EyePiano: Leveraging Gaze For Reflective Piano Learning

Citation for the original published paper (version of record):

Karolus, J., Sylupp, J., Schmidt, A. et al (2023). EyePiano: Leveraging Gaze For Reflective Piano Learning. DESIGNING INTERACTIVE SYSTEMS CONFERENCE, DIS 2023: 1209-1223.
http://dx.doi.org/10.1145/3563657.3596065

EyePiano: Leveraging Gaze For Reflective Piano Learning

Jakob Karolus
jakob.karolus@dfki.de
German Research Center for Artificial Intelligence (DFKI)
Kaiserslautern, Germany
RPTU Kaiserslautern-Landau
Kaiserslautern, Germany

Johannes Sylupp
sylupp-jojo@gmx.de
LMU Munich
Munich, Germany

Albrecht Schmidt
albrecht.schmidt@ifi.lmu.de
LMU Munich
Munich, Germany

Paweł W. Woźniak
pawel.wozniak@chalmers.se
Chalmers University of Technology
Gothenburg, Sweden

ABSTRACT

Mastering skills which involve high dexterity, such as playing the piano, requires extensive guidance through personal teaching. Understanding how we can leverage data from sensor-based systems to improve the learning process, allows us to build interactive systems which effectively facilitate skill acquisition. To explore such possibilities, we developed EyePiano—a gaze-assisted tool for reflective piano playing. EyePiano guides the practice process of learning piano scores through analyzing the pianist's gaze behavior. We based the design of EyePiano on requirements identified through interviews with piano teachers and a feasibility evaluation of gaze metrics. Our system illustrates that basic gaze metrics are sufficient to predict difficult regions for students. Thus, highlighting sections of the music piece which are particularly difficult for the pianist allows EyePiano to support piano rehearsals for students. Our work showcases the feasibility of using gaze data for reflective music education, enabling effective instrument practice.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**.

KEYWORDS

reflective learning, gaze, piano, proficiency

ACM Reference Format:
Jakob Karolus, Johannes Sylupp, Albrecht Schmidt, and Paweł W. Woźniak. 2023. EyePiano: Leveraging Gaze For Reflective Piano Learning. In *Designing Interactive Systems Conference (DIS '23), July 10–14, 2023, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3563657.3596065

DIS '23, July 10–14, 2023, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-9893-0/23/07.
https://doi.org/10.1145/3563657.3596065

1 INTRODUCTION

Being able to develop skills based on one's aptitudes is central to a fulfilling life, and a key component in finding meaning [47]. Traditionally, skills are acquired through organized education staffed by professional teachers. Technological developments offer new possibilities and advances in artificial intelligence [21] allow us to create intelligent systems that are able to support users in learning new skills. Such intelligent tutoring systems (ITSs) have long piqued the interest of human-computer interaction (HCI), exploring their potential for a variety of different domains, such as sketching [59], programming [28], and machine tasks [23]. While these examples show the feasibility of using computational methods for supporting skill development, it still remains a challenge for HCI to harness the power of computational performance assessment to develop effective learning support strategies. This is particularly relevant for domains where creativity is required, such as playing music.

The HCI field has an established interest in supporting creativity and, specifically, helping users play instruments with more enjoyment and develop skills. Work by Rogers et al. [46] used projections to help users improve their piano playing, a common approach for piano tutoring systems [7, 45]. Piano Genie [11] allowed users with no piano skills to improvise piano pieces. Chiang and Sun [7] developed a portable system for assisting piano play on the go. While these systems show that interactive assistance in playing the piano can improve playing performance, they do not enable the user to understand their performance and reflect on their skills, which is particularly valuable for advanced players. Our work is interestingly different from past research on interactive piano tutoring as it demonstrates that a data-driven approach can be used to identify particularly difficult parts of a piece. By doing so, the system enables reflective learning, i.e., facilitates understanding performance and promotes sensemaking.

In this work, we investigate the potential of leveraging gaze data for reflective piano learning, enriching existing learning concepts for musical education. In this domain, autodidactic approaches and digital alternatives to personal teaching can play a key role in making music accessible to a wider audience and enable more opportunities to practice and play, e.g., through remote sessions. While there already exists a number of systems supporting different kind of instruments, such as the guitar [26, 34], the violin [25], and the piano [46, 62], we envision our work not as a strict tutoring system — possibly substituting a personal teacher — but rather as a tool that provides pianists with the means for reflective learning by guiding their practice routine. Our goal is to understand how we can design systems that seamlessly integrate into existing learning concepts but also allow for new opportunities of autonomous learning.

To this end, we leverage the pianist's gaze data to infer their current practice progress. Gaze as a modality provides us with an unobtrusive way to monitor students and offers cognitive insights (see [62]) about their play, rather than just technical proficiency, i.e., correct play. We developed EyePiano—a gaze-assisted tool for reflective piano learning—in a systematic process of identifying design and technical requirements through interviews with piano teachers and a gaze feasibility study. The system offers a gaze-assisted feedback module for score learning, integrated into a guided practice routine. We identified access to an auditory gold standard and feedback customization as key features. In an evaluation of EyePiano, we found that participants approved of the reflective feedback of the gaze-based algorithm, which offered support for working on their individual weaknesses. EyePiano allowed them to focus on the challenging parts of a score, facilitating a better understanding of their own learning process.

In our work, we highlight the potential of data-driven systems for reflective piano learning. We contribute EyePiano's complete design process and its evaluation. Based on our findings, we derive implications for future systems, including key functionalities and adapted learning concepts for musical education.

2 RELATED WORK

In our work, we draw from findings in the domain of gaze analysis as well as from existing research work on intelligent tutoring systems. In the following, we provide an overview for both of these domains, how our investigation has benefited from them, and how they come together in our work on EyePiano.

2.1 Gaze and Music

Seminal works by Buswell [5] and Yarbus [61] have first indicated a relationship between gaze behavior and high-level cognitive processes. Since then researchers have connected eye movements to user activities and respective skills for a vast variety of scenarios, e.g., reading [43], personal interest [57], or language proficiency [27]. For a more extensive overview, in particular for gaze metrics in HCI, we refer the reader to review works by Jacob and Karn [24] and Duchowski [13].

Reading and playing music score notation, also known as sight-reading, is more cognitively demanding than text reading. Aggravating factors include the complex notation, appropriate transfer to motor commands, and the musician's head and body movements [44, 52]. Although there is a generally agreed on movement pattern [35] when reading scores, non-linear movements such as refixating on already processed score parts are challenging [42]. On the one side, characteristics of the score notation influence the gaze behavior of the player. Here, research has identified, among others, rhythm [40], tempo [51], note length [29, 40], structure [6, 58], and genre [58] to have an effect on the musician's gaze behavior. On the other side, the musician's proficiency plays a vital role in how they process score notation. Here, the ability to sight-read has been in the focus of research works. Generally, more skilled sight-readers fixate for shorter periods on average and read further ahead (greater eye-hand span) [35, 40], suggesting a faster music comprehension. Similarly, proficient sight-readers exhibit fewer overall fixations and fewer refixations on already seen parts of a score [4]. This multi-faceted nature makes it challenging to identify universal gaze characteristics of a musician's proficiency.

Consequently, in EyePiano we opted for a more feedback-focused approach. While both score difficulty and the musician's proficiency moderate the exhibited gaze pattern, it is irrelevant which aspect is responsible. In other words, for EyePiano it does not matter if the student struggles due to the high score difficulty or to their low playing proficiency. The resulting outcome that the student needs support from the system will be the same, and can be inferred from their gaze data. We adopt this approach to foster independent reflective learning [9, 50]. In HCI, reflection is considered a key design goal for system design to foster one's understanding of oneself or self-improvement [1, 2]. While past work primarily explored designing for reflection in domains where performance is readily quantifiable (e.g., physical activity [48] or smartphone use [19]), we designed a system that supports reflection in the creative, open-ended activity of piano play.

2.2 Intelligent (Piano) Tutoring Systems

The goal of intelligent tutoring systems (ITS) is to provide a more student-centered approach to facilitate learning. While being closely connected to the metaphor of human teachers in the beginning, ITSs nowadays often make use of artificial intelligence [8, 37]. Being able to manipulate the problem-solving environment [8] to create a situational awareness in students, e.g., with contextual illustrations [31] or interactive simulations [18], can significantly increase the learning performance of students.

Analogously, tutoring systems in musical education have applied the same concept. Interactive systems directly communicate and assess the learning process for a variety of music fields. For example, making use of augmented reality like guitAR [34] to project chord and note sequences on the guitar fretboard, leveraging sensing technologies to detect correct finger postures of guitarists in EM-Guitar [26], or to provide instant vibrotactile feedback for violinists on their posture [25].

In particular, the piano has been of great interest to researchers. Early work by Dannenberg et al. [10] contained a complete one-year curriculum, including video tutorials and low-level real-time feedback on wrong pitch, tempo changes, and interruptions; additionally containing automatic page-turning and audio playback. Similarly, Kitamura and Miura [30] focused on generating suitable exercises and feedback on detected weak points for amateur pianists. Commercial learning apps and websites offer a plethora of lessons, including listening features and progress visualizations, e.g., Playground Sessions[1], Song2See[2], Practice Bird[3], and Synthesia[4]. In EyePiano, we also explicitly focus on providing feedback for the student's weak points. Contrary to the presented systems here,

[1]https://www.playgroundsessions.com/
[2]https://www.songs2see.com/en/products/game/
[3]https://phonicscore.com/
[4]https://synthesiagame.com/

we do not use MIDI or audio as input for our detection algorithm. Instead, EyePiano leverages the student's gaze data to detect challenging parts of a score, allowing us to capture cognitive struggles rather than just wrong play.

Tutoring systems for the piano often make use of simplified learning methods, lowering the entry barrier for novices, in particular when it comes to sight-reading. Here, interactive projections play a vital role in communicating real-time feedback [45]. P.I.A.N.O [46] makes use of an interactive projection to display a rolling notation of the score, removing the need for sight-reading. Similar works (see [7, 22]) provide virtual finger postures for practicing the piano. Other feedback methods include tactile feedback supporting pianists in acquiring the necessary motor skills [12, 33]. While these interactive systems may allow users to improve their playing skills through novel learning methods, even improvising without prior piano skills [11], EyePiano—in contrast—focuses on delivering reflective feedback, providing the users with the means to improve their own learning process.

Instead of inferring play performance through detecting correct play, through MIDI and audio input, or even vision-based finger detection, a seminal work by Yuksel et al. [62] showed that it was possible to build tutoring systems that directly react to the pianist's cognitive workload. Using functional near-infrared spectroscopy (fNIRS), BACh [62] is able to adjust piece difficulty dynamically during play, always providing the perfect difficulty level for the pianist. Our work with EyePiano draws from this approach, identifying cognitively challenging score parts for the pianist rather than detecting technical proficiency, i.e., correct play. By doing so, we can identify not only parts that were played incorrectly but also parts that sound right but should be practiced nonetheless as identified by the player's cognitive demand. In contrast to BACh, our system leverages gaze data, allowing deployment in a variety of practice environments more easily.

3 METHODOLOGY

While we draw from existing designs of piano tutoring systems, we rather envision EyePiano as a tool for *reflective piano playing*. Such systems should allow pianists to reflect on and improve their own individual learning process by providing them with the means (potentially validated through data) to understand it. As such, EyePiano is mainly tailored towards advanced piano players. To that end, our system offers a data-driven approach to piano learning which aims to facilitate understanding performance and promotes sensemaking.

To allow for a holistic investigation (see Figure 1), we first interviewed experienced piano teachers to identify *design requirements* on how interactive systems could support score learning for pianists and integrate into existing musical education. In a second step, we confirmed the *technical feasibility* of gaze metrics to estimate piano playing proficiency by analyzing the gaze data of participants during piano play. This allowed us to derive implications for the design and implementation of EyePiano. Finally, we evaluated the system in a user study with advanced pianists, investigated the effectiveness of reflective piano learning. In particular, we observed whether participants systematically paid attention to their own playing difficulties and, most importantly, how EyePiano facilitated this recall process. This included a quantitative analysis on the amount of difficult bars over the course of the rehearsal with EyePiano as detected by the gaze algorithm. We complement this with a qualitative analysis through interviews, identifying specific aspects of EyePiano that contributed to reflective learning apart from traditional rehearsal. Consequently, we employed a mixed-method evaluation to assess requirements, constraints, and opportunities of reflective piano learning. To guide the individual parts of our research, we formulated three research questions:

RQ1: *What are design requirements and constraints for piano learning tools?* We first explored necessary design requirements and possible constraints for systems that support piano learning. Informed through related work and expert interviews with piano teachers (Section 4), we distilled key features as well as challenges to consider for such systems.

RQ2: *Can we determine the proficiency of a piano player using gaze data?* We addressed this research question in our gaze feasibility evaluation (Section 5). In conjunction with findings in related work, we hypothesized that piano players exhibit distinct gaze patterns that are influenced by, among others, their own playing proficiency and the current score difficulty. Since eye movements of beginner players are more volatile (see Section 2), we focus our investigation on players of advanced skill level. Here, we expect gaze patterns to generalize better. We evaluated whether it is feasible to implement an algorithm that detects where users experience difficulty in a given piano score. Eventually, this information is used to inform the user's learning process in EyePiano.

RQ3: *What are design implications for gaze-assisted reflective piano learning?* We identified necessary design requirements in our interviews with piano teachers. After assessing their viability in conjunction with our gaze feasibility evaluation, we implemented selected features in EyePiano (Section 6). In a subsequent evaluation (Section 7) of EyePiano, we first confirmed the feasibility of EyePiano for piano rehearsal in terms of usability and user experience. We further evaluated through post-hoc interviews how EyePiano can leverage reflective piano learning by allowing pianists to recall their mistakes through a guided rehearsal routine.

4 DESIGN REQUIREMENTS FOR AN INTERACTIVE TOOL TO SUPPORT PIANO LEARNING

We conducted a series of interviews with experienced piano teachers to identify a first set of requirements (**RQ1**) for the design of an interactive tool that could support pianist by means of reflective piano learning. We additionally inquired about their concept of playing proficiency to gather objective criteria to be later used in our gaze-based proficiency estimation.

4.1 Participants

We recruited three piano teachers with at least five years of experience through word of mouth. All participants were male and aged $\bar{x} = 43.3\,y$ ($s = 21.5\,y$). A remuneration of USD 12 (local equivalent) per hour was provided. The individual profiles of the participants

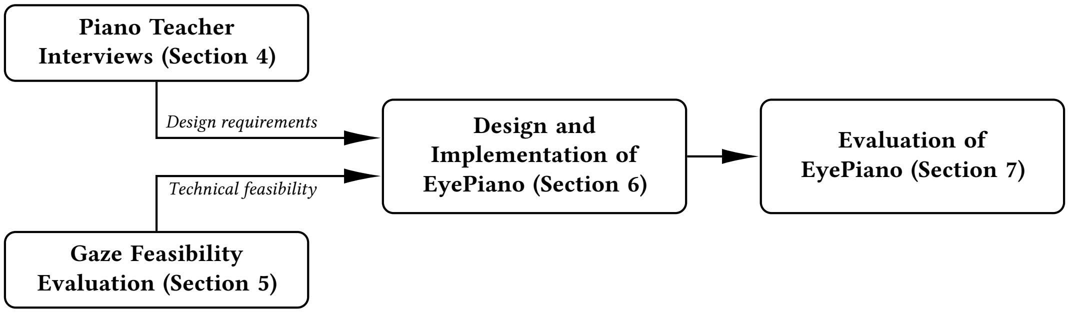

Figure 1: Structure of our holistic investigation, illustrating our initial *design requirements* analysis through piano teacher interviews (Section 4) and initial *technical feasibility* analysis (Section 5). Both these steps then informed the design and implementation of EyePiano (Section 6), which were subsequently evaluated in a final user study (Section 7).

are shown in Table 1, depicting their total years of teaching and average teaching hours per week.

ID	Age	Gender	Profession	Years of teaching	Avg. hours per week
1	38	male	piano teacher	8	15
2	67	male	retired piano teacher	42	20
3	25	male	music student	6	3

Table 1: Participant profiles in our interviews, including information on their teaching experience.

4.2 Interview Script

After the participants consented to the interview, we asked about their experience in teaching piano. We further inquired how they structured and conducted a typical piano teaching lesson as well as how they assessed a student's playing proficiency. To collect insights for our gaze feasibility evaluation (see Section 5), we also asked them how they determined the difficulty of a classical piano piece for their own lessons and had them evaluate three specific pieces later used for our own evaluation. Here, we identified how teachers use objective metrics to determine difficulty, highlighting challenging regions in the process. Finally, we discussed the potential of intelligent tutoring systems from their point of view, including their own experiences and thoughts about the possible limitations.

4.3 Analysis

All interviews were recorded and transcribed verbatim (total duration of 3:17 *h*). To analyze the interviews, we opted for a pragmatic approach to thematic analysis [3], known to be effective for conceptualizing requirements. After merging an initial coding tree based on one interview coded by two researchers, the rest of the interview material was evenly split and analyzed separately. Based on a final iterative discussion, we constructed three themes that describe the requirements discussed by the piano teachers.

4.4 Interview Findings

Our analysis resulted in the following themes: LEARNING AND TEACHING THE PIANO, PIANIST'S PROFICIENCY, INTELLIGENT TUTORING SYSTEMS.

4.4.1 Learning and Teaching the Piano. The piano teachers agreed on intrinsic motivation as a key factor to learn the piano. Here, it is the teacher's task to create new motivations for the students.

> *It is important to consider if I can set a new impulse. (P2)*

Most importantly, the teachers remarked on the existence of different learning strategies for students that need to be curated. For example, balancing an explorative with a more structured approach. Here, EyePiano's approach for reflective learning enables a variety of strategies.

4.4.2 Pianist's Proficiency. The teachers reported on the problem of objectively rating the difficulty of piano pieces and, in turn, assessing the relative proficiency of the pianist. Different interpretations and different ways of approaching a score influence the subjective feeling of difficulty for the pianist. However, a certain set of objective metrics still exists. Among others, rhythm, readability, amount of notes and polyphony, tempo, physical demand, and special techniques are of relevance.

> *(...) the physical, the demand for mobility, for speed, for polyphony, in which they then strike many keys simultaneously. (P2)*

In line with assessing objective elements of a score influencing the pianist's relative (influenced by score difficulty) proficiency and thus current performance, the teachers agreed that proficient pianists should be versatile.

> *(...) and above all, a pianist who is flexible. That is, if I can play not only one direction, but can do something with each direction. (P1)*

Additionally, a certain level of motor mastery is a prerequisite for any aspiring pianist.

> *(...) Virtuosity is always one of those things when they are simply technically very good. (P3)*

However, for a teacher, there will always be some form of subjectiveness involved when assessing a student's proficiency.

4.4.3 Intelligent Tutoring Systems. In a section of the interviews, we asked the piano teachers about their views on IPTSs. While they agreed that IPTSs lack a teacher's competence and would likely suffer from a too mechanical approach to teaching, all three saw the potential in combining such systems with traditional teaching methods — an approach we follow with EyePiano. Here, the possibility for instant feedback and preventing the consolidation of mistakes were highlighted.

> *The tendency is always, "Okay, I will play what I can already do and then I will play a thousand mistakes again" - afterwards (pianists are) just as dumb as before. (P1)*

Requested features were—among others—adjustable feedback for students, highlighting wrongly played notes and tempo, recommending first steps during rehearsal, and a form of gold standard to aim for. We will pick up these key features in our design of EyePiano.

> *Not only show where he had struggled with because maybe he has already noticed it himself. But you might also have a suggestion where he should start again because many people start at the beginning and that is often fatal. (P1)*

4.5 Implications for EyePiano

During our interviews, we identified specific design qualities and limitations for tools to support piano learning (**RQ1**). We summarize these findings below and highlight how they inform our gaze feasibility evaluation (see Section 5) as well as the design of our final prototype EyePiano (see Section 6).

Challenges in Determining Piano Playing Proficiency. Piano teachers often have their own subjective, qualitative metrics to assess a student's proficiency, each emphasizing different aspects (Pianist's Proficiency). The artistic nature of music [14, 63] warrants individual assessment. However, the teachers also reported on objective metrics that can be measured. Among others, these include motor skills, being able to express and interpret music, creativity and improvisation, and the ability to sight-read. For example, sufficient motor skills are a prerequisite to play multiple notes at once and to master large amplitudes and fast tempos (Pianist's Proficiency). Another aspect of proficiency is the ability to sight-read and play music directly from the sheet [52].

Consequently, we will focus on the ability to play by sight as an indicator for relative proficiency in EyePiano. Here, the idea is that piece difficulty and individual piano playing proficiency moderate and impact gaze patterns during sight-reading. We validate this assessment through a technical evaluation regarding the feasibility of gaze metrics to recognize when players are struggling with a specific part of the score (see Section 5).

Strengths and Limitations of Teachers and Intelligent Tutoring Systems. The piano teachers have identified the potential for automated feedback, e.g., on pitch and rhythm, as a major advantage for intelligent tutoring systems (Intelligent Tutoring Systems). The fact that these systems are always available, even between teacher lessons, makes them a valid asset for musical education. Statistics on learning progress allow a collaborative nature, ready to be shared with other students. In theory, this would allow IPTSs to possess a huge amount of accessible learning material. Yet, IPTSs can only provide a narrow view on learning achievements. The teacher, on the other hand, has a much more holistic view of a student's learning process (Learning and Teaching the Piano), able to provide explanations regarding mistakes, e.g., a wrong hand posture. Here, the experience of a teacher is a key element for proper motivation, allowing empathy and long-term success.

We argue that these complementary strengths of the experienced teacher and the data-driven nature of IPTS add value to a new blended way of learning — reflective piano learning (cf. [9]). With EyePiano, we explore the feasibility of this concept and how it integrates into existing concepts for musical education (see Section 7).

Key Features. When asked about potential features and requirements for a gaze-based system to support piano learning, the piano teachers highlighted the following key points. Firstly, the system needs to identify and display mistakes to students and, secondly, provide recommendations of a suitable start to practice (Intelligent Tutoring Systems). We later address these points in EyePiano, through a user-tailored practice routine that makes use of the recorded gaze data to identify difficult regions (see Section 6). Thirdly, feedback needs to be customizable, such as changing the detection sensitivity and allowing users to correct the system. EyePiano conforms to these requirements by implementing three distinct detection levels (see Section 5), also allowing for post-hoc correcting of system markings. Further minor features include a metronome and an overall score analysis.

5 FEASIBILITY EVALUATION OF GAZE METRICS

After drawing insights on proficiency aspects and potential metrics to quantify playing skill from our interviews with piano teachers, we take a closer look at **RQ2**, studying the feasibility of gaze metrics as an indicator for piano playing proficiency. We report on a small-scale dataset of gaze data during regular score play from pianists and subsequent analysis thereof that informed our final algorithm in EyePiano.

5.1 Gaze Data Collection

We employed a repeated measures design, where each pianist performed three different music pieces twice. The three blocks for the different pieces were counterbalanced, resulting in a total of six recorded performances and an additional test trial at the beginning. We captured gaze and MIDI data for post-hoc verification of the participants' markings of difficult parts within the scores. The three musical pieces were chosen from three intermediate levels in accordance with PianoBookGuide [36], IMLSP [41] and Wolters [60]. We chose a sonata by Clementi[5], Gnossienne by Satie[6], and a Prelude of Chopin[7].

[5]https://www.mutopiaproject.org/ftp/ClementiM/O36/sonatina-1/sonatina-1-a4.pdf, Level 5 (see [36]).
[6]https://www.8notes.com/scores/10611.asp, Level 6 (see [36]).
[7]https://www.mutopiaproject.org/ftp/ChopinFF/O28/Chop-28-6/Chop-28-6-a4.pdf, Level 7 (see [36]).

Figure 2: Apparatus showing the MIDI keyboard, two loudspeakers and a monitor with attached eye tracker connected to the stimulus and recording laptop (left).

Figure 3: Example excerpt of a pianist's gaze path on a score highlighting the challenge of non-linear gaze progression. Colors represent temporal progression. Red circles highlight non-linear gaze behavior (refixations to previous bars, gazing upwards) that we address with our algorithm to allow robust mapping of gaze data to respective bars.

5.1.1 Apparatus. The apparatus of the study is depicted in Figure 2. It consisted of an external monitor (2560x1440 px) used to display the score to the pianists. On the bottom rim of the screen, we attached a Tobii 4C eye tracker (90 Hz sampling rate). The participant was seated on a piano stool approximately 70 cm in front of the monitor. We used the full-key MIDI Masterkeyboard Doepfer LMK2+ coupled with the free synthesizer Piano One[8] integrated into the application Waveform[9]. Sound output was provided by two loudspeakers next to the screen. During performances of the participants, both the gaze data stream from the eye tracker as well as the MIDI data from the keyboard were recorded. An additional experimenter PC was connected to the apparatus as well, allowing the experiment to monitor the stimulus display's output, including the pianist's gaze, and additionally mark score regions that they identified as challenging for the participant. The experimenter and participant shared the same room separated by a distance of 2.5 m, including a physical separation through a perspex wall[10].

5.1.2 Procedure. After participants provided informed consent, we asked them about their demographics and calibrated the eye tracker. The calibration was accepted for a deviation level of ≤ 2 deg visual angle and additionally validated during the study and at its conclusion. Participants then played through the provided music pieces, activating each score by themselves by looking at a play button for two seconds. Page turning was handled automatically by the prototype, updating played half pages consecutively. To gather ground truth data[11] for the later gaze analysis, participants self-reported difficult passages by highlighting them on the screen immediately after each play-through. To add reliability, these markings were validated[12] by the experimenter based on the observed gaze data and musical performance. The study lasted approximately 60 minutes, and participants were compensated with the equivalent of USD 12 per hour. Ethical approval for this study was obtained from the Ethics Committee at the University of Constance.

5.1.3 Participants. We recruited six participants (4 female, 2 male; Age: $\bar{x} = 29\,y, s = 14\,y$) through personal contacts and participation in related studies. All participants were advanced pianists with on average $\bar{x} = 20\,y$ ($s = 16\,y$) of experience, being able to fairly play the Sonatine I (Op. 20 no 1-1) by Kuhlau[13]. Their sight-reading experience (7-item likert) was $\bar{x} = 5.8 (s = 0.8)$. None of them used digital notes regularly.

5.2 Gaze-Based Algorithm

Based on the collected gaze data, we implemented a robust detection algorithm for regions, in particular individual bars, that our participants found difficult to play. This metric serves as an indicator of their actual piano playing proficiency. Less proficient players will struggle with difficult bars, exhibiting characteristics gaze patterns during these sections (see Section 2). After eye event detection based on an I-DT algorithm [49], we further processed the gaze events to accommodate non-linear score reading behaviors [42], as detailed in the following sections. An example is detailed in Figure 3.

5.2.1 Line Detection. The first essential step is the correct allocation of gaze events to corresponding staff lines on the score sheet. Recorded gaze data do not necessarily fall within the boundary boxes of individual lines, e.g., due to body movements or mind wandering [52]. The main idea behind this processing step is based on including a "carriage return" detection of the collected gaze data, hence allowing a temporal allocation of gaze events to staff lines. The algorithm uses an adaptive threshold based on the initial 25th and 75th percentiles of the x coordinates of the fixation data. Additional safety checks on the y coordinates ensure robustness.

5.2.2 Fixation to Bar Mapping. Finally, a mapping of fixations to respective bars in the score notation was realized through constraining positional and temporal location of each fixation. Thus, a fixation might be assigned a different bar than its x/y coordinates

[8]https://neovst.com/piano-one/
[9]https://www.tracktion.com/products/waveform-free
[10]Note: the study was conducted during the COVID-19 pandemic. A full hygiene concept was compiled in compliance with the rules of the university.
[11]Whether or not a current passage is difficult for the player.
[12]The experimenter may adjust or delete participants' markings or add new ones. Changes were verified with the participants.
[13]https://www.mutopiaproject.org/ftp/KuhlauF/O20/sonatine-1-allegro/sonatine-1-allegro-a4.pdf, Level 5 (see [36]).

based on its temporal data. First, this allowed us to calculate the exact time a pianist spent looking at a specific bar, which we later used for an outlier detection algorithm. Secondly, bar-level detection of difficult regions is vital for appropriate feedback, as suggested in our interview with piano teachers.

In particular, we identified refixations of already read parts, gaze data outside of staff lines, gaze wandering, bar repetition or skipping, and crossing between bars at their borders as problematic cases where gaze events could be attributed to the wrong bar (see Figure 3). After initially allocating gaze events to staff lines, further clustering based on their location and time allowed us to discard outliers and to properly handle repetitions. Employed thresholds for the algorithms are based on the average bar duration for each individual participant due to the observed high variance across participants in terms of playing time and fixation patterns.

5.2.3 Classification of Difficult Bars. To cater to individual practice needs by students as suggested by the piano teachers, we opted to implement three sensitivity levels within our detection algorithm: a user-dependent machine learning algorithm (MLA), a simple outlier detection (SOD) based on average bar dwell time, and a combination of the two methods. Our classification algorithm will detect whether a given bar was difficult to play for a given pianist.

Fixations and associated bar mappings are used as initial input for the MLA. We identified that no single gaze feature was sufficient for a robust classification of difficult regions. Thus, we constructed a set of gaze features and optimized parameters by reviewing the resulting confusion matrices. Noteworthy features of our final set included: normalized refixation count, dwell time on 60-pixel clusters[14], and normalized horizontal movement between fixations (horizontal saccade length). By dividing the gaze data into several epochs (sliding window approach), our MLA can make use of aggregated gaze metrics, adding robustness. Epoch durations are adaptive to the average bar duration of participants and passes were submitted for 0.25, 0.5, and 1 times the average bar duration. A unanimous vote between all three passes was required to be considered a difficult passage for the player.

To detect the most challenging parts, we developed a simple outlier detection (SOD) that is based on the participants' average dwell time on a specific bar. We empirically validated that bars over one standard deviation longer than the average bar dwell time were troublesome for participants.

In a final algorithm, we combine both SOD and MLA to allow for bar-level estimation of playing difficulty. Depending on the configured sensitivity level, either algorithm takes precedence. On the highest sensitivity level, only the MLA is used, detecting most bars that were challenging for the player, albeit suffering from a higher false positive rate. The medium level only used the SOD, while the lowest sensitivity level combined both MLA and SOD. In this setting, bars are only considered difficult when flagged by both algorithms, ensuring a low false positive rate while still detecting the most challenging parts of a score. An overview showing cross validation results for user-dependent classification (averaged over all users) is given in Figure 4.

The purpose of our gaze-based algorithm lies within the robust detection of most difficult regions, allowing us to investigate new reflective learning paradigms supported by EyePiano (see Section 7). Developing a complex and sophisticated algorithm is not the focus of this work, neither is it necessary to investigate how gaze data can be leveraged for reflective piano learning.

[14]In accordance with [44] and empirically validated on our dataset.

5.3 Implications for EyePiano

Based on our analysis of the recorded gaze data, we have identified several implications for our design of EyePiano, addressing technical requirements (**RQ2**). We highlight these in the next section and refer to respective decisions made for EyePiano.

Gaze Is a Suitable Predictor of Difficult Score Regions. Our analysis highlighted that a set of basic gaze metrics are indicative of how challenging score regions are to play for a pianist. Variances in temporality and locality of gaze data allow us to derive the player's relative playing proficiency. Regions with a higher average refixation count, fixation count, and dwell duration correlate with greater difficulty for the pianist, e.g., due to high score difficulty or their own lack of playing proficiency. In our work, we employed a set of several gaze features to create a robust algorithm, counteracting potential noise artifacts as reported by related work (see [29, 35, 40]). However, our lack of tempo control [42] required us to carefully review the recorded gaze data, implementing fail-safes (repeated bars, slow start). The presented algorithm has been integrated into the final version of EyePiano, including the algorithm based on viewing time (SOD) and the more sophisticated machine learning algorithm (MLA) based on gaze metrics. It provides robust detection of students' difficulties on a bar level. Interviewed piano teachers have confirmed that this granularity is sufficient.

Favor of User-Dependent Difficulty Prediction Algorithms. To aid in robustness, we opted for a user-dependent prediction algorithm in EyePiano. This allows us to tailor the detection more closely to individual differences in eye patterns, such as body movements, overall varying proficiency [4, 52] and genre-specific notation styles influencing rhythm [40], note length [29], or structure [58]. Therefore, employing a user-dependent algorithm allows us to better cater to the individual preferences of pianists. By design, this loss of generality is not an issue for EyePiano. As pianists will typically use the system for a longer period of time, more gaze data is provided naturally, allowing an even better fit for the student. For this use case, it is most beneficial if the algorithm can learn the student's individual characteristics. While this choice results in an initial cold-start problem[15], a generic model can be used at first and trained over time, e.g., by fusing predictions with collected MIDI or audio data on play correctness.

6 EYEPIANO— A GAZE-ASSISTED TOOL FOR REFLECTIVE PIANO PLAYING

After confirming technical requirements, we further explore necessary design requirements for EyePiano. Identified by our interviews with piano teachers (see Section 4) and the results from our gaze feasibility study (see Section 5), we implemented four key features and subsequently evaluated the system in terms of usability and user experience. In concluding interviews, we additionally paid attention to changes and alterations to traditional score learning

[15]There is initially no data to train the algorithm for new users.

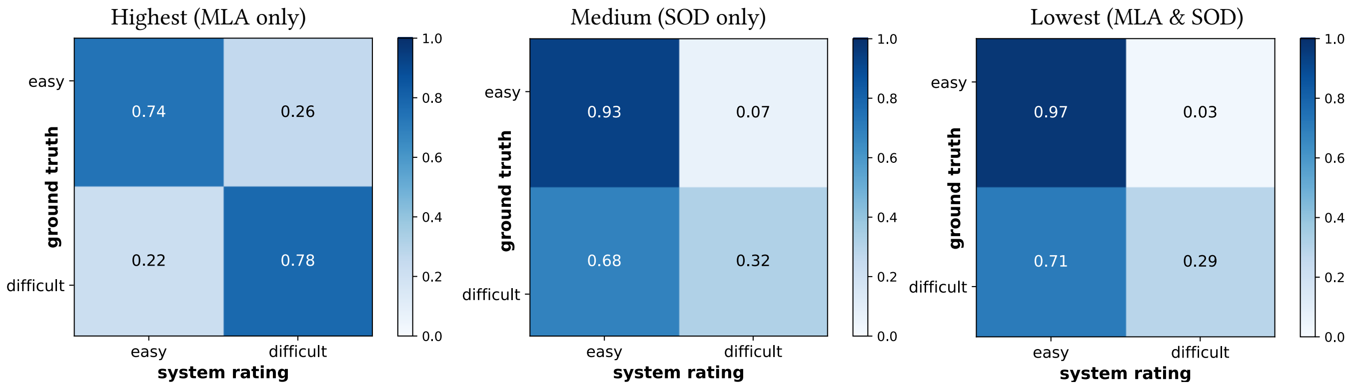

Figure 4: Confusion matrices depicting the accuracy of the final prediction algorithm for difficult passages grouped by three sensitivity levels (from left to right: highest to lowest). There is a total of N=1230 samples with 386 difficult parts.

and how these impacted the learning process (**RQ3**). The key features are detailed in the following. A complete overview of the user interface is shown in Figure 5.

Highlighting difficult bars. After a completed play through, our algorithm identifies difficult bars and highlights them on the score sheet (see the center of Figure 5). Additionally, the pianist has the option to visualize their own gaze data superimposed on the score, both for the whole score and for difficult bars only. The option to correct the system's markings is provided as well.

Adjustable sensitivity levels. Closely coupled to the first feature is the option to adjust the sensitivity level to individual needs, e.g., rigorous rehearsal vs. more lenient practice. Here, EyePiano allows changing between the three sensitivity levels dynamically. The selection is immediately updated and highlighted bars are recalculated. Pianists can thus play around with the different levels and reflect on their performance, i.e., selecting a lower sensitivity if only the most challenging bars are to be practiced.

Gold standard to support learning goals. Piano teachers remarked that students needed a gold standard of how a score sounded to aim for. To provide tangible goals for piano learners, we added an audio playback functionality, allowing them to listen to a gold standard of how to play the specific bar. The playback can be triggered by selecting a bar, e.g., via gaze.

Recommendations for practice. A last requirement was easy and ready access to a rehearsal strategy that would allow players to reflect on their individual weaknesses. EyePiano addresses this need by including its own user-tailored practice procedure, accessible in the training mode (see Figure 5). This simple step-by-step process of (1) selecting a suitable sensitivity based on highlighted bars, (2) listening to the respective playback before (3) rehearsing, guides piano learners through the score and helps them to identify, to reflect on and to systematically rehearse problematic sections. A more detailed description of this practice routine is given in Section 7.3.

In addition, EyePiano implements hands-free interaction and can be controlled via gaze alone. This eliminates the need to navigate via a separate mouse and keyboard, potentially hindering the learning flow. We implemented a standard dwell-time activation to avoid a random gaze selection. A circular progress bar provides feedback for the user.

7 EVALUATION

We evaluated EyePiano in a study employing a repeated measures design consisting of two blocks, where each block required the participants to practice a specific score using the features offered by EyePiano through following the given practice procedure. Each practice of a specific score was repeated twice, yielding three rehearsals per score. Including a test trial, participants played a total of seven times. The difficulty of the two scores corresponded to level five and seven [36], ensuring similar ratings in IMLSP [41] and Wolters [60]. We chose one prelude by Bach[16] and one by Chopin[17]. Note that this design allowed participants to use EyePiano in both blocks, potentially eliciting a richer qualitative feedback [3] and — most importantly — insights into how EyePiano can support the learning process [9]. We opted for this design rather than a comparative approach where quantitative measure would be of limited significance, especially when biased by individual learning strategies.

7.1 Apparatus

We utilized the same apparatus for the evaluation of EyePiano as we did for our gaze feasibility study (see Section 5.1.1), except the EyePiano software. A Tobii 4C remote eye tracker was attached to the stimulus monitor ($2560x1400\,px$) used to display the score during playing and EyePiano's training mode during practice. As EyePiano allows hands-free interaction throughout, we not only used the eye tracker to record gaze data but as an input device for participants as well. We again used the MIDI Masterkeyboard Doepfer LMK2+, external loudspeakers, and PianoOne in conjunction with Waveform for sound output. Seating arrangements and hygienic precautions for the experimenter and participants were the same. The setup is depicted in Figure 6.

[16]https://www.mutopiaproject.org/ftp/BachJS/BWV939/bwv-939/bwv-939-a4.pdf, Level 5
[17]https://www.mutopiaproject.org/ftp/ChopinFF/O28/Chop-28-4/Chop-28-4-a4.pdf, Level 7

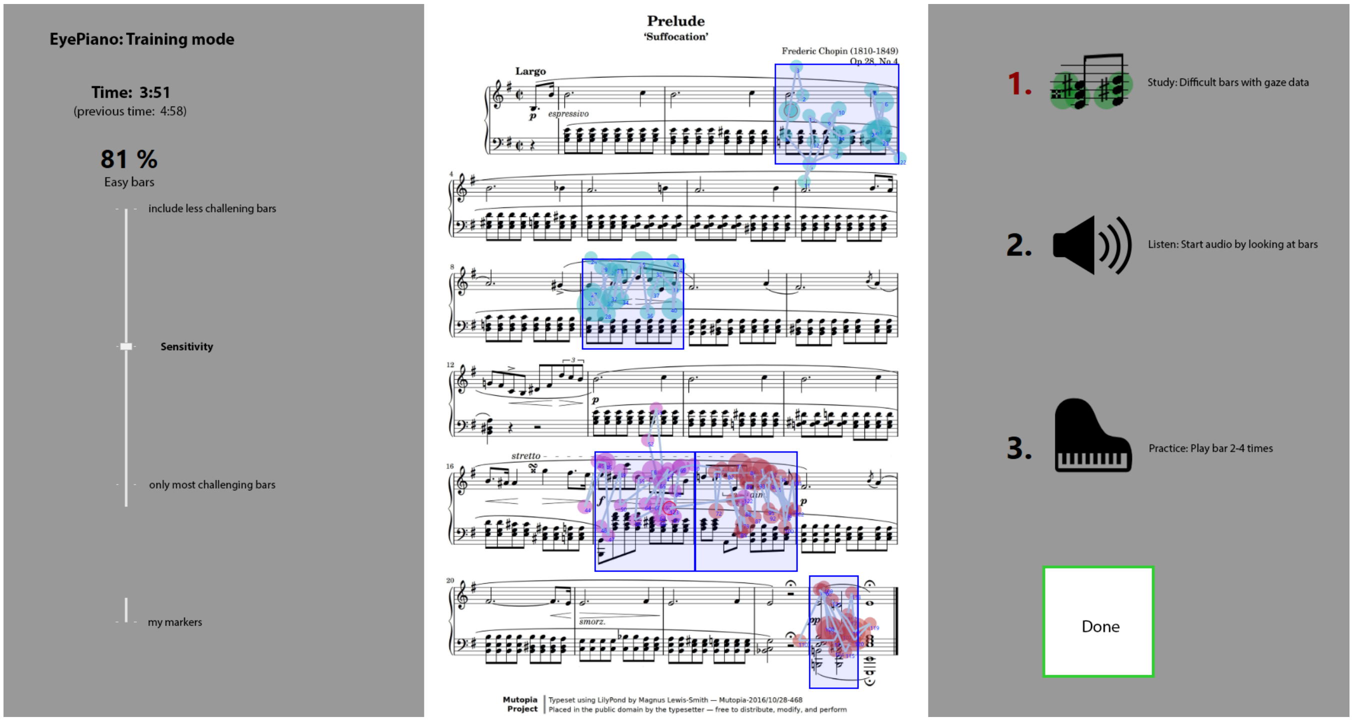

Figure 5: Training mode of EyePiano showing sensitivity settings (left side), highlighted difficult bars (center) and the practice procedure (right side). Step one (choosing an appropriate sensitivity level) is currently selected.

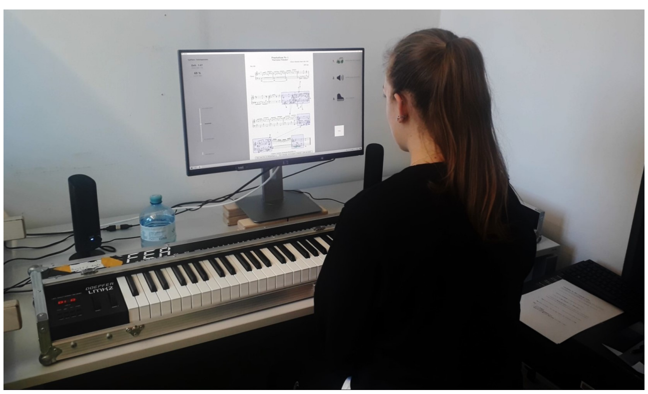

Figure 6: Study setup showing participant during the training mode of EyePiano.

7.2 Participants

We recruited four pianists who had previously participated in our gaze feasibility study (see Section 5). Note that no participant has seen the EyePiano software before. The tool used in our previous gaze feasibility evaluation was used for recording gaze data and displaying note sheets only. As such, participants were not involved in the design process of any of EyePiano's features. The participants' profiles are listed in Table 2. None of them had used digital tutoring systems before. We intentionally recruited participants from our previous study to make use of the person-dependent classifier of EyePiano. Thus, for each participant, we trained the classifier based on their gaze data from the gaze feasibility study. We believe this to be a sufficient sample size for our formative evaluation [56] since all participant exhibit adequate proficiency with the piano but still struggled with the difficult pieces.

7.3 Procedure

First, participants were informed about the study. After participants provided consent, we asked them to fill out a demographics questionnaire. Subsequently, we calibrated the eye tracker. The calibration was accepted for a deviation level of ≤ 2 deg visual angle and additionally validated during the study and at its conclusion. After this initial setup, participants were able to familiarize themselves with EyePiano in a test trial.

Participants then practiced the provided music pieces by first activating each score themselves through looking at a play button for two seconds. Analogously to the gaze feasibility study, page turning was handled automatically. Participants engaged in EyePiano's training mode after each play-through following an abstract practice routine as suggested by the piano teachers. The routine was provided to them on a sheet of paper and additionally implemented within EyePiano:

(1) Study the prediction results by EyePiano. Look at the averaged play statistics and select an appropriate sensitivity level. This step allowed the participants to reflect on their performance.

ID	Age	Gender	Years of playing	Playing frequency	General playing	Sight-reading	Technical	Favorite genres
1	20	f	14	1 / week	5	6	5	classical music, jazz
2	28	m	11	1 / week	5	6	4	classical music, jazz, pop
3	23	f	15	1 / week	5	5	5	classical music, film music
4	23	f	7	1 / month	5	6	4	classical music

Table 2: Demographics of the participants in EyePiano's evaluation. All participants took part in our previous gaze feasibility study. Proficiency in general playing, sight-reading, and technical skills were rated on a 7 point likert scale.

Participants' impressions of EyePiano	
Q1	The system accurately assessed my performance.
Q2	The practice routine was understandable.
Q3	The practice routine was helpful for me.
Q4	EyePiano helped me with learning the pieces.
Q5	I have used the feedback of EyePiano to improve my play.
Q6	The feedback of EyePiano changed how I addressed my rehearsal.
Q7	I have gained new insights on challenging parts through EyePiano.

Table 3: Additional questions on the participants' impressions of EyePiano: *strongly disagree* to *strongly agree*; all visual analog scale (0 to 20).

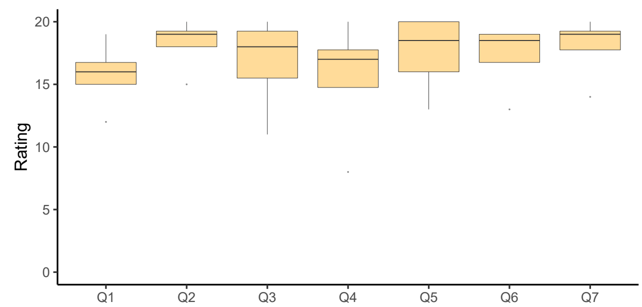

Figure 7: Additional custom questions assessing various features of EyePiano (all visual analog scale: 0 to 20). Please refer to Table 3 for the complete list of identifiers. Overall, EyePiano showed high accuracy (Q1), provided a helpful practice routine (Q2, Q3), and offered useful feedback (Q4, Q5) that influenced participants' practice (Q6) as well as offering new insights (Q7).

(2) Listen to the marked bars that were detected as being difficult for you by EyePiano.
(3) Practice each marked bar at least two times. Included tips: play slowly, practice each hand separately, play the previous bar as well.

We concluded the study with the UMUX [17] and flow-short-scale [15] questionnaire, assessing usability and perceived flow. An additional questionnaire on specific features of EyePiano (see Table 3) assessed the perceived accuracy of the system and its practice routine. A final semi-structured interview revealed additional insights on the potential and further challenges of reflective piano learning and EyePiano in particular. The study lasted approximately 60 minutes, and participants were compensated with the equivalent of USD 12 per hour. Ethical approval for this study was obtained from the Ethics Committee at the University of Constance.

7.4 Results

In this section, we report results on our evaluation of EyePiano, such as its user experience as reported by participants and qualitative insights from our post-study interviews. As we have already confirmed the technical feasibility of our gaze algorithm, we omit a detailed analysis here, instead focusing on the qualitative feedback from the users' impressions of the system's accuracy.

7.4.1 Usability, Feedback Assessment, and Accuracy of EyePiano. We administered two post-study questionnaires: UMUX [17] ($\bar{x}$ = 76.1, s = 13.8) and Flow-Short-Scale [15] (Flow: $\bar{x}$ = 4.8, s = 0.8, Worry (low is better): $\bar{x}$ = 2.2, s = 1.4). The results suggest no major usability issues. Our additional custom questions (see Figure 7) show that participants gave a high rating for the perceived accuracy of EyePiano (Q1). Additionally, the user-tailored practice routine was deemed understandable (Q2) and helpful (Q3). The system helped pianists to learn new pieces (Q4), whereby users made extensive use of EyePiano's feedback (Q5) and adapted their rehearsal accordingly (Q6). Finally, participants rated EyePiano's capabilities to facilitate new insights about their own weaknesses as very high (Q7).

7.4.2 Effects of Rehearsal. To evaluate whether the focused rehearsal impacted play performance in the participants' training sessions, we fitted a linear mixed model using the percentage of difficult bars (as predicted by our gaze algorithm) as a dependent variable. We submitted the rehearsal count as fixed effect[18] and added the participant ID as a random effect. We then tested the fitted model against a null model (without the rehearsal count as fixed effect) and found a significant difference ($\chi^2(1) = 8.76, p < 0.05$). Residuals plots did not reveal any deviations from homoscedasticity or normality. This result confirms that EyePiano's practice routine is effective and does not impede the learning process. Note that compared to traditional rehearsal, EyePiano employs a more focused rehearsal on difficult passages, guiding the user. In the post-study interviews, we further investigate what features of EyePiano were especially helpful for the pianists, compared to traditional rehearsal.

7.4.3 Interviews. We analyzed the post-study interviews (approx. 15 *min* per participant) using the pragmatic approach as detailed

[18]We additionally evaluated models using the played score as a fixed effect (and their interaction effects). Effects are analogous, which is why we only report the most simple model.

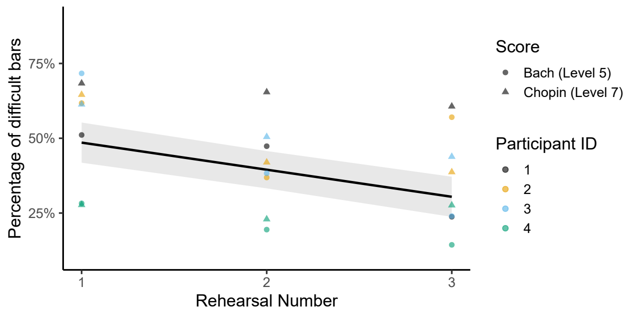

Figure 8: Linear mixed model fitted to assess whether our rehearsal routine had an effect on the amount of difficult bars. The solid black line resembles the fitted model (see Section 7.4.2); the gray shaded corridor marks the standard error. Additionally, data points from individual participants are illustrated, grouped by score.

by Blandford et al. [3]. Two researchers were involved in this analysis. Intermediate coding tree construction and final merge followed the method as employed in our other studies. We constructed four themes for the transcript data: Proficiency Insights, Goal-Driven Learning, Tutoring and User Experience. We provide insights for each theme below.

Proficiency Insights. Participants stated that EyePiano was especially useful in providing guidance during practice. It supported them in remembering difficult bars that would require more attention.

> *(...) just as a helpful reminder to know exactly where to start again. (P1)*

Similarly, EyePiano highlighted unnoticed sections and allowed the pianists to adjust their training. Here, the system elicited new insights for some participants that only recalled the most difficult bars.

> *So I find it totally helpful that it also shows regions I would not have noticed. (P4)*

Subsequently, EyePiano curated their practice through focused rehearsal — a key aspect as indicated by the piano teachers — by arranging a training sequence for students depending on the chosen sensitivity level.

> *I found it really helpful because you can focus exactly on the parts you found harder. (P1)*

Goal-Driven Learning. Piano teachers and related works [8, 39] have highlighted the importance of proper motivation during learning. To provide tangible and achievable goals while practicing, EyePiano incorporates a playback functionality for each score. Participants reported this feature to be very helpful, especially if scores were unknown.

> *I find listening to it useful. It helps me. Especially with pieces that you don't know. (P3)*

Additionally, the playback facilitated the recognition of mistakes. For the students, it was often easier to identify their mistakes through auditory comparison rather than studying the score sheet.

> *You may not even notice that you played incorrectly, but then you hear how it should sound. Then you might notice directly when you made mistakes. (P4)*

Tutoring. EyePiano adheres to a practice routine that adapts to the proficiency of the individual student based on their recorded gaze data. Albeit not as powerful and motivating as a real teacher, this feature allowed us to curate the learning process for each participant. Feedback on this feature was divided. While participants concordantly agreed on the usefulness of this feature, some participants expressed that it was still rather abstract. Nevertheless, interviewees especially saw the benefits for beginners and amateurs, offering them a more structured learning process.

> *There are always people who do not have their own strategy, for them being pointed to it is definitely a good thing. (P1)*

Participants also remarked on the potential of having a digital footprint of one's rehearsal and saw the opportunity of sharing results with the teacher and other students, allowing for competition among students as well as providing teachers with an in-depth look at a student's progress.

> *(So I could upload my score) to the cloud or share it with my student. (P2)*

User Experience. Feedback on EyePiano's gaze interaction was split. Increased demand due to its unfamiliarity was one of the stated reasons. Participants suggested touch- and gesture-based interaction as alternatives.

> *The handling was actually quite simple, but I find it a little exhausting using my eyes. (P4)*

The user interface was clearly structured and easy to use, but participants disliked that piano learning with EyePiano was limited to score learning. Other techniques such as playing by chords or improvising were missing, though all of them agreed that the functionality to support score learning was realized well and were positively surprised by the accuracy of the system in detecting difficult bars.

> *I think that among the difficult ones, what the system indicated and what I marked myself mostly matched, with a few exceptions. (P2)*

7.5 Summary

In our evaluation of EyePiano, we investigated the system's usability and user experience, particularly focusing on whether our previously identified requirements for reflective piano learning (**RQ3**) were fulfilled.

User Experience. Participants reported predominantly positive impressions of EyePiano's usability. The system is clearly structured and supportive, as indicated in questionnaires and interviews. The gaze interaction in EyePiano received mixed reviews, warranting alternative gaze selection modes [55] or alternative modalities

(User Experience). We note that EyePiano does support mouse and keyboard interaction.

EyePiano's Detection Algorithm Is Sufficiently Accurate. While our initial accuracy results were modest (see Section 5), our post-study interviews (see Section 7.4.3) and questionnaires (see Section 7.4.1) confirmed that the employed detection algorithm provides accurate recognition of bars that are challenging for the pianist. Often the lowest sensitivity was sufficient to allow a guided rehearsal within EyePiano's practice routine. We attribute this to the user-dependent classification enabling us to capture the individual characteristics of each participant and the low false-positive rate of the lower settings, as it allowed EyePiano to focus on the most challenging parts.

EyePiano Facilitates Score Learning. The feedback of EyePiano allowed participants to recall difficult bars and facilitated their score learning (see Sections 7.4.1 and 7.4.2). We confirmed in the interviews that pianists perceived the algorithm as accurate enough to support their rehearsal and appreciated the audio playback to allow for more effective training. In particular, EyePiano provides them with new insights about their own weaknesses, highlighting overlooked bars (Proficiency Insights). Through an effective rehearsal routine, EyePiano provides a systematic approach to score learning (Tutoring). Further, participants remarked that especially the option to select different sensitivity levels supported them in eliciting useful feedback (Proficiency Insights). Lastly, the audio playback supported goal-driven learning and prevented the consolidating of mistakes (Goal-Driven Learning).

8 DISCUSSION

Over the course of our investigation, we identified four key features that support reflective piano learning: **highlighting difficult bars**, **choice of sensitivity**, **access to a gold standard**, and a **user-tailored rehearsal routine**. We implemented and evaluated these features with EyePiano (see Section 7.5 for details). Based on our findings, we provide a set of implications and guidelines for future systems in the following, and discuss limitations of the current system.

8.1 Gaze Is an Indicator for Playing Proficiency

Our gaze feasibility study and our final evaluation of EyePiano confirmed the suitability of basic gaze metrics as an indicator for piano playing proficiency (**RQ2**). We have used the ability to sight-read scores as a proxy [20], which limits this approach to learning new scores. Our work suggests that a complementary teaching method including both teachers and automated systems allows for long-lasting learning success (**RQ1**, **RQ3**). Interestingly, our approach shows opportunities for other forms of musical education as well, allowing the ability to sight-read to become an indicator of playing proficiency (**RQ3**). This approach could be especially useful for proficiency assessment for cases where analyzing play performance is difficult for machines, such as polyphonic or non-MIDI instruments.

Though this approach has great potential, confound variables such as varying levels of play proficiency, or personalized play styles are likely to introduce noise and biases in the gaze data (cf. Section 2.1). EyePiano in its current form is tailored towards advanced players, limiting inter-user variance but also its scope of application. Future work is needed to identify algorithm adjustments for novice and expert pianists to allow a more general model. However, small-scale user-dependent models are a valid alternative as we have shown in this work. They require little training data and are able to cope with individual user variances. Future integration of other proficiency metrics, as discussed in Section 4 can help strengthen the robustness of the algorithm.

Our work shows engineering requirements for future gaze-based systems for piano learning. We highlighted the suitability of user-dependent classification algorithms, provided choices for sensitivity levels, and assessed the necessary locality (bar-level) for effective rehearsal [40]. Our choices were driven by empirical findings as well as by the design rationale to facilitate reflective piano learning, **making gaze a robust and suitable indicator of piano playing proficiency** (**RQ2**).

While deploying a gaze-based tutoring system such as EyePiano as described in this work is currently out of scope for most home users, there exists excellent solutions for camera-based gaze estimation using webcams, possibly even integrated in digital notebooks used to display the score sheet. Our call for user-dependent gaze detection complements this scenario, as little training data would be required to train the gaze detection algorithm. A single piano lesson with a teacher would be sufficient to train a sufficient model to support EyePiano's features.

8.2 EyePiano Facilitates Reflective Piano Learning

Our findings confirm that the user-tailored practice routine as provided by EyePiano was appealing to our participants and supported their learning process. We first confirmed that EyePiano's rehearsal routine provided positive learning benefits (cf. Section 7.4.2), likewise to traditional rehearsal routines. We further explored how the individual features of EyePiano supported the pianists in interviews (cf. Section 7.4.3). We found that it provided them with the means to reflect on their own learning process and subsequently improve their play. Here, we identified **the interplay of the key features** as essential (**RQ3**), as participants reported in our interviews. First, **highlighting difficult bars** allowed participants to recall and to reflect on challenging, often overlooked parts of a score, providing them with a starting point for their practice. The fact that EyePiano is based on objective measurements (difficult bars as detected by our gaze algorithm) supported evidence-based learning. Secondly, ready access to **different sensitivity levels** enabled the pianists to further adapt their own individual practice, such as focusing only on the most challenging parts. Thirdly, an **auditory playback** complemented the rehearsal process. Being able to listen to a gold standard for a particular piece facilitated goal-driven learning and mitigated the consolidation of mistakes. Lastly, all features are incorporated into **EyePiano's user-tailored rehearsal routine** that guides the learning process and supports a more systematic rehearsal of music pieces (**RQ1**, **RQ3**). Finally, we observe how EyePiano features a property noted previously as beneficial in computer-based learning systems [38]—providing the learner with the ability to self-correct. This feature was previously identified as important for independent music learning [32, 53].

8.3 Reflective Feedback Through Data-Driven Systems Enriches Musical Education

In this work, we highlight the potential of reflective piano learning for musical education. While this paradigm is becoming popular in other domains, predominantly in sports [16], musical education to date largely focuses on traditional teaching methods. We show how new methods for data collection and analysis can help build intelligent systems which provide an in-depth look into one's proficiency. However, we do not expect these systems to substitute personal teachers (**RQ1**). On the contrary, their capabilities can **enhance existing musical education where a mutual enrichment between teacher and system allows both to play to their individual strengths** (**RQ3**), e.g., the system provides tangible evidence for the teacher to curate the learning process. Research has already shown that having access to gaze data may potentially be beneficial to improve musical skills [54]. The ability to sight-read can thus serve as a crude proxy for a musician's proficiency and generalizes well across multiple instruments. Consequently, systematic reflection and guided rehearsal as implemented in EyePiano are applicable throughout a wide range of instruments in musical education.

In this new paradigm, the teacher provides the holistic approach to musical education as before, while EyePiano facilitates data-driven support for reflective piano learning. A traditional teaching scenario includes weekly sessions with a teacher. The teachers selects appropriate scores for the students to rehearse until the next lesson, and checks the learning progress at the beginning of each lesson. Here, we see the potential of EyePiano to monitor the learning process of students continuously throughout the week and to guide rehearsal, avoiding consolidation of mistakes between sessions. Moreover, the teacher has the option to review collected data and get a better understanding of their student's learning process, curating the learning progress more efficiently through a different selection of pieces or even altering their teaching methods. Thus, integrating EyePiano's rehearsal into musical education **broadens the potential audience (teacher and student)** and their **available opportunities** (**RQ3**).

9 CONCLUSION

In this paper, we described the design, and evaluation of EyePiano—a gaze-assisted tool for reflective piano learning. Our investigation into the potential of data-driven systems has contributed design implications and opportunities for reflective piano learning. Informed by findings from our piano teacher interviews and from our gaze feasibility evaluation, we identified and implemented four key features in EyePiano: highlighting difficult bars, choice of sensitivity, access to a gold standard, and a guided rehearsal routine. Our work shows that a data-driven system for reflective piano learning is not only feasible but also positively impacts the learning process.

We confirmed that EyePiano facilitated reflection on the learning process of piano scores through the interplay of these key features, initiating a potential paradigm shift in musical education, where both personal teachers and data-driven systems can play to their strength. We envision that this collaborative nature is not limited to just the student and their teacher. It will enable a variety of practice opportunities for musical education through readily available teaching, not limited to traditional in-person practice forms. Access to a wide range of students and their individual learning progress may allow us to build powerful systems that support reflective piano learning in the future.

ACKNOWLEDGMENTS

This research is supported by the European Union's Horizon 2020 Programme under ERCEA grant no. 683008 AMPLIFY and 952026 HumanE-AI-Net.

REFERENCES

[1] Eric P.S. Baumer. 2015. Reflective Informatics: Conceptual Dimensions for Designing Technologies of Reflection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 585–594. https://doi.org/10.1145/2702123.2702234

[2] Marit Bentvelzen, Paweł W. Woźniak, Pia S.F. Herbes, Evropi Stefanidi, and Jasmin Niess. 2022. Revisiting Reflection in HCI: Four Design Resources for Technologies That Support Reflection. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 1, Article 2 (mar 2022), 27 pages. https://doi.org/10.1145/3517233

[3] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. Qualitative HCI Research: Going behind the Scenes. *Synthesis lectures on human-centered informatics* 9, 1 (2016), 1–115.

[4] Diana Blom, Patricia Arthur, and Sieu Khuu. 2016. Music Sight-Reading Expertise, Visually Disrupted Score and Eye Movements. https://doi.org/10.16910/JEMR.9.7.1

[5] Guy Thomas Buswell. 1935. *How People Look at Pictures.* Chicago: University of Chicago Press.

[6] Zsuzsa Buzas, János Steklács, Damien Sagrillo, Ágnes Maródi, and Iván Devosa. 2016. Testing Music Reading with Eye Tracking in Three European Countries. (March 2016). https://orbilu.uni.lu/handle/10993/28797

[7] Pei-Ying Chiang and Chung-Hsuan Sun. 2015. Oncall Piano Sensei: Portable AR Piano Training System. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction* (Los Angeles, California, USA) *(SUI '15)*. Association for Computing Machinery, New York, NY, USA, 134. https://doi.org/10.1145/2788940.2794353

[8] Albert T. Corbett, Kenneth R. Koedinger, and John R. Anderson. 1997. Chapter 37 - Intelligent Tutoring Systems. In *Handbook of Human-Computer Interaction (Second Edition)*, Marting G. Helander, Thomas K. Landauer, and Prasad V. Prabhu (Eds.). North-Holland, Amsterdam, 849–874. https://doi.org/10.1016/B978-044481862-1.50103-5

[9] Leah Coutts. 2019. Empowering Students to Take Ownership of Their Learning: Lessons from One Piano Teacher's Experiences with Transformative Pedagogy. *International Journal of Music Education* 37, 3 (Aug. 2019), 493–507. https://doi.org/10.1177/0255761418810287

[10] Roger B. Dannenberg, Marta Sanchez, Annabelle Joseph, Peter Capell, Robert Joseph, and Ronald Saul. 1990. A Computer-based Multi-media Tutor for Beginning Piano Students. *Interface* 19, 2-3 (Jan. 1990), 155–173. https://doi.org/10.1080/09298219008570563

[11] Chris Donahue, Ian Simon, and Sander Dieleman. 2019. Piano Genie. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 160–164. https://doi.org/10.1145/3301275.3302288

[12] Rumen Donchev, Erik Pescara, and Michael Beigl. 2021. Investigating Retention in Passive Haptic Learning of Piano Songs. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 2 (June 2021), 60:1–60:14. https://doi.org/10.1145/3463513

[13] Andrew T. Duchowski. 2018. Gaze-Based Interaction: A 30 Year Retrospective. *Computers & Graphics* 73 (June 2018), 59–69. https://doi.org/10.1016/j.cag.2018.04.002

[14] David J. Elliott. 1990. Music as Culture: Toward a Multicultural Concept of Arts Education. *Journal of Aesthetic Education* 24, 1 (1990), 147–166. jstor:3332862 http://www.jstor.org/stable/3332862

[15] Stefan Engeser and Falko Rheinberg. 2008. Flow, Performance and Moderators of Challenge-Skill Balance. *Motivation and Emotion* 32, 3 (2008), 158–172. https://doi.org/10.1007/s11031-008-9102-4

[16] Bettina Eska and Jakob Karolus. 2022. Supporting Sportspeople in Gaining Bodily Insights through Reflective Feedback. In *Proceedings of the New Trends in HCI and Sports Workshop, Co-Located with the ACM International Conference on Mobile Human-Computer Interaction, NTSPORT@MobileHCI 2022, Vancouver, BC, Canada, October 1, 2022 (CEUR Workshop Proceedings, Vol. 3267)*. CEUR-WS.org.

[17] Kraig Finstad. 2010. The Usability Metric for User Experience. *Interacting with Computers* 22, 5 (May 2010), 323–327. https://doi.org/10.1016/j.intcom.2010.

04.004 arXiv:https://academic.oup.com/iwc/article-pdf/22/5/323/1992916/iwc22-0323.pdf

[18] Art C. Graesser, Patrick Chipman, Brian C. Haynes, and Andrew Olney. 2005. AutoTutor: An Intelligent Tutoring System with Mixed-Initiative Dialogue. *IEEE Transactions on Education* 48, 4 (Nov. 2005), 612–618. https://doi.org/10.1109/TE.2005.856149

[19] Miriam Greis, Tilman Dingler, Albrecht Schmidt, and Chris Schmandt. 2017. Leveraging User-Made Predictions to Help Understand Personal Behavior Patterns. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Vienna, Austria) *(MobileHCI '17)*. Association for Computing Machinery, New York, NY, USA, Article 104, 8 pages. https://doi.org/10.1145/3098279.3122147

[20] Bogdan Hoanca, Timothy C. Smith, and Kenrick Mock. 2014. Machine-Extracted Eye Gaze Features: How Well Do They Correlate to Sight-Reading Abilities of Piano Players?. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '14)*. Association for Computing Machinery, New York, NY, USA, 279—282. https://doi.org/10.1145/2578153.2578194

[21] Lars Erik Holmquist. 2017. Intelligence on Tap: Artificial Intelligence as a New Design Material. *Interactions* 24, 4 (June 2017), 28–33. https://doi.org/10.1145/3085571

[22] Feng Huang, Yu Zhou, Yao Yu, Ziqiang Wang, and Sidan Du. 2011. Piano AR: A Markerless Augmented Reality Based Piano Teaching System. In *2011 Third International Conference on Intelligent Human-Machine Systems and Cybernetics*, Vol. 2. 47–52. https://doi.org/10.1109/IHMSC.2011.82

[23] Gaoping Huang, Xun Qian, Tianyi Wang, Fagun Patel, Maitreya Sreeram, Yuanzhi Cao, Karthik Ramani, and Alexander J. Quinn. 2021. AdapTutAR: An Adaptive Tutoring System for Machine Tasks in Augmented Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3411764.3445283

[24] Robert J.K. Jacob and Keith S. Karn. 2003. Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises. *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (2003).

[25] Rose M.G. Johnson, Janet van der Linden, and Yvonne Rogers. 2010. MusicJacket: The Efficacy of Real-time Vibrotactile Feedback for Learning to Play the Violin. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. ACM, New York, NY, USA, 3475–3480. https://doi.org/10.1145/1753846.1754004

[26] Jakob Karolus, Hendrik Schuff, Thomas Kosch, Paweł W. Wozniak, and Albrecht Schmidt. 2018. EMGuitar: Assisting Guitar Playing with Electromyography. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, Hong Kong, China, 651–655. https://doi.org/10.1145/3196709.3196803

[27] Jakob Karolus, PawełW. Wozniak, Lewis L. Chuang, and Albrecht Schmidt. 2017. Robust Gaze Features for Enabling Language Proficiency Awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2998–3010. https://doi.org/10.1145/3025453.3025601

[28] Hieke Keuning, Bastiaan Heeren, and Johan Jeuring. 2021. A Tutoring System to Learn Code Refactoring. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. Association for Computing Machinery, New York, NY, USA, 562–568. https://doi.org/10.1145/3408877.3432526

[29] Veronica Kinsler and R.H.S. Carpenter. 1995. Saccadic Eye Movements While Reading Music. *Vision Research* 35, 10 (1995), 1447–1458. https://doi.org/10.1016/0042-6989(95)98724-N

[30] Tamaki Kitamura and Masanobu Miura. 2006. Constructing a Support System for Self-Learning Playing the Piano at the Beginning Stage. In *Proceedings of International Conference on Music Perception and Cognition*. Citeseer, 258–262.

[31] Kenneth R. Koedinger, John R. Anderson, William H. Hadley, and Mary A. Mark. 1997. Intelligent Tutoring Goes To School in the Big City. *International Journal of Artificial Intelligence in Education (IJAIED)* 8 (1997), 30. https://telearn.archives-ouvertes.fr/hal-00197383

[32] Marilyn J Kostka. 2004. Teach them how to practice. *Music Educators Journal* 90, 5 (2004), 23–26.

[33] Richard Li, Yingyan Wang, Chih-Pin Hsiao, Nicholas Davis, James Hallam, and Ellen Do. 2016. Tactile Teacher: Enhancing Traditional Piano Lessons with Tactile Instructions. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion (CSCW '16 Companion)*. Association for Computing Machinery, New York, NY, USA, 329–332. https://doi.org/10.1145/2818052.2869133

[34] Markus Löchtefeld, Sven Gehring, Ralf Jung, and Antonio Krüger. 2011. guitAR: Supporting Guitar Learning Through Mobile Projection. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 1447–1452. https://doi.org/10.1145/1979742.1979789

[35] Jaime Madell and Sylvie Héébert. 2008. Eye Movements and Music Reading: Where Do We Look Next? *Music Perception* 26, 2 (Dec. 2008), 157–170. https://doi.org/10.1525/mp.2008.26.2.157 arXiv:https://online.ucpress.edu/mp/article-pdf/26/2/157/190304/mp_2008_26_2_157.pdf

[36] Marcel Manshanden. 2021. Piano Books Divided into 15 Levels - PianoBookGuide. (Jan. 2021). https://www.pianobookguide.com/piano-book-level-search

[37] Roger Nkambou, Jacqueline Bourdeau, and Riichiro Mizoguchi. 2010. Introduction: What Are Intelligent Tutoring Systems, and Why This Book? In *Advances in Intelligent Tutoring Systems*, Roger Nkambou, Jacqueline Bourdeau, and Riichiro Mizoguchi (Eds.). Springer, Berlin, Heidelberg, 1–12. https://doi.org/10.1007/978-3-642-14363-2_1

[38] Sandra Y Okita. 2014. Learning from the folly of others: Learning to self-correct by monitoring the reasoning of virtual characters in a computer-supported mathematics learning environment. *Computers & Education* 71 (2014), 257–278.

[39] Richard Parncutt and Gary McPherson. 2002. *The Science & Psychology of Music Performance: Creative Strategies for Teaching and Learning*. Oxford University Press, Oxford and New York. 135–150 pages.

[40] Marjaana Penttinen. 2013. *Skill Development in Music Reading*. Ph. D. Dissertation. University of Turku, Finland.

[41] Project Petrucci LLC. 2021. Piano Pieces by Level – IMSLP. (Jan. 2021). https://imslp.org/wiki/Special:DiffPage/DiffMain/1

[42] Marjaana Puurtinen. 2018. Eye on Music Reading: A Methodological Review of Studies from 1994 to 2017. *Journal of Eye Movement Research* 11, 2 (2018), 1–16.

[43] Keith Rayner. 1998. Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychol Bull* 124, 3 (Nov. 1998), 372–422.

[44] Keith Rayner and Alexander Pollatsek. 1997. Eye Movements, the Eye-Hand Span, and the Perceptual Span during Sight-Reading of Music. *Current Directions in Psychological Science* 6, 2 (1997), 49–53. https://doi.org/10.1111/1467-8721.ep11512647 arXiv:https://doi.org/10.1111/1467-8721.ep11512647

[45] Liam Rigby, Burkhard C. Wünsche, and Alex Shaw. 2020. piARno - An Augmented Reality Piano Tutor. In *32nd Australian Conference on Human-Computer Interaction*. ACM, Sydney NSW Australia, 481–491. https://doi.org/10.1145/3441000.3441039

[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 *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces (ITS '14)*. Association for Computing Machinery, New York, NY, USA, 149–158. https://doi.org/10.1145/2669485.2669514

[47] Richard M Ryan and Edward L Deci. 2000. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American psychologist* 55, 1 (2000), 68.

[48] Herman Saksono and Andrea G. Parker. 2017. Reflective Informatics Through Family Storytelling: Self-Discovering Physical Activity Predictors. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5232–5244. https://doi.org/10.1145/3025453.3025651

[49] Dario D. Salvucci and Joseph H. Goldberg. 2000. Identifying Fixations and Saccades in Eye-Tracking Protocols. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*. ACM, 71–78. https://doi.org/10.1145/355017.355028

[50] Donald A Schön. 1987. *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* Jossey-Bass.

[51] Donald Jay Smith. 1988. *An Investigation of the Effects of Varying Temporal Settings on Eye Movements While Sight-Reading Trumpet Music and While Reading Language Aloud.* Ph. D. Dissertation. Pennsylvania State University.

[52] Tony Souter. 2001. *Eye Movement, Memory and Tempo in the Sight Reading of Keyboard Music.* Ph. D. Dissertation. University of Sydney / University of Sydney / Sydney Conservatorium of Music. http://hdl.handle.net/2123/8674

[53] Derjen Jeney Sun. 1993. *Teaching young children compositional concepts to enhance music learning in a computer learning environment.* The Ohio State University.

[54] Maria Timoshenko. 2018. Seeing into the Music Score: Eye-Tracking and Sight-Reading in a Choral Context. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications (ETRA '18)*. Association for Computing Machinery, New York, NY, USA, 1–2. https://doi.org/10.1145/3204493.3207415

[55] Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2013. Pursuits: Spontaneous Interaction with Displays Based on Smooth Pursuit Eye Movement and Moving Targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. Association for Computing Machinery, New York, NY, USA, 439–448. https://doi.org/10.1145/2493432.2493477

[56] Robert A. Virzi. 1992. Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough? *Hum Factors* 34, 4 (Aug. 1992), 457–468. https://doi.org/10.1177/001872089203400407

[57] Tina Caroline Walber, Ansgar Scherp, and Steffen Staab. 2014. Smart Photo Selection: Interpret Gaze As Personal Interest. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2065–2074. https://doi.org/10.1145/2556288.2557025

[58] Homer E. Weaver. 1943. Studies of Ocular Behavior in Music Reading. *Psychological Monographs* 55, 1 (1943), i–50. https://doi.org/10.1037/h0093537

[59] Blake Williford, Matthew Runyon, Wayne Li, Julie Linsey, and Tracy Hammond. 2020. Exploring the Potential of an Intelligent Tutoring System for Sketching Fundamentals. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376517

[60] Klaus Wolters. 2001. *Handbuch Der Klavierliteratur: Klaviermusik Zu Zwei Händen* (5., vollst. rev. aufl. ed.). Atlantis-Musikbuch-Verl., Zürich.
[61] Alfred Lukyanovich Yarbus. 1967. *Eye Movements and Vision.* New York: Plenum Press.
[62] Beste F. Yuksel, Kurt B. Oleson, Lane Harrison, Evan M. Peck, Daniel Afergan, Remco Chang, and Robert JK Jacob. 2016. Learn Piano with BACh: An Adaptive Learning Interface That Adjusts Task Difficulty Based on Brain State. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5372–5384. https://doi.org/10.1145/2858036.2858388
[63] Marcel Zentner, Didier Grandjean, and Klaus R. Scherer. 2008. Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement. *Emotion (Washington, D.C.)* 8, 4 (2008), 494–521. https://doi.org/10.1037/1528-3542.8.4.494