapping of notes to piano keys without prior sight-reading skills. Three learning modes support the natural learning process with live feedback and performance evaluation. We report the results of two user studies, which show that P.I.A.N.O. supports faster learning, requires significantly less cognitive load, provides better user experience, and increases perceived musical quality compared to sheet music notation and non-projected piano roll notation.

Author Keywords

Piano; music; instrument learning; interactive projection; piano roll notation; muscial expression; CAMIT.

ACM Classification Keywords

H.5.5 Information Interfaces and Presentation: Sound and Music Computing; K.3.1 Computers and Education: Computer Uses in Education—*Computer-assisted instruction*.

ITS 2014, November 16–19, 2014, Dresden, Germany.

http://dx.doi.org/10.1145/2669485.2669514

INTRODUCTION

Hallam [13] has shown that children and young people benefit from the positive effects of music-making on personal and social development. Learning to play piano as an adult is often motivated by self-actualization and enjoyment [18]. Active music-making has also been found to enhance the health and well-being of elderly people, and can even contribute to recovery from depression [5]. However, those positive effects will only occur when playing is enjoyable and rewarded [13].

For novices, learning to play piano is a prolonged challenge. One reason for this is that conventional sheet music notation (e.g., the western notation based on a five-line staff) does not support the mapping of notes to piano keys very well [32, 23]. First, reading sheet music, called *sight-reading*, requires a vertical to horizontal mapping of notes' pitch information to the respective piano keys. Second, it requires learners to map complex note symbols and articulation marks to the notes' duration and their individual expression. While trained pianists process this mapping automatically, it is a burdensome challenge for novices [28]. Furthermore, in order to play accurately, correct *fingering* is required, i.e., which fingers to use for specific notes and note sequences [27]. These burdens often result in frustration and high dropout rates from piano courses [11], which could be lowered by supporting motivation and faster learning progress [36].

In this paper, we present P.I.A.N.O. to effectively support learning to play piano with interactive projection. In order to avoid the need for sight-reading and to reduce a learner's cognitive load, we developed an enhanced note visualization based on *piano roll notation* (see Fig. 1a). The notation originates from paper rolls used in the late 19th century, on which position and length of holes indicated notes' pitch and duration [7]. Using this approach in a graphical visualization allows the direct inference of the duration of notes (represented as vertical bars). Furthermore, by projecting notes onto the piano, a note's pitch information (a bar's horizontal position) can be mapped directly to the respective piano key. Different colors of notes support correct fingering. We extended the basic piano roll notation with support for legato and staccato articulations, as well as trill and grace note ornaments (see Fig. 1b) as a first step towards fully visualizing sheet music.

P.I.A.N.O. provides three different modes to support the learning process (see Fig. 1c). In the *listen* mode, learners can listen to a song and follow its note visualization to become familiar with its rhythm and melody. The *practice* mode waits for correct key presses before advancing the note visualization. Learners are further supported by live feedback in finding correct keys. To evaluate the learning progress, the *play* mode measures the accuracy and errors of a performance. A detailed feedback screen facilitates identification of parts requiring further practice (see Fig. 1d).

We conducted two user studies to evaluate learning progress, required cognitive load, and user experience of learners when learning a song with P.I.A.N.O. compared to sheet music notation and basic piano roll notation without projection. The results of our within-subjects study (n_1=56) show that P.I.A.N.O. reduces the initial hurdle of starting to play piano by facilitating faster learning with less cognitive load. Results of a week-long between-subjects study with the three systems (n_2=18) show that P.I.A.N.O. facilitates playing with less errors and offers a steeper learning curve for novices practicing a song for one week (please note that the metaphor "steep learning curve" refers to a quick progress in learning throughout the paper). A qualitative rating of the final performances from the second study by 6 piano experts shows that P.I.A.N.O. also increases perceived quality and overall impression of the played music.

Contributions

The main contributions of our work are (1) results showing a significant reduction of cognitive load when learning piano playing with a projection-augmented piano that supports direct mapping of music notation to piano keys. (2) An enhanced piano roll notation conveying rich note information that does not require sight-reading and demonstrates the feasibility of visualizing complex sheet music notation. (3) Three learning modes (*listen*, *practice*, and *play*) that provide song preview, live feedback, and performance evaluation optimized for roll notation that effectively support the learning process. (4) Validation of the effectiveness of the aspects above with quantitative and qualitative evaluation in two user studies, and rating of perceived quality by piano experts.

We first summarize related work before describing the P.I.A.N.O. system in detail, followed by a presentation of its comparative evaluation. We conclude with a discussion of advantages, limitations, and potential extensions of our approach.

RELATED WORK

Music Games

Rhythm games, like *Guitar Hero* [1], reconcile a novice's desire to make music with speed of progress. In such games, colored notes scroll down the screen to a line of markers, which must be mapped to colored buttons on instrument-shaped controllers, and pressed in time to score points. Piano games use basic piano roll notation in a similar fashion. For instance, in the Android app *Pianist HD* [2], vertical bars of different length scroll towards a visual touch keyboard. The arcade game *Keyboardmania* [34] indicates pitch information by showing dots instead of bars on a display above a small keyboard similar to a real piano. The biggest drawback of those games is the use of controllers instead of real instruments and simulated music making, which provides entertainment to players but does not support learning to play real instruments. However, such games can nevertheless contribute to the development of some musical skills, such as visual tracking and rhythmic performance [12, 25].

Synthesia [30] merges rhythm games with actual piano learning on real instruments. A digital piano can be connected via a MIDI interface to a PC, which displays a basic piano roll notation of notes flowing towards a virtual keyboard. Scoring is calculated by correct pitch and duration of each note and can be shared online with other players. The high activity of the scoreboard[1] and the large number of tutorials for specific songs on Youtube[2] show that Synthesia is a widely adopted method for self-educated piano playing. However, players still need to map visual note representations to the instrument. P.I.A.N.O. projects notes directly onto the piano instead. Furthermore, Synthesia's basic piano roll notation provides no information about articulations or note ornaments.

Augmented Pianos

Basic approaches for augmenting a piano keyboard are *Disney's Piano Sound Book* [6] and the *Laugh & Learn Baby Grand Piano* [10] for preschool children. Here, light-up keys ease the mapping of a simplified sheet music notation to the keys of the toy piano. Casio also offers real digital pianos with light-up keys [4].

The main drawbacks of such approaches are that they only indicate the very next keys to play and that they rely only on sheet music notation, which does not support direct mapping of notes to keys [32, 23]. This direct mapping is supported in P.I.A.N.O. by projecting an enhanced piano roll notation directly onto the piano and its keys.

Direct Mapping of Notes to Keys

Toshio Iwai [17] artistically combined projection with direct note mapping to piano keys. Similar to the piano roll notation, light dots could be drawn as notes on a projected surface with a trackball. The light dots then moved towards the piano and were played automatically. While the projection method is similar to our approach, users could only draw dots and not directly play the piano. Yang and Essl [38] augmented a digital piano with a projection setup similar to P.I.A.N.O. They implemented different visualization methods, including a basic piano roll notation, to discuss their influence on the design space of an augmented piano. However, their visualization provided neither any information about articulation or fingering, nor performance feedback. P.I.A.N.O. significantly extends the projection approach by supporting these features and the learning process with interactive learning modes.

In contrast to our goal of supporting piano playing without sight-reading knowledge, Takegawa et al. [31] propose a projection-augmented piano to support learning of sheet music notation. Notes of the sheet notation are visually connected to piano keys

[1] `http://www.synthesiagame.com/scoreboard.aspx`
[2] 1,440,000 search results for "Synthesia" as of September 2014

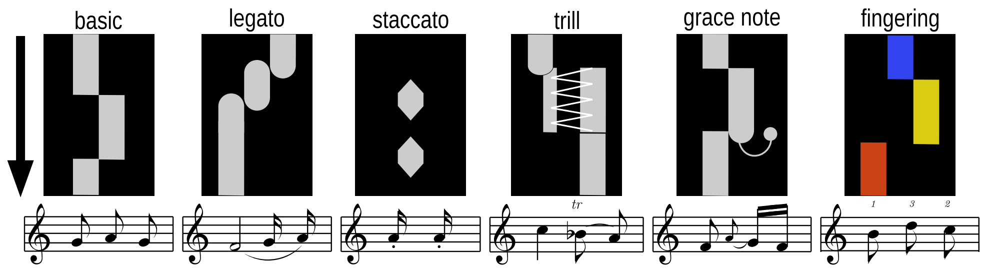

Figure 2. Different note and fingering representation in P.I.A.N.O. (*top*) compared to five-staff notation (*bottom*).

in order to support note to key mapping. The results of a between-subjects study (n=9) suggest that the method efficiently supports sheet music learning. However, the evidence is preliminary with only three learners per method and a short practice session of 30 minutes. Recently Raymaekers et al. [24] presented "The Augmented Piano" (TAP), with a setup similar to P.I.A.N.O.The system also provides a gamified alternative, wherein instead of note visualization the user shoots spaceships with the keys. However, no evaluation of their system was provided.

To the best of our knowledge, our work is the first to provide an extensive evaluation of an enhanced projected roll notation in a short-term study (n_1=56) and a one-week longitudinal study (n_2=18) to compare performance development between different learning systems. Furthermore, we do not aim to teach sight reading, but rather piano playing with a combination of our enhanced note visualization (with articulation marks and ornaments as a first step towards fully visualizing sheet music), correct fingering, and the three tightly coupled integrated learning modes optimized for roll notation.

Fingering Support

The *Concert Hands* [26] are finger sleeves worn during play, which signal correct fingering. Hands are also autonomously guided to the right position by a wrist pilot on a rail mounted in front of the piano. Huang et al. [15] propose a similar glove to passively learn correct fingering by indicated vibrations.

In Synthesia [30], correct fingering is supported by a color for each hand, and numbers (1–5) displayed with notes indicating which finger to use. However, numbers have to be assigned to notes manually beforehand, which requires learners to know which fingering approach is best suited for a song.

Xiao et al. [37] propose a remote piano tutoring system based on projected hands and autonomously moving keys of a player piano. Both student and teacher place a camera on top of the keyboard to capture hand positions and another one in front of the pianist to capture body language. Captured video is projected on to each others' piano. This way, piano teachers can remotely guide accurate fingering and body language.

P.I.A.N.O. uses different note colors to indicate correct fingering, similar to the *Glow Piano Lessons* app for iOS devices [16]. However, in contrast to the iOS app, only notes for important finger switches are colored, in order to draw attention to them without overstraining learners. While this approach requires a short learning phase of color-to-finger mapping, it facilitates fast fingering decisions during play.

THE P.I.A.N.O. SYSTEM

Rather than relying on sheet music, P.I.A.N.O. projects music notation directly onto the instrument. Notes are represented in piano roll notation, i.e., upcoming notes move towards the keyboard, which allows direct mapping of notes to the respective piano keys without any sight-reading knowledge. The current keys to play are illuminated to further ease this mapping. In contrast to basic roll notations common in music video games, our note visualization supports not only pitch and duration of notes, but also additional articulation marks, and recommended fingering, in order to provide similar information as conventional sheet music notation. The learning process itself is supported with different interactive modes following social learning theory [3].

Note Visualization

A major goal in the design of our note representation was to ease the learning process by shifting cognitive capacity from notation mapping to music playing. Traditional sheet music uses a static, symbolic notation, which is hard to match to the analogous actions required in playing the piano [32, 23]. McLachlan et al. [23] also found that spatial, graphical notations lead to improved performance in novices. Therefore, we represent notes with pictorial analogies to reduce extraneous cognitive load for the learner. We refined our note representation in an iterative process with many discussions within our research team (including two piano players and two band members with an average experience of 15 years in playing different instruments), as well as multiple practice sessions and semi-structured interviews with four participants [33].

In our system, upcoming notes are projected on the extended surface behind the keyboard to provide a preview (see Fig. 1a). On the keyboard itself, the current keys are highlighted to signal to the learner what to play. In our pre-study, many participants looked predominantly at the piano keys. Although this visual scope increased upwards with longer practice, we found a preview of upcoming notes on the keys to be very helpful as it indicates where hands have to move next. Therefore, upcoming notes are highlighted with thin extension lines, as shown in Figure 1 a) and b).

Figure 2 shows the main elements of our note visualization. The horizontal and vertical position of notes indicate what key to press (pitch) and when; their vertical length indicates the note's duration, i.e., how long the key has to be pressed. Shape and color of notes provide additional information. Our current visualization supports the articulations *legato* and *staccato*. Legato notes overlap and are extended by a half circle to indicate smooth playing without pause. In contrast, we encoded staccato with triangular ends to evoke "pointier" notes, which should be played

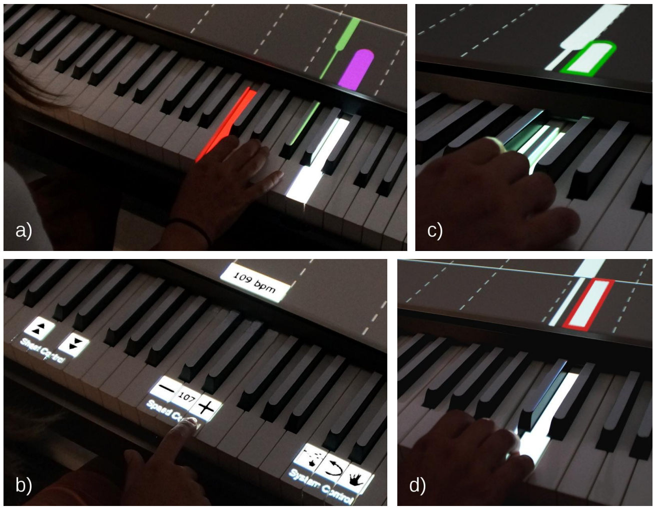

Figure 3. (a) In practice mode, an incorrectly played note is highlighted in red, the correct key in white. (b) The pedal-activated menu allows control of mode-specific settings, e.g., adapting the speed or scrolling through the visualization. (c) Live feedback in play mode highlights correctly played notes with green and (d) errors with red borders.

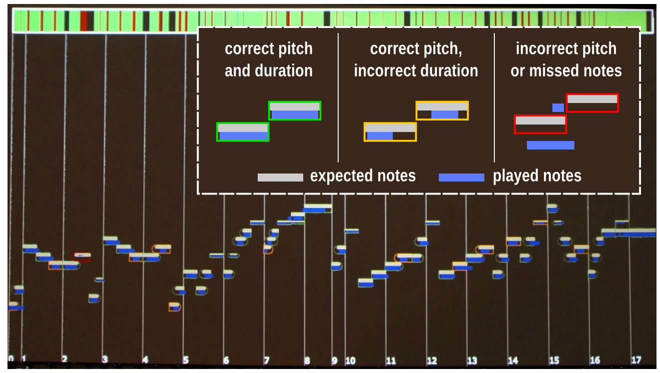

Figure 4. The play mode shows a detailed feedback screen. It shows notes played by the user (*blue bars*), notes played with correct pitch and duration (*green border*), correct pitch but incorrect duration (*yellow border*), and incorrect or missed notes (*red border*). The progress bar at the top provides a respective overview, with missed song parts in *black*.

short and disjointedly. Common ornaments in sheet music are *trill* and *grace notes*. Trills are visualized with interconnecting lines between two notes to indicate fast repeated switching between them. A grace note is a short note played directly before the next one; it is indicated by a connected dot.

Fingering is indicated through colored notes with one color for each finger of the right hand. We chose the colors for their contrast between each other in order to prevent mistakes due to chromatic similarity. Our pre-study showed that fingering information for every note overwhelms learners. Therefore, similarly to sheet music, P.I.A.N.O. highlights only important finger switches. Notes without fingering information are gray and can be played with any suitable finger.

Learning Modes

Social learning theory [3] distinguishes four steps of learning. The first is *attention*; learners observe a process. In the *retention* step, learners try to remember what they observed. With *reproduction*, learners perform the observed behavior themselves. Further practice leads to improvement and skill advancement. *Motivation* consists of reinforcement and punishment to ensure that learners continue practice; this is essential for observational learning to be successful. P.I.A.N.O. supports social learning theory by offering three learning modes: *listen*, *practice*, and *play* (accessible from the main menu, see Fig. 1c). System and navigation controls are available by pushing down any of the piano's foot pedals. Mode-specific controls, e.g., adjusting speed, are then projected onto the keys and can be activated via keypress (see Fig. 3b).

Listen mode (attention)

In the *listen* mode, learners can listen to the song and follow its note visualization (*attention*). Thus, learners become acquainted with the characteristics of different song parts. This auditory learning aspect was revealed as highly important in conversations with experts (i.e., a professional piano teacher and a professor of psychology). Following social learning theory [3], the synchronized playback of auditory and visual information should increase learners' attention.

Practice mode (retention)

The *practice* mode provides learners with an environment to practice accurate playing in order to support retention. The system encourages correct playing without haste by waiting for correct key presses and duration before continuing in the song. When a wrong note is played, the system brightly highlights the correct key and marks the wrongly-pressed key in red (see Fig. 3a). The pedal-operated context menu allows leaners to jump back and forth in the current song, change the speed or song part, or return to the main menu (see Fig. 3b).

Play mode (reproduction and motivation)

After practice, learners can assess their performance in the *play* mode by playing the song in one piece, hereby reproducing the previously observed and memorized behavior (*reproduction*). The speed of the song is adjustable to allow evaluation at various speeds. While playing, the learner receives with live feedback as part of the note visualization. A green border around the currently played note indicates a correctly played note; a red border indicates an incorrectly played note (see Fig. 3c & d). A progress bar at the top provides a summarized overview of correct (green), incorrect (red), and missed song parts (black). The progress bar is also shown in the final feedback screen (see Fig. 4), which provides detailed information of played notes (blue), and expected notes (grey). A colored border indicates correctly played notes (green), incorrectly played or missed notes (red), and notes played with correct pitch but incorrect duration (yellow). This feedback allows learners to identify which parts were played well and which require further practice, thus, it provides *motivation*.

Technical Aspects

We use a standard digital piano (Thomann DP-25), which supports MIDI communication. It is connected to a desktop computer (Intel Core i5-3470, 4 GB RAM, Intel HD Graphics 2500) via a MIDI/USB interface (Roland UM-One). The projection on the piano is created by a short-throw projector (BenQ MW516) displaying the computer's video output. It is mounted above the piano using a tripod. A smooth projection

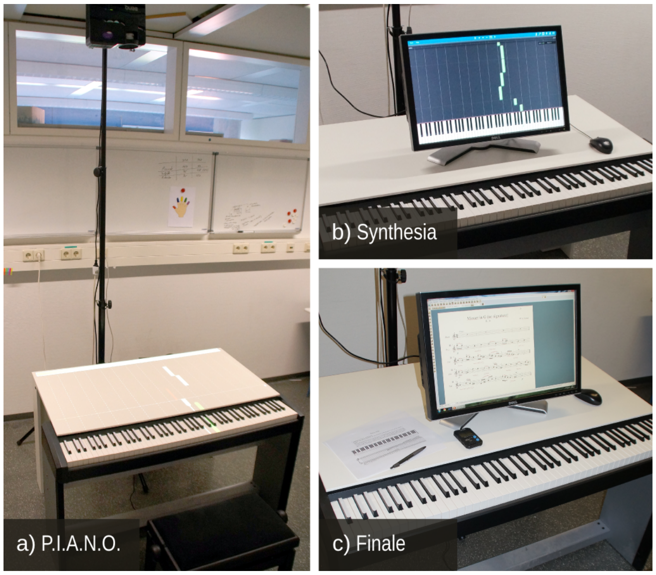

Figure 5. Setup of the compared systems in the user studies: (a) P.I.A.N.O. (interactive projection), (b) Synthesia (piano roll notation), and (c) Finale (sheet music notation).

surface is provided by a wooden plate attached to the piano, covering its top and extending towards the back (see Fig. 5a).

The MIDI protocol identifies key presses and their duration: each key press and release generates an event message , including a corresponding timestamp. This allows us to compare the system's representation of a song's expected notes with notes played on the piano in order to provide live feedback. The internal representation of a song, used for evaluation as well as visualization, is generated by parsing song files in MusicXML format [22]. Thus, any song available in this format can be played with P.I.A.N.O.

In listen mode, the computer generates MIDI messages to control piano playback. In practice mode, played notes are evaluated to generate the colored hints indicating correctly or incorrectly pressed keys (see Fig. 3c & d). In play mode, live feedback is generated based on real-time information of pressed keys. In addition, all key presses and key releases are logged throughout a learner's performance. This log is then compared to the ideal performance generated by the system to yield the result screen displayed at the end (see Fig. 4).

EVALUATION

In order to examine P.I.A.N.O.'s impact on the learning performance of novices, we conducted two experimental studies. We evaluated the learning performance with the interactive projection of P.I.A.N.O. in comparison to a more game-orientated *roll notation* and the traditional way of learning the piano with *sheet music*. As specific comparison systems, we selected commercially available software: *Synthesia* [30], a popular educational piano game, and *Finale* [21], a music software with traditional sheet music notation. Using software for all three systems enabled us to evaluate performance with a consistent quantitative measurement approach.

The first study (n_1=56) employed a within-subjects design to compare initial performance, experienced cognitive load, and perceived user experience of the three systems. A between-subjects study followed (n_2=18) to assess learners' performance over one week. We were further interested in the systems' impact on perceived quality and overall impression of performances. Thus, final recordings of the second study were blindly and independently rated by 6 piano experts.

System Setup

As the three systems were identical for both studies, we first outline their setup and the employed quantitative measurement approach, before discussing the specifics of the studies. Figure 5 shows the experimental setup of each system. The P.I.A.N.O. system was set up as described above (see Fig. 5a). For Synthesia and Finale, we used a 24"-display which was placed in front of the learners (see Fig. 5b & c). Sound settings were consistent for all systems and sound output was directly generated by the piano's built-in speakers.

Synthesia [30] was chosen due to its popularity with self-tutoring piano learners. Synthesia visualizes notes in a basic piano roll notation. It offers a *melody practice* feature, similar to our practice mode, as well as a *song recital* feature that demands a defined speed similar to our play mode. Fingering information is limited in both its availability and visualization, and there is no visualization of articulation techniques. When connected to a piano via MIDI, Synthesia can assess a learner's performance from piano key presses.

Finale [21] is a software for composing and displaying sheet music. To achieve comparable conditions to our play mode between all systems, it was essential that the dynamic indication of the current position within a song was supported. Finale displays a dynamic marker that runs through the depicted sheet music at a defined speed. In addition, learners were offered a metronome to support keeping a song's rhythm. In contrast to P.I.A.N.O. and Synthesia, sheet music requires learners to perform sight-reading. Thus, learners were given additional help in form of a notation guide and an image mapping each note in staff notation to the keyboard. The piano's middle C note was marked with a colored sticker.

In order to obtain consistent quantitative performance results of all systems, we implemented a measurement tool that recorded the piano's MIDI output, i.e., actually pressed keys, in synchronization with the evaluated system. Similar to Drake and Palmer [8], recorded session logs were analyzed against expected correct notes: We differentiated between *missed notes*, *incorrectly pressed notes* (pitch errors) and *correctly pressed notes* (correct pitch). Incorrectly pressed notes were counted independently from the number of expected notes, i.e., pressing a wrong key twice during the duration of one expected note resulted in two incorrectly pressed notes. This allowed for a more precise distinction of incorrect performances. Correctly pressed notes were further sub-classified regarding their duration, i.e., *correct* and *incorrect duration notes*. Duration was evaluated based on absolute length comparison and relative timestamp matching.

User Study 1: Initial Performance and User Experience

In a within-subjects study with the three systems, we evaluated the performance of novices with no previous piano experience in learning to play the right-hand part of a song in 15 minutes. We expected that *learning with P.I.A.N.O. would result in higher learning performance* (more correctly pressed notes, less incorrectly pressed and missed notes) (H_1) than Synthesia and Finale, due to the direct mapping of notes onto keys. For the same reason, we expected that *P.I.A.N.O. would create less intrinsic, less extraneous, and higher germane cognitive load* (H_2). *Intrinsic cognitive load* (ICL) results from the learning material itself, its element interactivity, and prior knowledge of the learner [29]. *Extraneous cognitive load* (ECL) occurs due to the design of the learning environment, whereas *germane cognitive load* (GCL) arises from the learner's intrinsic motivation and concentration on understanding the learning material. Hence, a system should strive for low ICL and ECL, but high GCL [29]. Finally, we expected that *P.I.A.N.O.'s user experience would be rated higher* (H_3) than those of the other systems.

Method and Material
With the help of two experienced pianists and a professional piano teacher, we selected three songs of similar difficulty and length (16 bars, 26 seconds) that have a dominant melody played with the right hand. The songs were chosen from a grade 1 syllabus of exam pieces from the Associated Board of the Royal Schools of Music [20]: "*Das Ballett*" by D. G. Türk, "*Minuet in G*" by W. A. Mozart, and "*Moderato*" by A. F. Gedike. To verify that the songs were of similar difficulty, we performed a treatment check based on learning performance with sheet music notation. We could not find a significant difference regarding correctly pressed notes. (song 1 vs. 2: *MD*=1.91, *SE*=4.40, n.s.; song 1 vs. 3: *MD*=3.62, *SE*=4.34, n.s.; song 2 vs. 3: *MD*=5.53, *SE*: 4.34, n.s.).

At the beginning of the session, participants completed a questionnaire regarding their demographic information, and musical background (i.e., experience with other instruments, and sight-reading skills). We further used a subtest of "the kit of factor-referenced cognitive tests" [9] to assess learners' spatial ability. Each participant interacted with all three systems and learned a different song with each. The order of systems and songs was counter-balanced (Latin square). For each system, participants were first given a written introduction to the specific system, and encouraged to try out the described features. Participants then listened to the assigned song three times (P.I.A.N.O.'s *listen* mode, Synthesia's *watch and listen only*, and Finale's *playback* mode). This was the only time when listening to the song was allowed. Playing performance was measured at four times: after listening to the song *without practice* (T_1), *after practicing the first half of the song (5 min.)* (T_2), *after practicing the second half of the song (5 min.)* (T_3), and a final measurement *after practicing the whole song (5 min.)* (T_4). However, each measurement corresponds to the learning performance for the whole song as measured in the system-specific play-through mode (P.I.A.N.O.'s *play* mode, Synthesia's *song recital*, and Finale's *dynamic-marker* mode).

After each system session, participants completed an 8-item survey [19], designed to differentiate between the three parts of cognitive load. Each item posed a statement to be rated on a 7-point Likert scale. Perceived user experience of each system was assessed with the AttrakDiff survey [14], which consists of 28 bipolar verbal anchors rated on a 7-point scale. AttrakDiff results in four scales: *perceived pragmatic quality (PQ)* measuring the support for achieving a goal, *hedonic quality – stimulation (HQ-S)* measuring perceived novelty and potential to grab users' attention,

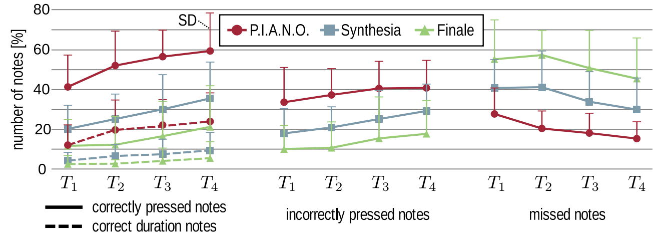

Figure 6. Learning performance for P.I.A.N.O., Synthesia, and Finale in first study for measurements T_1–T_4.

hedonic quality – identification (HQ-I) measuring potential of identification with the system, and *perceived attractiveness (ATT)*.

Sessions lasted about two hours on average, with 20-30 min. per system. Due to the long duration, participants were offered snacks and drinks, and afterwards allowed to choose from a variety of larger sweet collections as compensation.

Participants
A total of 56 right-handed learners participated in the study, all of them novices in playing piano. The average age of participants was 23 years (*SD*=3.92); they were fairly balanced in gender (32 female, 24 male), and most were students. Some participants had a basic level of ability to read sheet music; on average, participants achieved 6.98 out of 15 points in the test of sight-reading skills (*SD*=4.59).

Learning Performance
Figure 6 shows the results of the quantitative analysis of the learners' short-term learning performances. For all measurements (T_1–T_4), P.I.A.N.O. learners achieved a higher percentage of correct notes and also performed better regarding duration accuracy. An analysis of variance with repeated measures (rANOVA) revealed a significant difference in the number of *correctly pressed notes* ($F(2,88)=121.64$, $p<.001$, $\eta^2=0.73$) and *correct duration notes* ($F(1.37,61.85)=85.31$, $p<.001$, $\eta^2=0.66$) for all three systems. An analysis of contrasts showed that P.I.A.N.O. learners hit significantly more correct notes than Synthesia ($F(1,44)=120.42$, $p<.001$, $\eta^2=0.73$) and Finale learners ($F(1,44)=173.31$, $p<.001$, $\eta^2=0.80$), and also played more *correct duration notes* compared to Synthesia ($F(1,45)=78.04$, $p<.001$, $\eta^2=0.63$) and Finale ($F(1,45)=108.83$, $p<.001$, $\eta^2=0.71$).

With respect to *incorrectly pressed notes*, Figure 6 shows that P.I.A.N.O. learners also pressed more incorrect notes on average than learners of the other systems. An rANOVA revealed significant differences between all systems ($F(1.75,76.92)=60.11$, $p<.001$, $\eta^2=0.58$). The contrasts show that learning with P.I.A.N.O. results in more incorrectly pressed notes than learning with Synthesia ($F(1,44)=60.62$, $p<.001$, $\eta^2=0.58$) or Finale ($F(1,44)=85.70$, $p<.001$, $\eta^2=0.66$). However, on average, P.I.A.N.O. learners also tried to play more notes, which is reflected in the lower number of *missed notes*. Missed notes significantly differ between all systems (rANOVA: $F(1.73,77.79)=89.83$, $p<.001$, $\eta^2=0.67$) and P.I.A.N.O. learners missed significantly less notes than Synthesia learners ($F(1,45)=72.30$, $p<.001$, $\eta^2=0.62$) and Finale learners ($F(1,45)=165.04$, $p<.001$, $\eta^2=0.79$). In summary, with the exception of incorrectly pressed notes, the

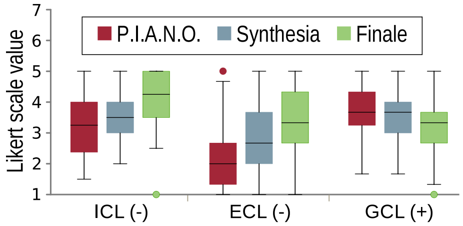

Figure 7. Intrinsic, extrinsic, and germane (positive) cognitive load of P.I.A.N.O., Synthesia, and Finale in study 1.

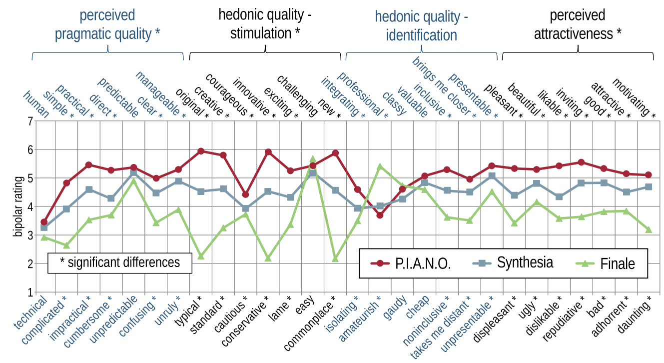

Figure 8. User experience results rated with AttrakDiff's 28 bipolar verbal anchors.

combined results confirm our hypothesis that P.I.A.N.O. improves initial learning performance (H_1).

Cognitive load

Figure 7 shows the cognitive load results. Internal consistency (Cronbach's α) of items was confirmed for intrinsic cognitive load (ICL) (P.I.A.N.O.: α=.81; Synthesia: α=.79; Finale: α=.70), extrinsic cognitive load (ECL) (P.I.A.N.O.: α=.81; Synthesia: α=.86; Finale: α=.87), and germane cognitive load (GCL) (P.I.A.N.O.: α=.62; Synthesia: α=.53; Finale: α=.58).

According to an rANOVA, significant differences existed for ICL (F(2,108)=18.39, p<.001, η^2=0.25), ECL (F(2,108)=16.88, p<.001, η^2=0.24), and GCL (F(2,108)=10.02, p<.001, η^2=0.16). P.I.A.N.O. induces significantly lower ICL than both Synthesia (F(1,54)=5.02, p<.05, η^2=0.09), and Finale (F(1,54)=38.00, p<.001, η^2=0.41), indicating that the complexity of the note notation is reduced by P.I.A.N.O.'s projected roll notation compared to Synthesia or Finale. ECL was also rated significantly lower with P.I.A.N.O. than with Synthesia (F(1,54)=9.64, p<.01, η^2=0.15) or Finale (F(1,54)=28.71, p<.001, η^2=0.35). This suggests that P.I.A.N.O. reduces split attention: In Synthesia, attention is split between the keyboard and the display. The metronome may further split attention in Finale. P.I.A.N.O. learners only needed to follow the projected notation which is seamlessly integrated with the keyboard.

Furthermore, P.I.A.N.O. induces significantly more positive GCL than Finale (F(1,54)=18.59, p<.001, η^2=0.26). We could not find a significant difference between P.I.A.N.O. and Synthesia (F(1,54)=2.88, n.s.); presumably their roll notations promote similar motivation. Thus, H_2 can be accepted, as P.I.A.N.O. induces lower intrinsic and extrinsic load than Synthesia and Finale, and higher germane load than Finale.

User experience

The per-item ratings of the three systems on the AttrakDiff scales are shown in Figure 8. P.I.A.N.O. was ranked highest in almost all cases, with Finale consistently ranked lowest. Significant differences exist for perceived pragmatic quality (PQ) (F(2,108)=19.97, p<.001, η^2=0.25), stimulation (HQ-S) (F(2,108)=117.76, p<.001, η^2=0.69), and attractiveness (ATT) (F(2,108)=24.75, p<.001, η^2=0.31). Differences were not significant for the identification scale (HQ-I). More specifically, for all scales P.I.A.N.O. was rated significantly higher than Synthesia (PQ: F(1,54)=8.64, p<.01, η^2=0.15; HQ-S: F(1,54)=46.78, p<.001, η^2=0.46; ATT: F(1,54)=11.85, p<.001, η^2=0.18), and Finale (PQ: F(1,54)=28.26, p<.001, η^2=0.34; HQ-S: F(1,54)=225.33, p<.001, η^2=0.81; ATT: F(1,54)=39.50, p<.001, η^2=0.42). As a result, H_3 can also be accepted.

User Study 2: One-week Performance

In the first study, P.I.A.N.O. and its interactive projection outperformed classic piano roll notation (Synthesia) and sheet music notation (Finale) for most of the quantitative performance metrics, cognitive load, and user experience. These results suggest that projecting an enhanced note visualization onto the piano significantly eases piano learning for novices. In order to validate P.I.A.N.O.'s superiority regarding performance (correct, incorrect, and missed notes), we conducted a second study with the primary goal of evaluating learner performance over a longer period of time. We recruited a new group of participants (n_2=18) that practiced with one of the systems each day for one week. We opted for a between-subjects design to avoid performance interference from multiple systems. Our hypothesis was that *P.I.A.N.O. would retain a steeper learning curve over one week than Synthesia or Finale*, based on measured performance (H_4).

Method and Material

In order to avoid a ceiling effect, the song for Study 2 was chosen through a preliminary study with 3 participants in order to select a song of which the right-hand part was neither too easy nor too hard to learn. Based on the results, we chose a slightly simplified version of Schumann's *Träumerei* (16 bars, 116 notes), wherein the difficulty consisted mostly of legato articulation and several two- and three-note chords. Three groups, each consisting of six participants were monitored while they practiced with one system (P.I.A.N.O., Synthesia, or Finale) for a total of 5 consecutive days. Each daily session consisted of a brief questionnaire regarding their pre-session physical condition, followed by listening to the song one time, a 15-minute practice session, and a daily play-mode measurement (D_1–D_5).

Participants

A total of 18 learners participated in this study; mainly students, 26 years old on average (SD=4.45), and balanced in gender (9 female, 9 male), which were evenly distributed across the three groups. The preliminary sight-reading test revealed no differences in the average sight-reading skills among learners of P.I.A.N.O. (M=8.8, SD=5.42), Synthesia (M=9, SD=4.69), and Finale (M=8.3, SD=4.32).

Learning Performance

The participants' learning curves for the three systems are depicted in Figure 9. For all measurements, P.I.A.N.O. learners achieved a higher percentage of correct notes, with better duration

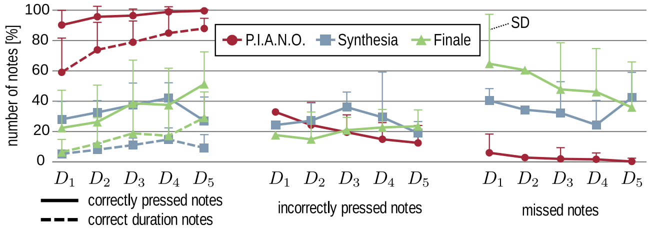

Figure 9. Development of learning performance in second study for 5 consecutive days (D_1–D_5).

accuracy (on average 88% *correct duration notes* on the last day) and lower percentage of *incorrectly played* and *missed notes* (12% and 0.3% on last day). An rANOVA revealed significant differences in the learning performance regarding *correctly pressed notes* (F(2,15)=34.26, p<.001, η^2=0.82), *correct duration notes* (F(2,15)=51.43, p<.001, η^2=0.87), and *missed notes* (F(2,15)=12.41, p<.01, η^2=0.62) at all five measurements. A significant interaction of system and measurement point was found, showing a steeper learning curve of P.I.A.N.O. learners regarding *correctly pressed notes* (F(3.62,3)=5.15, p<.01, η^2=0.41), *correct duration notes* (F(4.81,3)=4.24, p<.01, η^2=0.36), and *missed notes* (F(3.50,3)=3.59, p<.05, η^2=0.32) compared to those of Synthesia and Finale learners. For *incorrectly pressed notes*, an rANOVA showed no significant differences between systems (F<1, n.s.) or measurements (F(2.00,13)=1.88, n.s.). However, a difference in interaction was found (F(4.00,3)=3.69, p<.05, η^2=0.33); *incorrectly pressed notes* decreased over time for P.I.A.N.O. learners, were unstable for Synthesia learners, and increased for Finale learners. Together the results for *correctly/incorrectly pressed notes*, *correct duration notes*, and *missed notes* confirm our hypothesis that P.I.A.N.O. leads to an improvement in learning performance over time (H_4).

Expert Evaluation: Perceived Quality

6 piano experts were recruited (4 piano teachers, 2 professional players) with 28 years of experience on average, in order to gain qualitative performance ratings. We expected that *experts would rate recordings of P.I.A.N.O. learners higher than those of Synthesia and Finale learners in terms of perceived quality and overall impression* (H_5).

Method and Material

On the final day of study 2, we recorded one *practice* and one *play* session of each participant, resulting in 36 recordings (12 for each system). *Play* sessions were used for the expert evaluation because they provided consistent tempo, which is necessary for a realistic and comparable scenario. However, in some cases the play mode resulted in incomplete recordings when learners were unable to recover from playing mistakes. *Practice* sessions, therefore, allowed us to provide experts with more complete recordings for their ratings.

Each expert listened to all 36 recordings and filled out a survey for each. Recordings were anonymized regarding the system and playing mode. The three evaluated systems were presented to the experts in randomized order to avoid fatigue and assimilation effects. The rating criteria in the survey were a subset from a music performance adjudication set created by Wrigley et al. [35], and consisted of *pitch accuracy*, *duration accuracy*, *tempo*, *rhythm*, *continuity*, *dynamic*, *articulation*, *interpretation*, and *overall impression*. Each criterion had to be rated on a 7-point Likert scale (from *very poor* to *very good*). Furthermore, experts could comment on positive and negative aspects about the session in a text field. As a reward, each expert received a 10-Euro shopping coupon.

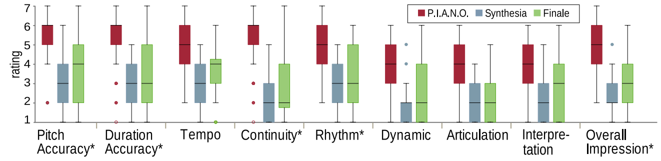

Figure 10. Expert ratings of perceived quality for 12 recordings (6 practice, 6 play) of each system.

Results

The ratings for associated play and practice sessions were combined by calculating their means. Calculated inter-rater reliability (Krippendorff's α) showed reliable ratings for 5 of the 9 scales: *pitch accuracy* (α=0.68), *duration accuracy* (α=0.66), *rhythm* (α=0.69), *continuity* (α=0.61), and *overall impression* (α=0.70). P.I.A.N.O. recordings were rated higher for all scales compared to recordings of Synthesia and Finale (see Fig. 10). For the reliable scales, ANOVAs revealed significant differences in pitch accuracy (F(2,15)=24.30, p<.001, η^2=0.76), duration accuracy (F(2,15)=11.92, p<.01, η^2=0.61), continuity (F(2,15)=15.61, p<.001, η^2=0.68), rhythm (F(2,15)=14.47, p<.001, η^2=0.66), and the overall impression (F(2,15)=13.67, p<.001, η^2=0.65). Thus, we can confirm our hypothesis H_5 for these scales.

DISCUSSION

Both user studies revealed significant differences in multiple variables between the three systems. P.I.A.N.O. learners were able to play more correct notes upon first play, and improved their performance more notably regarding duration accuracy, incorrectly pressed notes (pitch errors), and missed notes. The group of piano experts further confirmed that performances of P.I.A.N.O. learners achieved higher overall impression and perceived quality after one week of practice. Additionally, learning with P.I.A.N.O. resulted in less negative intrinsic and extrinsic cognitive load, more positive germane cognitive load, and provided a better user experience.

These results confirm the effectiveness of our projected piano roll notation, which allows direct mapping of notes onto the keys and does not require sight reading skills. Learners are able to concentrate on playing respective keys correctly (in terms of duration and basic articulation), rather than having to translate sheet music notation and find the correct keys. While trained pianists process mapping of sheet music notation automatically, it is a burdensome challenge for novices [28] and often leads to frustration [11, 36] – as also indicated by the lower average germane load measured for Finale. As one Finale learner commented: *"Without sight-reading skills, this task is very complex and absolutely frustrating."* P.I.A.N.O. eliminates this hurdle and allows song playing instantly. Most P.I.A.N.O. learners commented that time went by very quickly and that they had fun: *"Great when you want to learn playing piano in 5 minutes"*; *"very innovative and intuitive system which is fun to use."* In summary, learning piano with interactive projection appears to be easier and more fun than learning with sheet music notation or piano roll notation without direct mapping.

Limitations and Future Work

The study participants were drawn mainly from the current and former student population, and thus provide a fairly consistent age and educational background. Further studies are needed to investigate different populations, e.g., children or older adults. The second user study took place over the course of one week. While the measured learning curves proved highly promising for P.I.A.N.O., their further development over several weeks, months or even years is also of interest and might provide further valuable insights.

P.I.A.N.O. currently supports only right-hand song parts to target an early stage of learning. In order to provide support for more advanced learners and learning progression, we plan to extend our system to accommodate left-hand parts, as well as advanced features of conventional sheet music (e.g., crescendo or fermata). These visualization features could dynamically appear based on learners' performance. The system should also support fingering recognition, i.e., determining whether the learner pressed the key with the correct finger. We are currently integrating fingering recognition into our system, and plan to evaluate its influence on learning performance in a longitudinal study over several weeks.

Some learners were concerned about being too dependent on the system, and the lack of sheet music notation: *"Without the system I would not be able to play the song."* and *"You do not learn sight-reading."* While it was not our goal to support sight-reading skills, roll notation is unlikely to efficiently support the development of such skills. However, most learners in our studies were not interested in learning sheet music.

A common concern of consulted piano teachers was that the projected roll notation may lead to too much focus on correctness ("stiff playing"), thus hindering the development of richer musical expression. However, the results of our expert evaluation seem to refute those concerns. Although some qualitative scales (e.g., articulation and interpretation) did not provide significant differences, they do show higher tendencies. Some experts were initially concerned about rating novices who practiced for only one week, but were positively surprised (without knowing the used system) by recordings performed by P.I.A.N.O. players. This is also reflected in the significantly higher ratings on the overall impression scale. Although P.I.A.N.O. does not directly teach advanced musical expression, it could support learners in focusing on musical expression rather than on mapping notes to keys, by avoiding the burden of sight-reading. We plan to conduct extended longitudinal studies to examine whether these assumptions hold for practice spanning several weeks or months.

Developing musical expression is a prolonged process and playing correct notes is essential before proceeding to advanced skills. Our studies show that our approach supports quick success in terms of correctness at the first stage of piano playing. This success is essential for beginners when learning an instrument, as it motivates learners, could lower dropout rates and could even encourage people to start playing piano. While P.I.A.N.O. does not intend to replace individualized advice by expert piano teachers, it provides a significant improvement over current self-tutoring approaches. Piano teachers should further guide players with detailed advice, e.g., on fingering techniques and musical expression. This valuable advice remains irreplaceable by a system, although systems like MirrorFugue [37] try to support giving such advice remotely. In contrast, our aim was to support novices without sight-reading knowledge in faster learning and playing, focusing on developing basic playing skills and articulation. Our results show that this goal has been achieved, highlighting the validity of our chosen approach.

CONCLUSION

In this paper, we introduced P.I.A.N.O. as a novel approach to learning piano without sight-reading skills. Our enhanced piano roll notation is capable of depicting a note's pitch and duration, as well as a variety of articulation techniques (legato, staccato, grace notes, and trill notes). The projected notation allows a direct mapping of notes to respective keys, while three different modes further support the learning process. We conducted two user studies to compare the learning performance of P.I.A.N.O. to Synthesia and Finale, which use a basic piano roll notation without direct mapping and traditional sheet music, respectively. The results show that P.I.A.N.O. induces less cognitive load, better supports initial learning performance and faster progress over one week of practice, provides better user experience and leads to better perceived quality than Synthesia and Finale. Based on our results, we argue that the projected roll notation is a viable alternative to sheet music notation for beginners looking for a rewarding approach towards learning to play the piano.

ACKNOWLEDGMENTS

The authors would like to thank all study participants, the anonymous reviewers for their valuable feedback. This work was partially supported by the Transregional Collaborative Research Centre SFB/TRR 62 "Companion-Technology of Cognitive Technical Systems" funded by the German Research Foundation (DFG).

REFERENCES

1. Activision Publishing, Inc. Guitar Hero, accessed: Jun. 2014. `http://guitarhero.com/`.
2. Android appliaction. Pianist HD, accessed: Jun. 2014. `https://play.google.com/store/apps/details?id=com.rubycell.pianisthd`.
3. Bandura, A., and McClelland, D. C. Social learning theory. *General Learning Press* (1977).
4. Casio. Lighted keys keyboard LK-280, accessed: Jun. 2014. `http://www.casio.com/products/Digital_Pianos_&_Keyboards/Lighted_Keys/LK-280/`.
5. Creech, A., Hallam, S., McQueen, H., and Varvarigou, M. The power of music in the lives of older adults. *Research Studies in Music Education 35*, 1 (2013).
6. Disney. *Mickey Mouse Clubhouse Piano Sound Book: Mickey's Piano Party*. Publications International, 2009.
7. Dolan, B. *Inventing entertainment: the player piano and the origins of an American musical industry*. Rowman & Littlefield Publishers, Lanham, Md., 2009.
8. Drake, C., and Palmer, C. Skill acquisition in music performance: Relations between planning and temporal control. *Cognition 74*, 1 (2000), 1–32.
9. Ekstrom, R. B., French, J. W., and Harman, H. H. *Manual for kit of factor-referenced cognitive tests*. Educational Testing Service, Princeton N.J., 1976.

10. Fisher Price. Laugh & Learn Baby Grand Piano, accessed: Jun. 2014. http://www.fisher-price.com/en_AU/brands/babytoys/products/40014.
11. Giomi, E. C. "I do not want to study piano": early predictors of student dropout behavior. *Bulletin of the Council for Research in Music Education*, 161 (2004).
12. Gower, L., and McDowall, J. Interactive music video games and children's musical development. *British Journal of Music Education 29*, 1 (2012), 91–105.
13. Hallam, S. The power of music: Its impact on the intellectual, social and personal development of children and young people. *International Journal of Music Education 28*, 3 (2010), 269–289.
14. Hassenzahl, M., Burmester, D. M., and Koller, F. AttrakDiff: a questionnaire to measure perceived hedonic and pragmatic quality. In *Mensch & Computer*, Springer (2003).
15. Huang, K., Starner, T., Do, E., Weiberg, G., Kohlsdorf, D., Ahlrichs, C., and Leibrandt, R. Mobile music touch: mobile tactile stimulation for passive learning. In *Proc. CHI '10*, ACM (2010).
16. iOS application. Glow Piano Lessons, accessed: Jun. 2014. https://itunes.apple.com/us/app/glow-piano-lessons/id441970188.
17. Iwai, T. Piano–as image media. *Leonardo 34*, 3 (2001).
18. Jutras, P. J. The benefits of adult piano study as self-reported by selected adult piano students. *Journal of Research in Music Education 54*, 2 (July 2006), 97–110.
19. Klepsch, M., and Seufert, T. Subjective Differentiated Measurement of Cognitive Load. In *Proc. 5th Intl. Cognitive Load Theory Conf.* (2012).
20. London-Based Examinations Board. Associated Board of the Royal Schools of Music (ABRSM), accessed: Jun. 2014. http://de.abrsm.org/en/home.
21. MakeMusic, Inc. Finale - music notation software, accessed: Jun. 2014. http://www.finalemusic.com/.
22. MakeMusic, Inc. MusicXML, accessed: Jun. 2014. http://www.musicxml.com/.
23. McLachlan, N. M., Greco, L. J., Toner, E. C., and Wilson, S. J. Using spatial manipulation to examine interactions between visual and auditory encoding of pitch and time. *Frontiers in Psychology 1* (2010).
24. Raymaekers, L., Vermeulen, J., Luyten, K., and Coninx, K. Game of tones: Learning to play songs on a piano using projected instructions and games. In *Proc. CHI EA '14*, ACM (2014), 411–414.
25. Richardson, P., and Kim, Y. Beyond fun and games: A framework for quantifying music skill developments from video game play. *Journal of New Music Research 40*, 4 (2011), 277–291.
26. Rubato productions. Concert Hands, accessed: Jun. 2014. http://www.concerthands.com/products.html.
27. Sloboda, J. A., Clarke, E. F., Parncutt, R., and Raekallio, M. Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance 24*, 1 (1998), 185–203.
28. Stewart, L., Walsh, V., and Frith, U. Reading music modifies spatial mapping in pianists. *Perception & psychophysics 66*, 2 (2004), 183–195.
29. Sweller, J., Van Merrienboer, J. J., and Paas, F. G. Cognitive architecture and instructional design. *Educational Psychology Review 10*, 3 (1998), 251–296.
30. Synthesia LLC. Synthesia - piano game, 2013. http://www.synthesiagame.com/.
31. Takegawa, T., Terada, T., and Tsukamoto, M. A piano learning support system considering rhythm. In *Proc. ICMC '12*, Michigan Publishing (2012).
32. Tan, S.-L., Wakefield, E. M., and Jeffries, P. W. Musically untrained college students' interpretations of musical notation: sound, silence, loudness, duration, and temporal order. *Psychology of Music 37*, 1 (2009), 5–24.
33. Weing, M., Röhlig, A., Rogers, K., Gugenheimer, J., Schaub, F., Könings, B., Rukzio, E., and Weber, M. P.I.A.N.O.: Enhancing Instrument Learning via Interactive Projected Augmentation (demo). In *UbiCom '13 Adjunct Proceedings* (2013).
34. Wikipedia. Keyboardmania. version ID: 542589199, 2013. http://wikipedia.org/wiki/Keyboardmania.
35. Wrigley, W. J., and Emmerson, S. B. Ecological development and validation of a music performance rating scale for five instrument families. *Psychology of Music 41*, 1 (2013), 97–118.
36. Wristen, B. Demographics and motivation of adult group piano students. *Music Education Research 8*, 3 (2006), 387–406.
37. Xiao, X., Pereira, A., and Ishii, H. MirrorFugue III: conjuring the recorded pianist. In *Proc. NIME'13* (2013).
38. Yang, Q., and Essl, G. Visual associations in augmented keyboard performance. In *Proc. NIME '13* (2013).