A Reactive Environment for Dynamic Volume Control

Dalia El-Shimy
Centre for Intelligent Machines
3480 University Street
Montréal, Québec
dalia@cim.mcgill.ca

Thomas Hermann
CITEC, Bielefeld Univeristy
Universitätsstraße 25
Bielefeld, Germany
thermann@techfak.uni-bielefeld.de

Jeremy R. Cooperstock
Centre for Intelligent Machines
3480 University Street
Montréal, Québec
jer@cim.mcgill.ca

ABSTRACT
In this paper, we discuss the design and testing of a reactive environment for musical performance. Driven by the interpersonal interactions amongst musicians, our system gives users, i.e., several musicians playing together in a band, real-time control over certain aspects of their performance, enabling them to change volume levels dynamically simply by moving around. It differs most notably from the majority of ventures into the design of novel musical interfaces and installations in its multidisciplinary approach, drawing on techniques from Human-Computer Interaction, social sciences and ludology. Our User-Centered Design methodology was central to producing an interactive environment that enhances traditional performance with novel functionalities. During a formal experiment, musicians reported finding our system exciting and enjoyable. We also introduce some additional interactions that can further enhance the interactivity of our reactive environment. In describing the particular challenges of working with such a unique and creative user as the musician, we hope that our approach can be of guidance to interface developers working on applications of a creative nature.

1. INTRODUCTION
According to Jordà, the fact that it is not easy to define the role of a computer in live performance the way one can with traditional acoustic instruments is an indication that we are still in the "Stone Age" of technology-aided music creation [11]. In fact, this challenge is a recurring theme when examining New Musical Interfaces (NMIs) in general, a term by which we describe novel gestural controllers, sound installations and sonic environments. When listening to traditional musical instruments, there is a concrete, visible and mechanic relationship between the movements of an instrument's body and the qualities of the resulting sound, allowing both musicians and listeners to develop a clear cognitive link between the excitatory input and the ensuing auditory output. On the other hand, digital musical instruments exhibit a decoupling between the their gestural controllers and sound generators. In fact, Croft likens this distinction to the practice of acousmatic music, asserting that "what is known to be the source is visible but remains perceptually detached" [?]. Such a phenomenon places the additional responsibility of designing a mapping between input and output onto the instrument's creator. While this can lead to a wealth of creative experimental NMIs, many are in reality peculiar, and are only used by their respective creators. Fels et al. attribute this problem to a lack of transparency, one of the qualities of mapping that provides an indication of the psychophysiological distance, in the mind of the player and the audience, between the input and the output of an instrument [5]. The more transparent a mapping, the more expressive a device can potentially be, and the more rewarding is the experience of playing for both musicians and audiences. Naturally, this begs the question: What makes a new instrument easy to learn, enticing or, most importantly, rewarding?

Many authors advocate for the use of metaphor as a driving force behind a mapping's design: by referring to elements that are considered "common knowledge", the mapping can ideally be made transparent to all parties. We wish to take this approach further, and argue that arriving at the appropriate metaphors can be best accomplished through a User-Centered Design (UCD) methodology based on the key principles of usability outlined by Gould: early and continual focus on users, empirical measurement of usage and iterative testing [8]. While the field of music technology has long benefited from research in Human-Computer Interaction (HCI), we believe in the merit of reversing this relationship: there is much insight that interface developers stand to gain from the design of New Musical Interfaces. In fact, we regard the development of NMIs as a highly specialized challenge in User-Centered Design. Musicians are unique users. Their needs can be difficult to establish, given that new instruments, controllers or sonic environments do not exist to serve a concrete purpose in the same way as, for example, a document editor. Furthermore, the nature of performance imposes strict constraints on any interaction design: musicians' hands, eyes and ears are almost always occupied. As a result, all considerations of usability design bear an added level of complexity, and many traditional input and output paradigms can be deemed unsuitable.

We chose to explore these challenges by applying the key principles of usability described by Gould to the design of a reactive environment for musical performance. Rather than introducing a novel musical instrument, our objective was to create a new performance environment that captilizes on the interpersonal interactions amongst musicians, allowing them to directly influence each other's volumes by moving about their space. Through the design, development and evaluation of our system, we realized that no single discipline provided all the answers, and found ourselves incorporating techniques from various fields into our methodology: engineering design to implement the system, HCI to methodologically involve our users, ludology to evaluate how enjoyable they found our reactive environment, and social science research to develop a better understanding of musicians. The remainder of this paper describes how our system evolved in a manner that ensured our target user, the musician, remained front and center through all phases of the project.

2. BACKGROUND AND RELATED WORKS

2.1 Collaborative Musical Interfaces

NIME'12, May 21 – 23, 2012, University of Michigan, Ann Arbor.

Ensemble performance has always been been considered a highly social and interdependent art form. In a closed causal loop, a musician's state is continuously influenced by the feedback he receives from his peers and audience and, in turn, his output will come to affect those around him. Nonetheless, while musicians can influence each other a great deal, the level of control over this influence is rather limited. For example, a soloist can steer her collaborators towards a musical idea in which she is interested, but this type of influence is more of a suggestion; she has no direct control over the other musicians' instruments, and there is no guarantee that they will consent to her desire [15].

The increasing use of computing technology in performance, however, has made possible the construct of direct electronic communication channels between instruments, allowing perfomers to take fully active roles in determining not only their own musical output, but that of their peers as well [16]. To describe collaborative interfaces where players can influence, share and shape each other's music in real-time, Weinberg coined the term "Interconnected Music Networks" (IMNs). Such networks, he posits, "bear the promise of using technology to enhance the social context of music performance and enrich its social ritual roots" [15]. This notion was popularized in the 1970s by the League of Automatic Music Composers, who became the first group to write interdependent computer compositions where frequencies were mapped from one computer to generate notes in another. The group evolved into The Hub in 1986, and improved their communication schemes through the use of MIDI data exchanged by a central computer. However, participation in early IMNs was typically not a simple process. In the case of The League, for instance, the majority of interdependent connections between players were based on low-level elements, requiring possession of specialized musical skills and technical knowledge in order to partake meaningfully in the process [16]. Furthermore, the interactions proved to be overly complex for audiences to understand.

In recent years, more approachable varieties of collaborative interfaces have emerged to offer even novices the possibility of collectively creating music [3]. For instance, Jordá's Faust Music On-Line (FMOL) allowed users to create their own compositions using a simple Graphical User Interface (GUI) before uploading them for others to access and manipulate [11]. The goal of the project was "introducing the practice of experimental electronic music to newcomers while trying to remain attractive to more advanced electronic musicians". Similarly, Barbosa's Public Sound Objects allowed participants to partake in an on-going collaborative sonic performance by manipulating simple objects in a public online space [1].

Interactive sound installations are another example of collaborative interfaces designed with public accessibility in mind. For instance, as an alternative to the often undesirable "Muzak" heard in public spaces, the Intelligent Street allowed users to request changes via mobile text messages. The overall result was to turn visitors of a space from passive consumers to active participants creating their own aural landscape [12]. As another example, the Control Augmented Adaptive System for Audience Participation (CAASAP) was a project designed to examine a variety of ways in which audience members could make use of mobile phones to become part of the music-making process [14]. Finally, Feldmier et al. created low-cost wireless motion sensors that enabled them to estimate the level of activity of a large scale crowd . The data could be used subsequently to generate music and lighting effects, thus essentially allowing members of the crowd to drive the music to which they danced [4].

As collaborative interfaces respond to the interpersonal interactions amongst participants, their design should ideally be informed by a thorough understanding of common user behaviours. We believe that this can best be achieved through UCD methodologies. In fact, a number of developers have successfully taken a user-centric approach to the design of novel musical interfaces, as we further discuss in the following section.

2.2 Musical HCI

As Gentner explains, system developers and engineers often fall into the trap of believing that an ideal interface is one that reflects a system's underlying model [7]. A user, however, generally has no interest in or understanding of a system's inner workings, but is more concerned with completing a particular task using the system. Thus, in spite of the designer's best intentions, the result from the user's point of view can become a "bad interface". Developers of novel musical interfaces are not immune to this phenomenon. Jordá identifies idiosyncrasy as the biggest problem with new musical controllers, stating that many NMIs wind up only being used by their own creators [11]. Furthermore, Geiger et al. explain that since mapping strategies for novel controllers suffer from "missing interface standards and little design experience", a "try-and-error approach" is more often than not adopted by developers [6].

The search for a solution to such issues has led to the emergence of "music-oriented HCI" research, where the development of new sensing tecnologies, creation of mapping strategies and user involvment in design are heavily driven by HCI know-how. For instance, Bau et al. relied on participatory design methods from HCI for the development of the A20, a polyhedron-shaped, multi-channel audio input/output device. Throughout the design of the A20, the authors held participatory workshops where non-technical users were invited to explore the system's potential as a collaborative personal music player [2]. Similarly, Geiger et al. employed participatory design techniques in the early design phase of the VRemin, a set of 3D interfaces for a virtual Theremin [6]. The Do It Yourself Smart Experience (DIYSE) Project is another another example of HCI methodologies used in the design of a musical interface. Johnston et al. also used participatory design throughout the development of an interactive environment that encourages musical exploration [10]. Consistent with Geiger et al. [6], the authors found "no clear pre-existing requirements for software of this kind" and therefore adopted an exploratory approach.

2.3 Applying UCD to Music

While the systems described in the previous section were designed through user-centered approaches, they were evaluated iteratively against pre-conceived benchmarks established by the developers themselves. In other words, the designers had decided *a priori* what would be important for the user, rather than starting with the necessary "early focus on users" to establish these requirements. Our approach, instead, had the first author gain a full understanding of her users through lengthy observations, interactions and non-leading interviews, all described below. As a result, all design criteria, benchmarks and goals were tailored to reflect exactly what the user, rather than the designer, found important. Furthermore, our experience shows that while a user-centered approach to musical interface design can be effective, it does not paint the full picture. Iterative user testing may tell us the *how*, but its ability to explain the *why* behind a user's actions, decisions and behaviours can be rather limited. As a result, we have drawn on know-how from a number of disciplines to arrive at a thorough understanding of our user, an understanding we believe is critical towards desgining a truly satisfying, rewarding and engaging performance environment.

3. UNDERSTANDING THE USER

3.1 User Observation

In keeping with Gould's principles, our first step was to develop a thorough understanding of our users and the tasks we expect them to accomplish. To this end, we gathered extended "fly-on-the-wall" style video footage of musicians playing together in a relaxed environment. The participants we worked with and filmed varied in terms of expertise and the length of time they had been practicing music with each other. Since our focus was not on the creation of a new instrument, but on augmenting effective group performance with *existing* instruments, we were particularly interested in the interpersonal interactions amongst the musicians rather than the musician-instrument interactions. We worked with a total of 15 different musicians in five bands, who were filmed over a period spanning a few months.

As a general pattern, musicians who had performed together for longer periods tended to interact with one another through more physically pronounced movements. For instance, they were more likely to move closer during parts of a song where they felt a desire to "groove" with one another. During such periods, they also commonly assumed similar body postures. On the other hand, musicians who had played together less frequently mostly communicated through sustained glances and synchronized head bobbing movements. Finally, we noted that adjusting volumes mid-session was often a cumbersome task for ensemble performers. Typically, all levels were tuned before the start of performance, and any desired adjustments could only be undertaken between songs, after having been discussed amongst and agreed upon by band members. During stage performance, even this level of control is removed from the musicians, and typically relegated to a soundman.

3.2 Non-leading Interviews

While the video footage provided insight regarding the *what* and *how*, we had yet to fully appreciate the *why* behind many of the musicians' actions. Thus, to uncover this type of information, we decided to speak directly to a sample of our target population. In order to avoid any bias, we decided against a question-and-answer style of discussion, opting instead to conduct non-leading interviews. We spoke with six musicians, one female and five male, ranging in age from 18 to 42. Two were jazz musicians, while the rest performed various types of rock music.

3.2.1 Content Analysis

Each interview was transcribed, and a Grounded Theory (GT) methodology was applied. GT, a common technique used in qualitative social science research, operates on the notion that theories can be extracted from a given data set, in contrast to the scientific method, where hypotheses are formulated prior to data collection. Since we wanted to avoid forming or validating any pre-conceived notions about our users' needs and desires prior to fully analyzing the content of the interviews, this made GT suitable for our needs. We began by performing a content analysis. During a process known as "coding", any quotes alluding to motivations or values held by the musicians, behaviours, or preferences were assigned a descriptive tag. After all interviews had been coded, a list of all tags used during the process was compiled. Any tags we considered sufficiently related were then grouped and assigned a new encompassing tag, thereby reducing the overall set of tags down to a more manageable size of seven: Creative Engagement, Interaction with other musicians, Improving Technical Ability, Putting on a Live Show, Self-Expression, Professionalism and Enjoyment. Subsequently, all quotes representing a particular value were grouped together, and assigned a weight between 0 and 1, representing the importance of that value to the user, or how strongly they felt about it. The weights were then summed, with the ranking of the importance of these values shown under the "Interview Analysis" column of Table 1.

3.2.2 Validation

Validation is a component of qualitative research that involves checking at some level the accuracy of one's outcomes. Since our goal was to determine what exactly is important to musicians, we decided to survey a number of them to determine how well the outcome of our content analysis procedure matched the consensus of actual musicians. Our brief online survey presented musicians with the list of seven values identified during our content analysis, and asked that they rank these values in order of importance. The survey was completed by 21 students, six female and 15 male, between the ages of 21 and 40.

Rank	Interview Analysis	Musician Survey
1	Interaction with other musicians	Enjoyment
2	Enjoyment	Self-expression
3	Self-expression	Creative engagement
4	Creative engagement	Interaction with other musicians
5	Improving technical ability	Improving technical ability
6	Putting on a live show	Putting on a live show
7	Pursuing a professional career	Pursuing a professional career

Table 1: Values of importance to performers ranked in accordance to the outcomes of our interview analysis and survey of musicians.

Although not matching exactly, the results shown under the "Musician Survey" column of Table 1 correspond reasonably closely to our initial analysis. The only significant difference is the survey ranking of "Interaction with other musicians" in fourth place, whereas the interview analysis placed this first. The values listed in Table 1 would later become important when establishing benchmarks for our formal user experiment, described in Section 6.

4. DESIGN AND IMPLEMENTATION

4.1 System Overview

Having gained a clearer understanding of our users, our subsequent goal was to begin defining some useful functionalities for our system. As a primary guideline, we wanted the system to be driven by the interpersonal interactions between musicians. Furthermore, we noted that a common problem for musicians is to balance and adjust the mix they receive, as they often wish to hear their own or a peer's instrument a bit louder than the rest for self-monitoring. During live performance, it is normally impossible for them to accomplish this in an interactive manner, or without affecting all other members of the ensemble. Thus, we anticipated that providing greater levels of control than traditional performance environments afford would be a desirable feature for our users. Finally, we wanted all such controls to be transparent and, therefore, our mappings had to adhere to a clear metaphor. As a result, we identified Dynamic Volume Mixing (DVM) as the principal feature for our reactive environment: as two musicians get closer to one another, they perceive each other's volumes to become louder.

Assume that M musicians are interacting, each located at position $\vec{x}_i \in \mathbb{R}^2, i = 1, \ldots, M$. Furthermore, assume that the musicians produce the source audio signals $s_i(t)$.

Using DVM, we describe the mix $m_i(t)$ that the musician i receives by $m_i(t) = \sum_{j \neq i}^{M} a_{ij} s_j(t)$, where $a_{ij} = f(\|\vec{x}_i - \vec{x}_j\|)$ is given by a function that increases monotonously if the argument falls below a threshold θ. Practically, an exponential function models a linear increase on the decibel scale and matches users' expectations.

Naturally, to experience such changes, musicians must receive their audio mix through headphones, which is typically the case in studio situations. DVM increases the musicians' overall sense of control by enabling them to create their own individualized mixes, a job typically assigned to a soundman during live performance. It capitalizes on common behaviours exhibited by musicians to address particular needs that became apparent to us during our observations. Furthermore, DVM is based on an exaggerated property of sound that we experience everyday: sound sources closer to us are perceived to be louder in volume. Therefore, it satisfies the design guidelines established above.

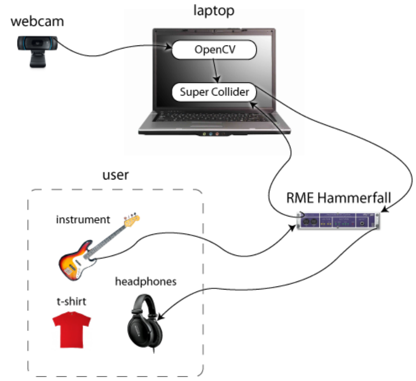

Figure 1: System Overview for Current Prototype. The dashed box titled "user" represents components given to each participating musician.

4.2 System Implementation

After painting a clearer picture of our users, our subsequent goal was to implement and validate our Dynamic Volume Mixing feature. We developed our prototype using SuperCollider for audio manipulation, and a ceiling-mounted webcam, providing its video output to the OpenCV library, for user tracking. The configuration of our preliminary prototype can be seen in Figure 1.

4.2.1 Position Detection

Our colour detection algorithm, written in C++, utilizes a number of functions from the OpenCV library. First, a Logitech HD Pro C910 wide angle USB webcam is attached to the ceiling. Then, each musician is given a bright t-shirt of a different colour (such as red, blue or green) to wear, taking care to ensure that the area seen by the camera does not include other objects containing the colours we wish to detect. We then specify a range that encompasses each colour's hue, saturation and value (HSV). Subsequently, images captured by the camera are processed by a thresholding function that returns a binary image, representing the pixels where the specified HSV values were detected. Using OpenCV, we can calculate the area occupied by these pixels (zeroeth order moment), as well as their centroids (first order moments) along each of the image's two axes. Finally, dividing each first order moment by the area yields the detected object's x and y co-ordinates. The position values are then broadcast via Open Sound Control (OSC), and can be used by SuperCollider to determine each player's invidual audio mix.

4.2.2 Audio Setup

Each musician is asked to wear a pair of Sennheiser HD Pro 280 closed headphones, which are plugged into the output channels of an RME Hammerfall Multiface II audio interface. To evaluate our Dynamic Volume Mixing function, the instruments used with this prototype were all electric rather than acoustic, and fed to the input channels of the Hammerfall audio interface. Thereby, we could ensure that the modified audio mix played back to each musician is not overshadowed by the actual sound of the instruments themselves. All audio streams are then processed, individualized mixes are created in SuperCollider, and played back through each musician's headphones. The distance between performers is continually calculated as the Euclidean distance between their respective color blob centroids, as returned from the Position Detection algorithm. If two musicians move closer to one another, decreasing this distance below a pre-determined threshold, they perceive each other's volumes as louder than the rest of the group.

5. PRELIMINARY USER FEEDBACK

We asked a jazz trio consisting of a singer, guitarist and keyboardist to test the DVM feature and give us their thoughts. At the end of the performance, the singer eagerly told us that she found the system to be very exciting, and that she had in fact moved closer to each musician during their solos in order to better "focus" on what they were playing. She added that during traditional rehearsal, she was often frustrated at her lack of control over other musicians' volumes: the sound levels, while optimal for other players, could at times be less than ideal for her. Having the ability to create her own personalized sound mix was therefore quite helpful. Finally, the musicians reported having fun while interacting with our system. Seeing as the preliminary feedback was mostly positive, we decided to conduct a formal experiment in order to quantitatively assess the effects of DVM on musical performance.

6. FORMAL USER EXPERIMENT

Our first objective was to determine the benchmarks against which our system should be tested. Since the procedure described in Section 3.2 helped uncover a number of values musicians found important, we knew it was critical to determine how well our system performed against each of those categories. Nonetheless, Hix and Harston advise that the number of usability goals tested in qualitiative experiments be kept low, citing 2-3 as an ideal figure that helps prevent testing and analysis from overwhelming developers [9]. Furthermore, only the top 4 entries could, to a certain extent, be measured during an experiment, as the remaining ones would require long-term monitoring of participants. Thus, we decided that our formal user experiment should help determine whether the Dynamic Volume Mixing feature met the following benchmarks:

- **Enjoyment:** Musicians should enjoy themselves while interacting with our system. In order to quantify enjoyment, we turn to ludology, where "flow" and "immersion" are often evaluated during game studies and have successfully been used as indicators of a player's overall sense of pleasure (see [**?**][**?**][**?**]). In particular, we used a modified version of IJsselsteijn's Game Experience Questionnaire [**?**], which evaluates a player's state through a series of general questions that can be extended easily beyond gaming applications.
- **Creative Engagement:** Musicians should be able to explore new grounds and enhance their sense of creative engagement. A questionnaire was specially created, where users were asked to rate their perceptions of the most basic components of creative engagement. Examples include whether they discovered or learned anything new, felt inspired, did something unexpected or took risks.
- **Self-Expression:** Musicians should also be helped to express their musical moods and ideas. Similar to our evaluation of creative engagement, a questionnaire was created to elicit the musicians' perceptions of basic qualities of self-expression. Examples included whether they felt understood, whether they were able to express their moods, feelings and ideas either verbally or musically, whether they felt their individuality had been preserved within the group and whether the performance reflected any aspects of their personalities.
- **Interactions amongst musicians:** One of our primary objectives, of course, is to support an increased level of interactions between musicians. In addition to the GEQ's Social Presence questionnaire component, which focused primarily on the behavioural involvement amongst musicians, we also looked at the position data collected throughout the performance to determine how often the musicians moved closer towards one another.

Musicians were asked to choose a number of songs familiar to them, and jam for approximately half an hour, once in a traditional, non-augmented fashion, and once with our system's Dynamic Volume Mixing functionality. This notion is an example of the "Adjection/Intensification" strategy described by Ravasio et al. for conducting qualitative research in HCI [13]. The idea is to isolate and determine the effects of dynamic volume control on the musician's perceptions of creativity, enjoyment, self-expression and interaction. For better comparability, the musicians heard each other through closed headphones in both cases, although, natually, the volume mix was static in the non-augmented case. Furthermore, since our experience indicates that it often takes musicians some time to gain momentum and feel comfortable, or "warm up", we did not want to interrupt them between every song to switch experimental parameters. Therefore, we only switched conditions once during the session, after approximately half an hour of performance, at the end of the currently played piece. This helped ensure that each jam session peaked and ended organically, thereby preserving the ecological validity of the performance. Although a balanced order of presentation of conditions would be superior from the perspective of a valid comparison, we wanted to make sure that in the limited time frame of the experiment, the musicians had ample opportunity to play under the DVM condition, identifying any significant usability issues that may be of concern.

6.1 Band 1

We first invited a 4-piece band, consisting of vocals, guitar, drum machine, and keyboard synthesizer to test our system. The musicians were between 22 and 28 years old, two female and two male, and had performed together in the past. However, the test session quickly uncovered a number of areas that needed improvement. The vocalist took the most advantage of the DVM feature, as she moved around the room to explore the shifting volumes. Her actions, however, seemed to frustrate the guitarist and drummer, who felt that they could not "get away" from what was happening around them. Analysis of position data helped shed some light on the issue: throughout the session with DVM, a threshold of 250 cm was set, meaning that two musicians would begin experiencing volume changes when the distance between them dropped below that value. This threshold was appropriate during our work with the jazz trio, described ealier. However, it proved to be unsuitable for a four-piece ensemble, as the initial distances between the musicians were already below the threshold when they began playing. Thus, we learned that any threshold used needed to be tailored to the size of the ensemble using the system. Furthermore, we concluded that the musicians needed a clearly marked default position to which they could return, should the volume changes become overwhelming. After these changes had been incorporated, we set out to evaluate our improved system.

6.2 Band 2

We tested our improved system with a 3-piece rock band consisting of bass, lead guitar, and rhythm guitar. The musicians were between 27 and 31 years old, all male, and had performed together in the past, rehearsing and playing live shows regularly for nearly two years. We monitored the distances between members of the ensemble during the session without DVM and found that they appeared to be comfortable at separations of approximately 230–250 cm, i.e., not trying to move closer together at these distances. Therefore, we determined 225 cm to be a reasonable threshold to set for the DVM session.

Data was collected post-session through the questionnaires, as well as in-session through video footage and position tracking. The participants were highly encouraged to think out loud, and express any feelings or concerns they had regarding their performance.

6.2.1 *Questionnaire Analysis*

Figure 2 compares the musicians' perceptions of flow, behavioural involvement, creativity and self-expression with and without DVM. The scores assigned to each factor were tabulated from the musicians' responses to questions pertinent to the various facets of that factor. Overall, the DVM feature fared quite well across the board, leading to a majority of the performers reporting an equal or improved experience with DVM on all factors.

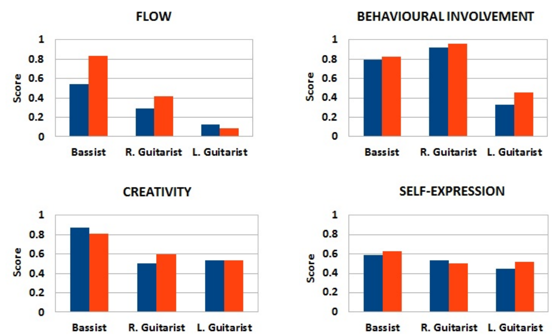

Figure 2: Comparison of levels of flow, behavioural involvement, creativity, and self-expression, reported by musicians, without DVM (in blue) and with DVM (in orange).

6.2.2 *Data Analysis*

As seen in Figure 3, during the session without DVM, the musicians did not venture far from their starting positions. The only notable exception was an instant when the rhythm guitarist briefly wandered across the performance space, before returning to his original post. In contrast, however, when DVM was used, however, all three musicians were far more adventurous, making full use of the performance space.

6.2.3 *Footage Analysis*

Through a content analysis of our video footage, we were able to gain more insight into the musicians' impressions of our system. First, even though the musicians were given a description of the DVM feature before the start of the session, they were quite pleasantly surprised when they began interacting with the system. They began by moving all around the space to "get a feel" for the volume shifts. When they were more comfortable, they started taking better advantage of DVM, with the rhythm guitarist and bassist, for instance, huddling around the lead guitarist as he played a solo, as can be seen in Figure 4. Thus, our system helped increase the level of interpersonal interactions amongst them. All three musicians reported finding the system quite novel and exciting. The rhythm guitarist commented explicitly that he had never experienced anything similar, and was quite happy to be given the opportunity to participate in our test session. Finally, the lead guitarist suggested that we use a metronome in the future, as he found it a bit difficult to keep time at certain points.

7. INTERACTION ENHANCEMENTS

The next step in the development of our system is to continue with the iterative process of testing and modification. While DVM met our established benchmarks reasonably well for the rock trio described above, we have to confirm these results by evaluating it further with other ensembles.

In addition, motivated by the positive feedback from the musicians, we have developed new ideas how to enhance the interactivity of the system further. We introduce and outline two approaches: (i) Enhanced Stereo Panning, and (ii) Orientation-based Sound Mixing.

7.1 Enhanced Stereo Panning

Enhanced Stereo Panning (ESP) takes the formalism for DVM established in Section 4.1, and extends the mix for musician i to a 2D-vector $\vec{m}_i = (m_{Li}, m_{Ri})$, representing the left and right audio channels. This allows us to create interactive, spatially structured sound mixes. As we naturally orient towards sound sources we are

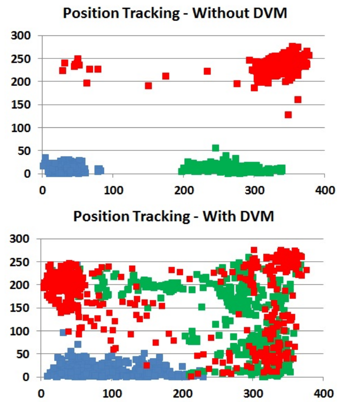

Figure 3: Overview of positions of rhythm guitarist (in red), lead guitarist (in blue) and bassist (in green) without and with DVM, sampled at 1-second intervals.

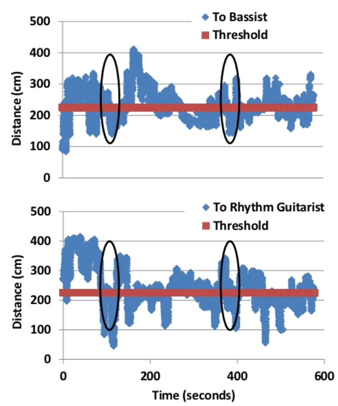

Figure 4: A 10-minute overview of the distance between the lead guitarist and the other musicians. Two instances of solos have been circled.

particularly interested in, ESP likewise enables an intuitive navigation of the mix. Other musicians deemed to be of less interest are in turn routed to one's spatial periphery, left or right, according to their position. The formalism to create this effect is to compute $m_{ki}(t) = \sum_{j \neq i}^{M} a_{kij} s_j(t)$where $k \in \{L\text{=left}, R\text{=right}\}$, and a_{kij} are now channel-wise mixing coefficients that depend both on the distances to other musicians j and the orientation of musician i. We introduce the unit length vector $\vec{e_i}$, which points from the right to the left ear. An intuitive approach is to set

$$a_{kij} = \frac{a_{ij}}{2} \left(1 + b_k \frac{(\vec{x}_i - \vec{x}_j)}{\|\vec{x}_i - \vec{x}_j\|} \cdot \vec{e}_i \right)^2, \qquad (1)$$

with $b_L = -1, \quad b_R = 1$. The scalar product between the difference vector and the ear-connection vector is within the range $[-1, 1]$ and the $(\ )^2$ ensures that the overall energy of the source signal remains constant when orienting the head towards a musician.

Certainly, the implementation of this technique demands the sensing of the musician's head orientation in real-time, which we have already tested using custom-built sensors. Although it is debatable whether this approach is appropriate for mixing, particularly when users move their heads quickly to the rhythm, e.g., "head banging", we argue that this ESP effect can be muted gradually and dynamically if the vector $\vec{e}_i$ changes too fast.

7.2 Orientation-based Sound Mixing

We can even go a step further by combining ESP and amplitude mixing into an Orientation-based Sound Mixing (OSM) feature, modifying Equation 1 so that the head rotation not only causes a stereo panning of the mixed sound signals, but also emphasizes the level according to its angular distance to the frontal position. In this scenario, the better a musician fits into an angular bandwith σ, the louder their source signal will be. For that only the coeffients a_{ij} in eq. 1 need to be redefined as

$$a_{ij} = f(\|\vec{x}_i - \vec{x}_j\|) \cdot g\left(\frac{\vec{x}_i - \vec{x}_j}{\|\vec{x}_i - \vec{x}_j\|} \cdot \vec{n}_i\right) \qquad (2)$$

where a good choice for $g()$ is a bell-shaped function such as $g(y) = b + \exp(-y^2/\sigma^2)$. The σ parameter allows adjustment of the angular width of the level emphasis area and b allows adjustment of the ambient level of sound sources outside the peak. This orientation-based sound mixing is particularly interesting for those musicians whose mobility is rather limited by their instrument, such as the keyboarder or drum player.

8. CONCLUSION

We described the development of a reactive environment for performance that allows musicians to change each other's perceived volumes dynamically by moving about their space. More than an exercise in engineering design, our approach was truly multidisciplinary, drawing on user-centered approaches from Human-Computer Interaction, qualitative research methodologies from social sciences, and evaluation techniques from ludology. Through extensive user observations and non-leading interviews, we were able to create a thorough portrait of our target users that remained front and center during all phases of development. A quantitative evaluation of the resulting system with a rock trio proved that our system has the potential to increase musicians' sense of flow, creativity, self-expression and behavioural involvement. In addition, it can increase the interpersonal interactions amongst members of an ensemble. We also introduced additional interactions, such as Enhanced Stereo Panning and Orientation-based Sound Mixing, that we plan to integrate into our reactive environment. Through the iterative testing and modification process mandated by UCD, we hope to continue refining and expanding our reactive environment, thereby offering musicians additional novel functionalities that extend beyond traditional performance.

9. ACKNOWLEDGEMENTS

The authors would like to thank Karola Pitsch of the Conversation Analysis Group at Bielefeld University for sharing her expertise, and Adriana Olmos from the Shared Reality Lab at McGill University for her invaluable help and advice. The research described here was supported by the Natural Sciences and Engineering Research Council of Canada.

10. REFERENCES

[1] Àlvaro Barbosa and M. Kaltenbrunner. Public Sound Objects: A Shared Musical Space on the Web. In *Proceedings of International Conference on Web Delivering of Music*, pages 9–15. IEEE Computer Society Press, 2002.

[2] O. Bau, A. Tanaka, and W. E. Mackay. *The A20: Musical Metaphors for Interface Design*, volume 8, pages 91–96. 2008.

[3] T. Blaine and S. Fels. Contexts of collaborative musical experiences. In *NIME'03*, pages 129–134, 2003.

[4] N. Bushnell. Relationships between fun and the computer business. *Commununications of the ACM*, 39:31–37, August 1996.

[5] J. Croft. Theses on liveness. *Organised Sound*, 12(1):59–66, Apr. 2007.
[6] M. C. Feldmeier. *Large Group Musical Interaction using Disposable Wireless Motion Sensors*. PhD thesis, Massachusetts Institute of Technology, 2002.
[7] S. Fels, A. Gadd, and A. Mulder. Mapping Transparency Through Metaphor: Towards More Expressive Musical Instruments. *Organised Sound*, 7:109–126, August 2002.
[8] C. Geiger, H. Reckter, D. Paschke, F. Schutz, and C. Poepel. Towards participatory design and evaluation of theremin-based musical interfaces. In *Proceedings of the 2008 Conference on New Interfaces for Musical Expression*, pages 303 – 306, Genova, Italy, June 2008.
[9] D. R. Gentner and J. Grudin. Why Good Engineers (Sometimes) Create Bad Interfaces. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, pages 277–282, New York, NY, USA, 1990.
[10] J. D. Gould and C. Lewis. Designing for Usability: Key Principles and What Designers Think. *Communications of the ACM*, 28:300–311, March 1985.
[11] D. Hix and H. R. Hartson. *Developing user interfaces: ensuring usability through product & process*. Wiley professional computing. J. Wiley, 1993.
[12] W. IJsselsteijn, W. Van den Hoogen, C. Klimmt, Y. De Kort, C. Lindley, K. Mathiak, K. Poels, N. Ravaja, M. Turpeinen, and P. Vorderer. Measuring the experience of digital game enjoyment. In *Proceedings of Measuring Behavior*, 2008.
[13] A. Johnston, L. Candy, and E. Edmonds. Designing and evaluating virtual musical instruments: facilitating conversational user interaction. *Design Studies*, 29(6):556–571, 2008.
[14] S. Jordà. FMOL: Toward User-friendly, Sophisticated New Musical Instruments. *Computer Music Journal*, 26:23–39, September 2002.
[15] H. Lörstad, M. d'Inverno, and J. Eacott. The intelligent street: responsive sound environments for social interaction. In *Proceedings of the SIGCHI International Conference on Advances in Computer Entertainment Technology*, pages 155–162, New York, NY, USA, 2004.
[16] P. Ravasio, S. Guttormsen-Schar, and V. Tscherte. The qualitative experiment in HCI: Definition, occurrences, value and use. *Transactions on Computer-Human Interaction*, pages 1–24, 2004.
[17] P. Sweetser and P. Wyeth. Gameflow: a model for evaluating player enjoyment in games. *Computer Entertainment*, 3:3–3, July 2005.
[18] K. Tahiroglu. Dynamic social interaction in a collective mobile music performance. In *Proceedings of the 2009 International Conference on Computational Science and Engineering*, pages 854–859, Washington, DC, USA, 2009.
[19] G. Weinberg. The Aesthetics, History, and Future Challenges of Interconnected Music Networks. *Proceedings of the 2002 International Computer Music Conference*, pages 349–356, 2002.
[20] G. Weinberg. Interconnected Musical Networks: Toward a Theoretical Framework. *Computer Music Journal*, 29:23–39, June 2005.