DJeye: Towards an Accessible Gaze-Based Musical Interface for Quadriplegic DJs

Fabio Bottarelli
Università degli Studi di Milano
Laboratorio di Informatica Musicale
fabio.bottarelli@studenti.unimi.it

Nicola Davanzo, Giorgio Presti
Federico Avanzini
Università degli Studi di Milano
Laboratorio di Informatica Musicale
{name.surname}@unimi.it

ABSTRACT

Despite the recent advancements in the development of accessible musical interfaces for individuals with limited motor capabilities such as quadriplegia, DJing still remains a relatively inaccessible musical activity. To address this issue, we propose the design and implementation of DJeye, an eye tracking-based software musical interface which allows typical basic mixing operations such as crossfading, filtering, looping, track seeking, and more. The interface is founded upon established design principles for gaze-based musical interfaces, and introduces specific eye interaction methods involving winking. Although currently in a prototype stage, we conducted case studies to evaluate the proposed interaction methods and explore which functions may be of interest to end-users. The study was conducted with the participation of amateur DJs without disabilities, who were subjected to interviews, think-alouds, and questionnaires. The results of the case studies are analyzed in the paper to provide insights into future directions and developments.

1. INTRODUCTION

DJing refers to a vast set of practices characterized by the live performance of pre-recorded music tracks, which typically involves a variety of tools, including turntables or other hardware decks, software consoles, drum machines and groove boxes, mixers and effects, to control the playback and manipulation of these tracks [1]. Common actions performed by a DJ using their console include the loading and playing of tracks, adjustment of volume, pitch and tempo, mixing of multiple tracks, application of effects such as reverb or delay, scratching and beatmatching for seamless transitions, as well as interaction with the audience to gauge their musical preferences and adjust the music selection accordingly.

The purpose of this paper is to present DJeye, an accessible interface for DJing performances, which uses solely gaze pointing and winking (i.e. one eye blinking) as physical interaction channels, and can consequently be used by any musician without requiring the active use of the upper limbs. As such, DJeye can be classified as an Accessible Digital Musical Instrument (ADMI), a term commonly used in the literature to refer to instruments designed for persons with diverse abilities. ADMIs are receiving a growing interest from both the research community and the industry [2]. In particular, the use of several physical interaction channels (eye/head movements, breath, etc.) has been explored to design musically expressive instruments that can be accessed also by persons with motor impairments, including quadriplegia [3]. In this general context, however, very few previous studies have focused on the design and evaluation of accessible DJ interfaces.

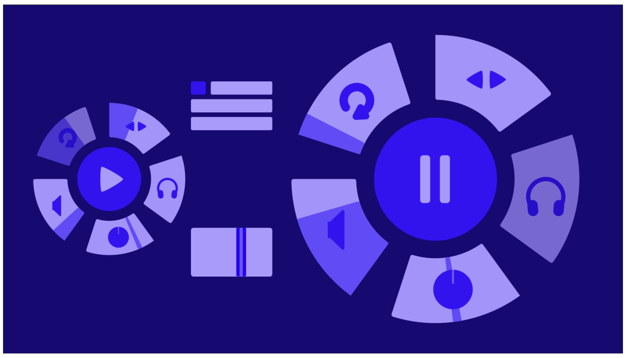

Figure 1. DJeye Interface at full screen, with focus on the right deck and the corresponding buttons/sliders.

DJeye is essentially a controller which outputs MIDI events directed towards an underlying DJing software, namely MIXXX, which was chosen as it is FOSS [1] and integrates all the features needed by the controller [4]. From the user side, DJeye receives input from an eye-tracking device and converts it into events to control the various elements of the interface. After a review of related works in Sec. 2, the design and current implementation of DJeye are presented in Sec. 3, while Secs. 4 and 5 discuss methods and results of an initial evaluation, respectively.

2. RELATED WORKS

The literature in the domain of DJing reveals several projects exploring novel interfaces for musical expression. As an

[1] Free Open-Source Software.

example, the ColorDex system [5] employs a new operational metaphor that allows a DJ to prepare up to six tracks and perform mixes between up to three of them, while minimizing complexity and enhancing the performance and engagement for both the performer and audience. The *Wearable DJ System* by Tomibayashi et al. [6] utilizes wearable computing and gesture recognition technologies to enable DJ techniques through intuitive gesture operations. The system's accuracy and effectiveness was confirmed through demonstrations.

Other notable digital interfaces in the DJing and NIME (New Interfaces for Musical Expression) field include D'Groove [7], ReacTable [8], and the MultiTouch DJ prototype [9]. D'Groove is a turntable equipped with torque/force feedback, which is powered by the computer during music playback. ReacTable is a synthesizer that uses a fully lit table, which interacts with tangible objects placed on top by the performer. The MultiTouch DJ prototype is a DJ controller that aims to merge traditional DJ setup with modern touch interface technology, while preserving the relevant technical abilities of disc jockeys.

Gaze pointing and blinking are established techniques in accessible interaction design for individuals with quadriplegia [3]. The unique characteristics of human eye movements must however be taken into account in the design of eye-tracking based musical instruments. Gaze typically moves through rapid and short *saccades*, still *fixations* with a duration of 100-400ms, and *smooth pursuits* for following moving targets [10]. Gaze pointing has various applications in Human-Computer Interaction, including mouse emulation, gaze-based text entry, web browsing, gaze-controlled games, attention-aware interfaces, user modeling and monitoring. Previous experiments have evaluated its speed and stability for its use in interactive interfaces [11]. Blinking and winking are still minimally explored as interaction channels. The ability to selectively close one eye seems to be linked to personal abilities and may be affected by motor impairments. A spontaneous blink typically lasts for 300-350 ms [12]: distinguishing voluntary blinks is important for interaction design. Winks are, on the other hand, rarely involuntary.

Malloch et al. [13] provided a model to classify three different behaviors which characterize Digital Musical Instruments (DMIs) interaction. *Skill-based* behaviors are automatic and performed without conscious attention (e.g. playing an acoustic instrument). *Rule-based* behaviors follow learned rules or procedures, with information perceived as signs (e.g. sequencing or live diffusion). *Model-based* behaviors are directed towards a conceptual goal and information is perceived as symbols (e.g. algorithmic music composition or presentation of recorded material). If we exclude turntablism (which involves scratching, beatmatching, beat juggling and other skill-based techniques, it can be stated that DJing is primarily oriented towards rule- or model-based interaction behaviors. While we can find several notable examples of skill-based performance DMIs based on eye tracking [14–19] eye-based interaction remains largely under-explored in the field of DJing, and rule-based instruments in general.

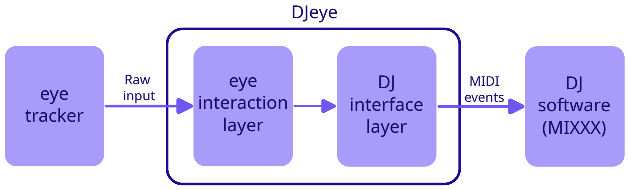

Figure 2. Flow of events during the interaction with the DJeye system, from eye tracker output to musical events on MIXXX DJ software.

A related contribution can be found in the report by Hornof [20], where the Ableton-Live-Adapted Interface developed by Margulies and Anderson in 2012 is described. This interface was built upon the Digital Audio Workstation (DAW) Ableton Live: rather than creating a new graphical interface, an existing one was modified to suit gaze based interaction.

Davanzo and Avanzini [21] presented a compilation of indications and strategies for the design of gaze-based interactions. The Midas Touch problem, which refers to an unintended interaction or selection that occurs due to the user's gaze either intentionally or unintentionally encountering an interactive surface, is a crucial issue to consider in this context. Among the proposed solutions is the separation of pointing gaze-aware elements and their activation into separate interaction channels (gaze and winks in DJeye), which the authors refer to as *hybrid interaction*. As further explained in Sec. 3, DJeye interaction includes the emulation mouse dragging actions by alternating the closure of one eye and the movement of the other pupil. Inspiration was drawn from a study conducted by Ramirez Gomez et al. [22], which however acknowledges the potential for some inaccuracies.

3. DESIGN AND IMPLEMENTATION

We refer to DJeye as the combination of two different software components, namely the *DJ Interface Layer* and the *Eye Interaction Layer*, which receive input from an eye tracking device, and output MIDI events directed to a DJing software running in the background, namely *MIXXX* [2]. This data flow is depicted in Fig. 2. While *Eye Interaction Layer* receives the signal from the eye tracker and converts it into mouse cursor events, the DJ Interface Layer provides the actual graphical interface of the instrument (Fig. 1).

DJeye is published under GNU-GPL v3 free software license, and the source code, as well as the MIDI mapping scripts for MIXXX (ref. Sec. 1) are available on a GitHub repository [3]. A demo video showing a DJeye performance can be found on YouTube [4].

3.1 Eye Interaction Layer

Although some sliders are oriented horizontally, dragging in DJeye always occurs in the up or down directions. The

[2] MIXXX official website: `https://mixxx.org/`
[3] DJeye GitHub repository: `https://github.com/LIMUNIMI/DJeye`
[4] DJeye demo on YouTube: `https://youtu.be/-bs08Ohdr7w`

motivation behind this choice, which may result in a slightly unnatural interaction for the user, is that we noticed a slight horizontal ”spike movement” of the gaze input upon closing one eye, which would result in unwanted dragging. Gaze movement controls the position of the mouse cursor, applying a simple lag-less fixation discrimination filter to improve its accuracy: the cursor is not moved unless a minimum threshold distance from the previous position is exceeded (thus a saccadic movement happened). A left wink causes a left mouse button click, while a right eye closure causes hold and drag events. An eye closure is detected when a number of input samples from the eye tracker are “null” or “invalid” (the eye is not detected), corresponding to at least 80 ms of eyelid closure. Many interface controls (sliders and crossfaders) support drag actions, which occur in the following manner: (a) the user stares at the target element and closes the right eye; (b) the user moves the gaze point slightly upwards (a few tens of pixels) relative to the target, to drag upward, or downward to drag downward; (c) the slider is dragged up or down at a constant rate, at discrete intervals, by a predetermined value; (d) the user reopens the right eye and the dragging action ends (Fig. 3). As shown in the figure, there is a small “inactive zone” where no dragging action takes place.

Separating the Eye Interaction Layer from the DJ Interface Layer allows to abstract on the type of eye tracker employed by the user: using a different device involves a simple replacement of this Layer. This way, the DJ Interface can potentially be controlled with any device that can simulate mouse events (e.g. an head tracker with appropriate mappings). Cross-system compatibility of the DJ Interface is also potentially assured, making it independent from device drivers compatibility. The default Eye Interaction Layer we developed currently supports Tobii Eye Trackers. [5]

3.2 DJ Interface Layer

The Controller is designed as a multi-platform application, which outputs MIDI events directed towards an underlying DJing software.

A typical design for a DJ controller includes two (or four) decks and a mixer. The decks contain controls for individual tracks (such as tempo, loop, and play/stop) while the mixer generally contains controls for volume, filters, and gain of all tracks, as well as the crossfader. As reported in [21], the gaze output of consumer eye trackers is not particularly accurate and usually requires spatial filtering or other mitigation solutions. Several strategies should be adopted to curb the problem, while displaying all the key interface elements in one screen. We chose not to hide controls in overly deep menus, because during mixing we considered it important to be able to perform many actions in a short amount of time. We therefore chose to reduce the number of controls to the essentials. Tab. 1 lists the controls available in each deck, as well as their reaction to click or drag inputs. The decks feature the necessary *Play/Pause* button as well as additional controls we

[5] We tested it with Tobii Eye Tracker 5, but it’s compatible with Tobii 4C, EyeX, PCEye Mini and others.

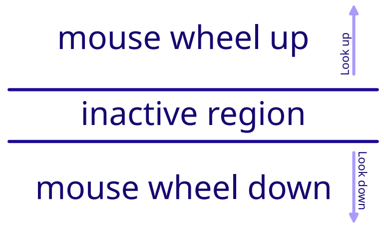

Figure 3. A depiction of the eye dragging system. Between the upper and lower dragging zones, an inactive (or “dead zone”) region is present, where no action occurs.

deemed essential for a DJing performance. These include: a *Volume* button for adjusting the volume of the two tracks; a *Loop* button for looping a small section of a track (e.g. in house music mixing); a *Highpass/Lowpass Filter* button for cutting frequencies from a track; a *Track Seek* button for selecting the starting point of a track; and a *Listen on Headphones* button for pre-listening to the next track. Although the Volume button is not strictly essential due to the presence of a crossfader, it is commonly found in mixer-style consoles. The Loop button, while not a fundamental control, is useful for blending tracks and creating a “growing energy” effect. The Highpass/Lowpass Filter is often used in mixing and enables creative applications beyond track blending. The Track Seek and Listen on Headphones buttons are considered fundamental as they allow the DJ to decide the starting point and pre-listen to tracks.

The circular arrangement of the deck buttons and associated sliders can be seen in Fig. 1. The play button is located at the center and surrounded by sliders. The value of the sliders can be modified by clicking and vertical dragging. For some sliders, a single or double click can trigger different events, as indicated in Table 1. The central section between the two decks includes a button to open the track selection window and a crossfader (see Fig. 1). The crossfader is a slider that adjusts the volume of the two tracks in a cross-fading manner: when placed at the right/left extreme, only the corresponding track will be heard, while at the center both tracks will be heard at the same volume.

Following the design cues reported in [21], In DJeye, the cursor is hidden to avoid involuntary movements and draw the user’s attention to the interface elements. Each control has an icon at its center, serving as a target for gaze focus. To compensate for the relatively low accuracy of eye tracker-based gaze pointing (compared to other pointing devices such as a mouse), interface components have been designed with a wide surface. Moving the cursor over a deck enlarges it while reducing the size of the other through a fluid animation. Figure 1 illustrates this when the focus is on the right deck.

DJeye’s decks layout draws inspiration from various musical interfaces, including The EyeHarp [14], which features a distinctive pie-shaped keys layout. This ensures a close spatialization of the visual elements. The concept has also been endorsed by [21] as a potential solution to the Midas Touch issue (see Sec. 2). However, due to the presence of the *Play* button in the central area of each deck, the mitigation of the Midas Touch problem is not achieved in

Control	Click (left eye wink)	Click and drag (holding right eye closed)
Play/Stop	Single click starts and stop the loaded track	-
Volume	Double click resets the volume to default value (0)	Output volume up/down
Loop	Single click enables/disables a loop, double click resets the loop to default duration (2 bars)	Changes loop duration, in bars
High/Low pass filter	Double click resets the filter to default position (center)	Adjusts the filter
Track Seek	-	Navigate the track, moving the playback to the desired point
Headphones	Single click enables/disables listening the track on headphones	-

Table 1. Controls included in each of the two decks, and events associated with the two available interaction methods.

Figure 4. Big Library view in MIXXX, with superimposed gaze operable buttons from DJeye.

this case. A hybrid interaction strategy is used, separating item selection and activation between gaze pointing and winking. This approach is faster than dwell-time methods, which introduce a delay between gaze and visual element activation. Being able to complete more tasks in a given time-frame is valuable for a DJ controller. The number of radial-arranged sliders can be increased by simply reducing their size, thereby allowing for future introduction of new functions. Furthermore, this shape was selected due to its similarity to a vinyl record, which is commonly seen in DJ consoles.

As mentioned earlier, DJeye is a MIDI controller, and as such it communicates with an underlying DJing software. MIXXX was chosen because it allows for MIDI mappings scripting. During the track selection operation, DJeye provides a partially transparent interface with eye operable arrow buttons to select tracks on the underlying "Big Library" view in MIXXX, and load the selected track into the left or right deck (see Fig. 4). The software also provides track quantization and synchronization, two useful functions typical of many mixing applications.

The DJ interface layer is implemented using the C++>11 programming language. Both graphical user interface design and MIDI messages transmission leverage the JUCE [6] cross-platform framework for audio application/plug-in development. The program is designed to be compatible across multiple operating systems, although Android, IOS, and IpadOS are not supported due to the lack of compatibility with MIXXX and the limited support for pop-up windows on these platforms (necessary for DJeye to work).

[6] JUCE official website: `https://juce.com/`

4. EVALUATION

4.1 Background

Although there is a growing interest in evaluating music interfaces within the NIME community (as indicated by Barbosa et al. [23]), this task poses significant challenges, with a lack of consensus among the community regarding the appropriate methods.

The objective of user experience (UX) evaluations, as stated by Springett [24], is to provide meaningful insights into the affective factors involved in interaction, supporting iterative system development for designers and stakeholders. Reimer and Wanderley [25] emphasize the value of measuring task performance and the importance of exploratory evaluation approaches that examine how individuals adapt and utilize technology. Music interface evaluations often use a combination of open exploration and specific task-based methods, according to the same authors. Exploratory evaluations, which assess user perceptions and interaction strategies, are prevalent in literature and provide rich data for informed future evaluations. Self-report methods, including open comments, questionnaires, and interviews, are widely used for evaluating musical interfaces and produce qualitative data. Wanderley and Reimer point to a range of psychometric analysis tools adapted from a general Human-Computer Interaction context, and tools created instead directly for the assessment of DMIs.

Among the latter, the authors advocate the use of the MPX-Q questionnaire (formulated by Schmid [26]) as a formal psychometric assessment instrument for digital music interfaces. This was created through multiple interviews with musicians to determine the most critical aspects to evaluate in a musical instrument. A dimensionality reduction procedure was carried out to reduce the number of aspects to a few core areas of inquiry, resulting in 43 extracted items/questions. The aspects evaluated can be summarized into three overarching categories: (1) Experienced freedom and possibilities, (2) Perceived control and comfort, and (3) Perceived stability, sound quality, and aesthetics. The MPX-Q is founded on the general idea of musical instruments and interactive devices as presented by Bongers [23] and Malloch et al. [13], where physical gestures are directly converted into sound production by the instrument. However, based on Malloch et al.'s interaction model discussed in Sec. 2, some of the questions in the MPX-Q may be not suited for rule/model-based software interfaces such as DJeye, and as such they were omitted. Omitted questions concerned sound quality, the ability to evoke different musical styles (which both depend on the employed musical repertoire rather than the instrument itself), and the quality and solidity of building materials (impossible to evaluate in a software instrument).

4.2 Evaluation method

The purpose of our study was to perform an exploratory investigation, as DJeye is still a prototype (i.e. at a preliminary stage of its development). The study aimed to address the following research questions: (a) Which capabilities and functions required by DJs that are currently not included in the application; (b) Perception of the interface design in terms of experienced freedom, control and comfort, and aesthetics (covered by the MPX-Q); (c) Comparison of the interface with a traditional DJ console, to highlight its limitations; (d) Naturalness and effectiveness of the proposed eye interaction method; (e) Level of fatigue associated with the gaze- and wink-based interaction method; (f) Impressions, emotions, and ideas of the users. The findings of the study should inform and prioritize the development of the DJeye interface in subsequent iterations.

Based on the research questions and the background exposed in Sec. 4.1, we defined an evaluation method that consists of: observation of users while they interact with the interface; interviews and think-aloud procedures during use; a combination of open and closed evaluation questionnaires, including scored questions, to direct the user's focus on specific aspects of the experience. Our experimental sessions consisted of the following phases:

1. **Preliminary phase**: Participants were introduced to the experiment, and demographic information was collected, including their experience with DJ interfaces and eye tracking. The eye tracker was also calibrated during this phase.
2. **Free phase I**: Participants were given a brief overview of the gaze- and wink-based interaction mechanism, and were then allowed to interact with DJeye for a designated period of time (7 minutes) without any instructions or prior knowledge of the interface. Participants were encouraged to "think aloud" and verbally express their intentions, emotions, and feelings while interacting with the system.
3. **Tasks phase**: Participants were guided in using each component of the DJeye interface to perform a simple DJing performance. This phase involved mixing live tracks using crossfading, filters, loops, headphone listening, track seeking, and volume control.
4. **Free phase II**: Similar to the previous "Free Phase", participants were again instructed to "think aloud" and perform free interaction with DJeye for a designated period of time (5 minutes). Since the previous phase, we expected the participants had acquired knowledge on how to operate the interface.

The evaluation process takes inspiration from a three-stage approach outlined by Stowell [27], suitable for discourse analysis, comprising free exploration (where participants are encouraged to freely explore the interface), guided exploration (where participants are provided with tasks and materials to create their own), and a semi-structured interview (focused on eliciting participant experiences from the free and guided exploration phases).

Our study employed a standardized script, administered by the test moderator, to instruct participants. The test sessions were recorded and the conversations were transcribed for analysis. Key extracts of the dialogues and significant observations made by the testers were documented. The script and transcriptions of the conversations, anonymized, as well as the list of tasks performed during Free Phase I and II, are accessible online as supplementary material [7].

Through a questionnaire, each tester was then asked to evaluate different facets of their experience. This questionnaire was comprised of 7-value Likert scale questions (degree of agreement with the sentence), open-ended and semi-open-ended questions. Those were partially inspired by the questions and concepts proposed by the aforementioned MPX-Q (Sec. 4.1), but featured survey topics appropriate to the analysis of a DJ interface, eye interaction, eye fatigue and interface completeness. To further analyze the experience, testers were invited to provide explanations for the assigned scores, as well as additional comments, suggestions and observations. Tab. 2 summarizes the Likert scored questions, indicating for each which aspects/domain they were aiming to investigate. Investigation domains included *freedom* of expression and experiencing; *explorability* of the interface; level of *engagement* with the instrument; *creative* possibilities; *conformance* with the expectations; *playing comfort*; *control* precision; *ergonomics*; *stability* of the system; *aesthetics* pleasure; *eye interaction* related issues.

Open-ended questions, not reported in the table but asked by the test moderator at the end of the questionnaire, included:

(a) Using between 1 and 4 adjectives, how did you feel during the performance with the instrument?

(b) Did you find that the interface gave you the same expressive possibilities as a traditional DJ interface? What would you add or change?

(c) Would you integrate an eye tracking-based interface into your DJing setup? Why?

(d) Express freely some comments about the experience.

5. RESULTS AND ANALYSIS

A sample of 5 musicians with previous amateur level experience in DJing volunteered for the study. The participants' ages were between 22 and 28 years old. The participants reported no prior experience with eye tracking technologies. Due to difficulties in accessing a sample of quadriplegic individuals, the chosen participants did not have any disabilities. The study was conducted partially in the soundproof room at the Laboratory of Music Informatics (LIM) at the University of Milan and partially at a private location. We utilized a system that comprised a

[7] Zenodo link: https://doi.org/10.5281/zenodo.7782513

Question	Score	All
Freedom domain		
The instrument allows me to learn new things	6	15677
The instrument offers me new facets of playing	6	34667
The instrument expands my experience of musical interaction	6	16677
I sense that the instrument offers a lot of variety	4	14445
The instr. offers me interesting possib. for sound manipulation	4	24455
I found the control system to be sufficiently comprehensive*	5	35567
Explorability domain		
I can continuously discover new things using the instrument	4	33456
Engagement domain		
The instrument keeps me interested	7	66777
It is easy to get into the flow of performance with the instr.	5	25566
The instrument keeps me engaged while playing it	7	67777
Creativity domain		
The instrument stimulates my creativity	6	25667
Conformance domain		
The instrument does what I want it to do	6	56677
The instrument responds well to my actions	5	55566
I can control the instrument intuitively	6	35667
The instrument works the way I expect it to	5	55567
I feel like I am controlling the behavior of the instrument	6	46667
Playing Comfort domain		
I perceive the instrument as comfortable to play	5	44567
The system allows me to play with comfortable gestures	5	55567
I feel relaxed when i play the instrument	5	34556
Control domain		
I can precisely control the instrument	5	55556
Ergonomics domain		
I feel comfortable when i play the instrument	6	55666
I feel that the instrument is an extension of my body	4	34466
Stability domain		
I think the instrument is well made	6	66667
the instrument is reliable	6	55667
Aesthetics domain		
The instrument looks appealing	6	46677
The instrument is aesthetically pleasing	7	55777
Eye Interaction domain		
I did NOT experience eye fatigue during the interaction*	6	34667
I found the eye interaction to be accurate and precise*	5	55566
Involuntary movements or unwanted interactions did NOT occur during the performance*	6	46667

Table 2. Questionnaire with Likert-7 scoring, subdivided by domain of investigation. All questions (except those marked with a '*') were taken from the MPX-Q questionnaire. Median is used as a measure of central tendency. Due to the low number of subjects, we reported all the results in the "All" column instead of a measure of dispersion.

laptop, a sound card for audio output management, professional studio headphones, a speaker system, and a Tobii Eye Tracker 5. [8] We did not record nor collect visual traces or gaze data through the sensor. Each session lasted approximately 60 minutes.

We analyzed interview transcripts by manual clustering of users' comments. During *Free Phase I*, we collected 69 observations divided into 5 clusters (+ 5 outliers). Those include:

- **Failure to achieve desired goals**. Because of unknown strategy (4 obs.), or because of wrong strategy employed (15 obs.). In particular in 7 cases the problem was about the tracks menu; in 5 cases about sliders and faders; in 4 cases the identification of the correct deck; in 3 cases the usage of the loop function.
- **Interaction fatigue** (15 obs.): in particular 5 observations were about dragging fatigue, among which 4 specifically concerned the crossfader; 3 observations about clicking fatigue; 2 about eye tracking calibration issues.
- **Complaints about the graphical user interface** (8 obs.). Spanning across different topics (buttons don't look like buttons; zooming decks are distracting; missing waveforms; unhappy with dragging strategies; inability to read while scrolling/dragging).
- **Achieved goals** (17 obs.). Spanning across different topics. Some of them were related to issues solved during the same session.
- **Verbal manifestations of frustration** (6 obs.), all from the same user.

Outliers included randomly achieved goals (2 obs.); unknown functionalities (2 obs.); random/unexpected behavior of the interface (1 ob.).

During *Free Phase II*, we collected 30 observations divided into 4 clusters.

- **Failure/difficulty to achieve goals** (7 obs.). Mainly regarding crossfader and seek functions, as well as wink-based clicking reliability.
- **Interaction fatigue** (6 obs.). 3 regarding wink-based interaction; 1 regarding eye fatigue; 1 reporting that "interacting with hands is easier than with eyes"; 1 reporting that "it is unnatural to always look at the DJ console".
- **Complaints about the graphical user interface** (14 obs.).
- **Achieved goals** (3 obs.). 1 about achieved double clicking; 1 about getting used to the system, and 1 appreciation of the completeness of the interface.

Results of the 7-values Likert scores questionnaire are reported in Tab. 2. Answers to the final open-ended questions can be summarized as follows.

(a): four out of the five users demonstrated positive emotions, feeling "ecstatic, surprised, excited, engaged, interested, curious, confident, relaxed, in the flow of music". One user felt "amused" but also "slightly fatigued in the eyes";

(b): users showed interest in the addition of low, middle and high frequency knobs in the filter, as well the introduction of additional effects (echo, reverb, flanger). One of them noted the lack of an indication about which tracks are being played in the main interface, as well as a revision of the drag direction (details in Sec. 6);

(c): two of them think the system could open up new multi-tasking possibilities during live sessions; one of them thought that it would be a "cool" technology to show to the audience; one user would use it just for track selection; one of them wouldn't need it;

(d): two users expressed positive emotions, finding the system useful, innovative, and well-structured. one of them would like to try a full and final version of the instrument again; one users remarked the desire for more graphical indicators; two users did not answer.

[8] Tobii Eye Tracker 5 (with specs) on Tobii's website: `https://gaming.tobii.com/product/eye-tracker-5/`

6. DISCUSSION AND CONCLUSIONS

Despite the results of the questionnaire indicating that the proposed interaction method is intuitive and comfortable, several users expressed eye fatigue in the open-ended comments. To improve the ergonomics of the interaction, it may be useful to explore additional hybrid interaction modes suitable for quadriplegic users [3], such as a combination of gaze pointing with two-eyed blinks, head tracking, or nodding gestures detected through eye tracking [22].

In terms of aesthetics the interface was found to be pleasing, but some reported interaction problems could be attributed to eye interaction difficulties, as well as the proposed visuals. To address these issues, colors with greater contrast could be used and different layouts for the keys could be explored [21]. Some users expressed the need for more on-screen indicators, showing in particular the name of the track currently being played. This requires a re-implementation of the controller or the creation of new display strategies as, since DJeye is a MIDI controller, the data flow with MIXXX is currently one-way. To enhance the user experience, new eye interaction strategies, such as continuous and vertical bars, pop-up bars, or different types of dragging methods could be introduced.

The emotions expressed by users during interaction were generally positive, and the freedom of expression was felt to be good, although some complaints were expressed about the low explorability and limited controls and options for sound manipulation. Despite this, users were interested in the multitasking possibilities offered by the interface, which does not require the use of hands.

Although inspired by *Gaze+Hold* wink-based techniques to click and drag object proposed by Ramirez Gomez et al. [22], DJeye's slider dragging technique works in a slightly different way. Therefore, it may be interesting to assess the usability and reliability of this new technique. The interviews suggest that winking is a robust interaction method for this interface. However, there is a lack of research on the proportion of the population proficient in winking and its rhythmic capabilities. An experiment could measure these factors by instructing participants to close their eyelids in time with a metronome set to an increasing frequency, starting at 50 beats per minute. This would enable measurement of time delay, consistency, accuracy, and user fatigue.

It should be noted that the laboratory setting used for the test better simulates private use of the instrument, rather than a live setting, which is an intrinsic aspect of DJing as noted by Ahmed et al. [28]. As indicated by the authors, DJing activities also include collection and preparation of material as well as promotion and performance; DJeye improves the accessibility of only the latter. DJeye gaze- and wink-based interaction does not engage the mouth, although a multitude of mouth-related features could have been exploited as interaction channels [3]; this is to maintain the ability to interact vocally with the audience as usually happens during live DJing performances (e.g. by cheering the audience, or accepting requests for tracks).

Our evaluation has a preliminary and short-term nature. While it provided useful insights, according to Reimer and Wanderley [25] the development of skills and expertise in musical instrument use is a longitudinal process that occurs over an extended period of time, as performers establish relationships with their instruments. As pointed out by the same authors, however, this type of evaluation is often neglected, and could be carried out in the future with a complete and stable version of DJeye. Furthermore our evaluation only considered the perspective of the performer and it would be important to also consider the views of other stakeholders, such as the audience, as advocated by O'Modhrain [29]. This could be accomplished through live performance evaluations, accompanied by a questionnaire to gauge audience reception (as carried out for The EyeHarp in [14]).

Finally, our evaluation did not include quadriplegic musicians. Interaction with DJeye do not require the use of upper or lower limbs; however, quadriplegic individuals may have unique characteristics that could complicate proper eye tracking setup, such as spasmodic head movements. Further case studies should include them to study its effectiveness with this population and their individual and specific needs.

7. REFERENCES

[1] S. Webber, *DJ Skills – The essential guide to Mixing & Scratching*. Oxford: FOcal Press, 2008.

[2] E. Frid, "Accessible digital musical instruments – a review of musical interfaces in inclusive music practice," *Multimodal Technologies and Interaction*, vol. 3, no. 3, p. 57, 2019.

[3] N. Davanzo and F. Avanzini, "Hands-Free Accessible Digital Musical Instruments: Conceptual Framework, Challenges, and Perspectives," *IEEE Access*, vol. 8, pp. 163 975–163 995, 2020.

[4] T. H. Andersen, "Mixxx: Towards novel dj interfaces," in *Proc. Int. Conf. New Interfaces for Musical Expression*, Montreal, 2003, pp. 30–35.

[5] N. Villar, H. Gellersen, M. Jervis, and A. Lang, "The ColorDex DJ system: A new interface for live music mixing," in *Proceedings of the 7th International Conference on New Interfaces for Musical Expression*, ser. NIME '07. New York, NY, USA: Association for Computing Machinery, Jun. 2007, pp. 264–269.

[6] Y. Tomibayashi, Y. Takegawa, T. Terada, and M. Tsukamoto, "Wearable DJ system: A new motion-controlled DJ system," in *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, ser. ACE '09. New York, NY, USA: Association for Computing Machinery, Oct. 2009, pp. 132–139.

[7] T. Beamish, K. Maclean, and S. Fels, "Manipulating music: Multimodal interaction for DJs," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '04. New York, NY, USA: Association for Computing Machinery, Apr. 2004, pp. 327–334.

[8] S. Jordà, G. Geiger, M. Alonso, and M. Kaltenbrunner, "The reacTable: Exploring the synergy between live music performance and tabletop tangible interfaces," in *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, ser. TEI '07. New York, NY, USA: Association for Computing Machinery, Feb. 2007, pp. 139–146.

[9] P. Lopez, A. Ferreira, and J. A. M. Pereira, "Battle of the DJs: An HCI Perspective of Traditional, Virtual, Hybrid and Multitouch DJing," in *11th International Conference on New Interfaces for Musical Expression (NIME 2011)*. Oslo, Norway: nime.org, May 30, pp. 367–372.

[10] P. Majaranta and A. Bulling, "Eye Tracking and Eye-Based Human–Computer Interaction," in *Advances in Physiological Computing*, ser. Human–Computer Interaction Series, S. H. Fairclough and K. Gilleade, Eds. London: Springer, 2014, pp. 39–65.

[11] N. Davanzo and F. Avanzini, "Experimental Evaluation of Three Interaction Channels for Accessible Digital Musical Instruments," in *Proc. '20 Int. Conf. on Computers Helping People With Special Needs*. Online Conf.: Springer, Cham, Sep. 2020, pp. 437–445.

[12] F. VanderWerf, P. Brassinga, D. Reits, M. Aramideh, and B. Ongerboer de Visser, "Eyelid movements: Behavioral studies of blinking in humans under different stimulus conditions," *Journal of Neurophysiology*, vol. 89, no. 5, pp. 2784–2796, May 2003.

[13] J. Malloch, D. Birnbaum, E. Sinyor, and M. M. Wanderley, "Towards a New Conceptual Framework for Digital Musical Instruments," in *Proc. 9th Int. Conf. on Digital Audio Effects*, Montreal, Canada, September 18-20, 2006, 2006, pp. 49–52.

[14] Z. Vamvakousis and R. Ramirez, "The EyeHarp: A Gaze-Controlled Digital Musical Instrument," *Frontiers in Psychology*, vol. 7, p. article 906, 2016.

[15] N. Davanzo, P. Dondi, M. Mosconi, and M. Porta, "Playing music with the eyes through an isomorphic interface," in *Proc. of the Workshop on Communication by Gaze Interaction*. Warsaw, Poland: ACM Press, 2018, pp. 1–5.

[16] OpenUpMusic, "The Clarion," http://openupmusic.org/the-clarion/, 2021.

[17] S. Valencia, D. Lamb, S. Williams, H. S. Kulkarni, A. Paradiso, and M. Ringel Morris, "Dueto: Accessible, Gaze-Operated Musical Expression," in *The 21st International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '19*. Pittsburgh, PA, USA: ACM Press, 2019, pp. 513–515.

[18] N. Davanzo, M. De Filippis, and F. Avanzini, "Netychords: An Accessible Digital Musical Instrument for playing chords using gaze and head movements," in *In Proc. '21 Int. Conf. on Computer- Human Interaction Research and Applications (CHIRA '21)*, Online conf., 2021.

[19] D. Kandpal, P. R. Kantan, and S. Serafin, "A Gaze-Driven Digital Interface for Musical Expression Based on Real-time Physical Modelling Synthesis," Jun. 2022.

[20] A. J. Hornof, "The Prospects For Eye-Controlled Musical Performance," in *Proc. 14th Int. Conf. on New Interfaces for Musical Expression (NIME'14)*, ser. NIME 2014, Goldsmiths, University of London, UK, Jul. 2014.

[21] N. Davanzo and F. Avanzini, "Design Concepts for Gaze-Based Digital Musical Instruments," in *Proceedings of the 2022 Sound and Music Computing Conference*, Saint-Etiénne, France, Jun. 2022, pp. 477–483.

[22] A. R. Ramirez Gomez, C. Clarke, L. Sidenmark, and H. Gellersen, "Gaze+Hold: Eyes-only Direct Manipulation with Continuous Gaze Modulated by Closure of One Eye," in *ACM Symposium on Eye Tracking Research and Applications*, ser. ETRA '21 Full Papers. New York, NY, USA: Association for Computing Machinery, May 2021, pp. 1–12.

[23] J. Barbosa, J. Malloch, M. M. Wanderley, and S. Huot, "What does "Evaluation" mean for the NIME community?" in *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME'15)*, Louisiana State University, Baton Rouge, LA, USA, May 2015, pp. 156–161.

[24] M. Springett, "Evaluating cause and effect in user experience," *Digital Creativity*, vol. 20, no. 3, pp. 197–204, Sep. 2009.

[25] P. J. C. Reimer and M. M. Wanderley, "Embracing Less Common Evaluation Strategies for Studying User Experience in NIME," in *International Conference on New Interfaces for Musical Expression*, Jun. 2021.

[26] G.-M. Schmid, *Evaluating the Experiential Quality of Musical Instruments*. Wiesbaden: Springer Fachmedien, 2017.

[27] D. Stowell, M. D. Plumbley, and N. Bryan-Kinns, "Discourse Analysis Evaluation Method for Expressive Musical Interfaces," Genoa, Italy, pp. 81–86, Jun. 2008.

[28] A. Ahmed, S. Benford, and A. Crabtree, "Digging in the crates: An ethnographic study of DJS' work," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '12. New York, NY, USA: Association for Computing Machinery, May 2012, pp. 1805–1814.

[29] S. O'Modhrain, "A Framework for the Evaluation of Digital Musical Instruments," *Computer Music Journal*, vol. 35, no. 1, pp. 28–42, Mar. 2011.