Designing for group music improvisation: a case for jamming with your emotions

Citation for published version (APA):

DOI:
10.1504/IJART.2016.10002396

Document status and date:
Published: 28/12/2016

Document Version:
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Designing for Group Music Improvisation: a Case for Jamming with Your Emotions

Gustavo Ostos Rios, Mathias Funk, Bart Hengeveld

Department of Industrial Design
Eindhoven University of Technology
Eindhoven, The Netherlands

g.a.ostos.rios@student.tue.nl,
{m.funk, b.j.hengeveld}@tue.nl

Abstract. During improvisation, musicians express themselves through live music. This project looks at the relationship between musicians during music improvisation, the processes of expression and communication taking place during performance and possible ways to use musicians’ emotions, to influence a digital instrument – and in consequence improvisation. To illustrate this, a three-layer model is described studying the relationship between band members and the audience as a system, where emotions, expressivity and generation of sound give shape to improvisation. Focus is applied specifically on how individual emotional arousal can be used as input to control as a group a musical instrument: EMjam. The instrument is illustrated, describing the design and implementation as well as the evaluation of it, to subsequently discuss the results.¹

Keywords: Music, improvisation, composition, instrument, performance, emotions, arousal, expressivity, electro dermal activity, skin conductivity.

Introduction

Designing for systems

In HCI we are more and more moving away from designing interactions for single users, towards designing interactions for networked groups of single users; we move from a ‘one user-one technology’ paradigm towards a ‘multiple users-multiple technologies’ paradigm. Examples one can think of are the many social networks that have emerged, but also websites such as Wikipedia, where value is essentially co-created.

Successful though these and other examples have proven to be, we find that many current interactions for groups of users are limiting in terms of human expressivity, partly due to a mismatch between how computers work (i.e., discrete) and how humans work (i.e., continuous). We see new opportunities emerge for interaction and UX designers: opportunities for interactive products that can benefit from this connectedness without the limitations of bandwidth and latency. As such these opportunities allow for going beyond mere functionality towards interactive products that embody and communicate more qualitative aspects of interaction between users. This might lead to a new digital expressivity or even a new aesthetics of interaction [2]. Through our research we want to investigate how we can better design for expressive interactions in future hardware and software systems.

¹ This paper is an extended version of a previous conference paper [1].

Copyright © 200x Inderscience Enterprises Ltd.
Ostos Rios et al.

Systems design for group music improvisation

In this, we take the angle of group music. Through a sequence of designs—one of which we describe in this article—we want to gain insight in how design for distributed systems with shared semantics [3] for novel expressive interactions can be carried out effectively, what needs to be distributed and shared, what tools might apply, and what process can guide designers towards system designs that simply work and deliver a great user experience taking the advance of embedded and networked technologies and recent developments in user interfaces into account. As a further nuance we take a different approach to systems design via group music than usual: not only human actors, but also the instruments themselves can take on a more active role.

The main reason for using the context of group music improvisation is that group music improvisation, as we know it, already embodies systemic qualities (networked people and technologies), and that it stands in a long tradition of facilitating behavior and personalization. For example, in jazz music, the rendition of a composition—which is in essence a musical design—is highly dependent on the size of the performing band, the people in the band bringing their individual musical backgrounds, heroes, styles, fortes and flaws, and even on the type of gig or venue. Nonetheless, as a system, a jazz band functions really well; the whole is genuinely more than the sum of the parts.

Transferring this phenomenon to Human-Computer-Interaction (HCI) research, we may ask: can we design for such a setting within HCI; a setting in which people can express themselves while being part of a computational, self-directing, co-creating organism? Or better: can we even go further? Can we design a network of interactive designed agents that are specifically aimed at ‘the whole’ rather than ‘the sum of the parts’? In other words, establishing the design of interfaces between intelligent system parts that allow more or less direct users or even more remote stakeholders as a new paradigm in the overall context of interaction design [4].

We foresee that designers will face several sub-challenges in order to tackle the main one: First of all, as this is a novel design area a team of designers can probably only achieve their goal by behaving as a system themselves. For example, as the term ‘music’ is extremely wide, the designers will have to find a way to converge on musical preferences, instruments, skills, performance-concepts, through mechanisms of self-organization. From a research perspective this provides valuable, first-hand insights into mechanisms for designing for systems. Furthermore, systems design is considerably complex, by acting as a system, this complexity can be explored and ways to cope with this can be learned and adopted while designing.

Secondly, looking at the structure of bands—and more prominently orchestras—the resulting group music improvisation system will most likely incorporate multiple feedback-loops, i.e., loops of action and response. The most probable feedback-loops in a musical system include the individual ‘instrumentalist’ loop (i.e., a musician playing his instrument and adjusting his playing to what he hears himself play), the sub-group ‘section’ loop (e.g., a drummer and bass player adjusting their individual playing to each other’s playing for the sake of the section’s performance) and the band-sized ‘whole’ loop (e.g., all players adjusting their individual playing to everybody else’s for the sake of the whole band performance) [2]. Having these different feedback loops in place determines how the influence or control of an individual node in the system will (1) be able to shift forms as a consequence of being in different musical roles and (2) be perceived (and appreciated) by the audience.
The feedback loops referred to above are largely unidirectional and depend highly on social processes: interaction between musicians usually happens non-verbally using body language, gestures, mimics, and such, directed by listening to other instruments and watching other players play their instruments. In this process any player can only perceive sound coming from another player’s instrument, but only influence it indirectly by influencing the other player. However, we are interested in a situation where players can also influence the instruments of other players, as such taking a step beyond ‘multiple players playing a modular instrument’. Modularity is taken to a higher level, in which any constellation of players and instruments form a self-organizing (now: musical) body. The network of connections between players and instruments becomes denser and entirely bi-directional: players interact both with their own instruments, other players, and other players’ instruments. The result is a more unpredictable interaction, which necessarily adds to the complexity of the performance. In addition, this interaction mode changes how players perceive the setting of the performance: there is less a strong and exclusive bond between players and their own instruments. Instead, instruments become shared artifacts—and the performance opens up.

Thirdly, at a more personal level, we face the challenge of enabling musicians to communicate emotionally through music, which has the focus of the case study presented in this paper, while being part of a networked musical setting; in other words: how can musicians express themselves in a situation where other musicians can access your instrument? Emotions in musical performance

The case we describe in this paper, called EMjam, builds on the construct that when paying attention at a concert it is possible to see performers’ expressions; a guitarist playing a solo and reaching a peak at a certain point of it; a bass guitar player following with his face the lines played; a drummer making accents with the whole body; and in addition to this, the audience responding to the performance. At these moments we can see that making music is not only playing the instrument. There are factors that have an important role on performance; individual expressivity of the emotional experience through physical gestures allows musicians to communicate their intentions and express their emotions towards others. In the same way, interaction during performance has an effect in the arousal of emotions on performers and listeners [5]. Even though the process of making music can be individual, meaning that one person plays one instrument, in a band it is, as mentioned earlier, composed by a group playing different instruments. Hence, group members influence each other during performance, either in a physical way through non-verbal communication or through expression of their emotional state. EMjam builds on this as it is a musical instrument that uses individual variations of musicians emotional arousal during group improvisation, to generate a parallel second layer of sound as a result of every individual input [6].

Emotions are an important part of music improvisation [7, 8], and musical improvisation in connection with psychology [9, 10] or brain research [8, 11, 12] are active areas of research. Musicians play a unique composition that is created in its final form at the specific moment of improvisation, resulting in a single musical design that would sound different on other sessions [13]. Although many musicians draw on patterns and “trademark” sequences of their own when improvising, the composition happens in place, and musicians can externalize the emotions felt at that time. The use of emotions in improvisation to control digital instruments through sensor technology can offer new opportunities that are yet to be explored in the field.
The work described in this paper focuses on the process that requires musicians to communicate their intentions and express their emotions towards band members and audience, by making use of physical gestures and consequential computer generated sound. The work as such builds on the fact that, while we can use traditional instruments to make sound, there are other possibilities that can add expressivity to the performance, such as the manipulation of sound through physical gesture to control digital instruments [14, 15]. This project goes beyond the use of physical interfaces and musicians’ expressions to control sounds, exploring the possibilities of using individual physiological measurements of emotional components, such as arousal, to be used as an input for a digital musical instrument controlled by the whole band; giving musicians a second layer of improvisation and expression.

This paper is organized as follows: first, the connection to related work is established, after which the general theoretical approach to musical improvisation shall be outlined as a framework of three intertwined layers: first, the musical layer as the process and result of improvisation itself, a communication layer reflecting the expressions that happen continuously between musicians and the audience, and, finally, the emotion layer which is strongly related to the individual experience of musician and audience. These three layers are generally hard to separate and distinguish, however, in the context of this work this is necessary to underline the role of emotions and measurable bodily reactions in the generation of new music. In the following section, we move from the rather theoretical point of view of musical layers towards the EMjam concept, which implements a new musical instrument that captures bodily signals corresponding to emotional arousal during an improvisation session and turns these signals into sounds that blend into the overall improvised music. Practically speaking this means that part of the music originates from musicians playing their instruments, and part from physiological measurements translated into ambient soundscapes that move between foreground and background. The EMjam concept is evaluated using a working prototype as described in the evaluation section, after which the paper concludes with a discussion and an outline of future steps.

**Related work**

To understand the underlying concepts behind the design of EMjam and the approach that was taken, it is important to review a number of aspects within the areas of expressivity and emotions. These aspects represent the foundation of the theoretical research that was later on translated in the form of an experienceable prototype, following a research-through-design approach.

**Expressivity of musicians**

Performing musicians communicate emotions by means of musical instruments in a non-verbal way\(^1\), through a variety of modalities. For example, according to Camurri et al. expressivity suggests a process in which musicians externalize their emotions through “expressive gestures” [16] that contain expressive content and convey information related to the emotional domain such as: feelings, moods, affect and intensity of emotional experience; either in an acoustic or visual modality. Gestures can be expressive not only through physical gestures of performer, but include also the produced sound and visual media. Finding new ways of controlling or influencing music using a performer’s expressivity has been addressed in the past [14, 15], introducing a new layer of

---

\(^1\) Even vocalists use gestures and other modalities to better express their emotions.
interaction in the context of musical performance. These approaches focus on proposing new instruments and systems, as well as influencing the sound generated by traditional instruments through the use of a musicians’ physical gesture. Furthermore, new applications start to focus on the inclusion of emotional responses through physiological signals for the control of digital instruments. Related work can be seen in [17] where physiological signals (electroencephalogram, electromyogram and heartbeats) are used to control sound in order to build a biological driven musical instrument. Another example is the BioMuse system [18], which uses on-body sensors to measure motion and physiological signals to determine emotional states integrated into what is defined as an Integral Music Controller. The BioMuse is based on a model describing the number of existing interface devices for controlling a digital musical instrument [19]. Its importance is the addition and consideration of emotions as a controlling interface, which introduces a new feedback loop within musical performance. Additionally, the BioMuse system covers three layers of feedback that can be achieved in musical performance:

- Layer 1 represents the internal emotion and thoughts of the performer.
- Layer 2 is the physical interface layer.
- Layer 3 represents the creation of music as a consequence of the gesture.

Here, thoughts and emotions are represented through physical gestures that are used to control a sound generator. The resulting sound and interaction create a direct feedback loop. It also considers emotional state and physical gestures from the audience and the possibility to be used to manipulate the generated sound. However, the model proposed by Knapp and Cook [19] focuses on individual feedback loops and interaction with a controller and does not consider the influence of other band members during performance. Interaction within the band is important for musicians and has an influence on improvisation [20], and looking at the structure of bands, a band incorporates several feedback loops, e.g., an individual loop, a band loop, and an audience loop; demanding musicians to be aware of different settings at the same time [13]. Designing new musical instruments for a band behaving as a system requires a different approach: understanding the organization of musicians during music improvisation sets a starting point to develop new musical instruments for a band behaving as a system. In consequence, it is possible to propose systems aimed to create richer interactions that could influence feedback loops between musicians. Our approach explores the use of emotional components, specifically arousal to control a digital musical instrument as a group, offering band members a tool to help them reinforce the communication of emotions.

**Theories of emotions**

Authors might disagree when defining emotions, but they do agree on emotion components. According to Schachter’s theory of emotions [21], an emotional state can be seen as the result of the interaction between two components [22]: physiological arousal and cognition about the arousing situation. Physiological arousal determines the intensity of an emotion but not the quality of the emotion itself. The cognition determines which emotion is experienced. Both are considered necessary for the appearance of an emotional state.

Within the study of emotions, several approaches and definitions can be described. According to Frijda [23], emotions are elicited by significant events. These events are of relevant significance to persons when they touch upon one or more individual concerns or goals. To Frijda, emotions are then “the result from the interaction of an event’s actual
or anticipated consequence and the subject’s concern” [23]. Emotions emerge in a process including the occurrence of events, the presence of concerns or goals that make the event significant to a person and the appraisal of whether concerns are or are not accomplished in the interaction.

Sloboda and Juslin [24] describe how an emotion being a scientific construct is inferred from three kinds of evidence: (a) self-reports including adjective checklists, rating scales, questionnaires or free descriptions; (b) expressive behavior including facial expressions, vocalizations and body language such as gestures; and (c) physiological measurements including heart rate, respiration, skin conductivity, muscle tension, electrocardiogram, blood pressure and electroencephalograph. Based on this kind of evidence, Sloboda and Juslin [24] mention a study, which shows a consensual definition of emotion based on a review of different author’s definitions. Here, emotions are considered as “a complex set of interactions among subjective and objective factors…” [25], including aspects such as the rise of arousal, generation of cognitive processes, physiological adjustment to the arousing conditions, and behavior “that is often, but not always, expressive, goal directed, and adaptive” [25].

Models of emotions

Laurier et al. [26] make a distinction of works presenting three different ways that have been used to study emotions in relation to music: (1) The categorical representation is based on basic emotions such as happiness, sadness, anger, fear and tenderness. (2) The multi labeling approach makes use of a wider set of adjectives to describe categories of emotions. (3) The dimensional representation models emotions in a two dimensional space. On the approach of dimensional representation, Russell’s Model of Affect [27] has generated the most research [28]. This model represents affect (an observable expression of emotion) into a two dimensional space. Here, arousal represents the intensity of the emotion on a vertical axis (arousal – sleep). On the horizontal axis the valence (pleasure – displeasure) represents the quality of the emotion. One of the strengths of this model and its use on the music context is that it represents a simple way of organizing emotions “in terms of their affect appraisals (pleasant or unpleasant) and physiological reactions (high or low arousal) [28].

Our work focuses on a dimensional representation by measuring musicians' emotional arousal to use data as a controlling input for a digital musical instrument. The way musicians experience improvisation gives place to individual emotional states, which can vary within the session. Improvisation is a dynamic process that fluctuates between soft and relaxed compositions to more active and explosive ones as a “fluctuating wave that rises and falls” [29], which can affect also musicians' arousal. In this sense, arousal being a physiological component can be measured; in order to relate the intensity of musicians' emotional state to the improvisation session.

Music as an elictor of emotions

Studies have linked music to emotions, showing that music has the ability to evoke emotions in the listener, and that the qualities that allow this have been used by composers and performers to convey emotions to listeners. According to Zatorre [30] music is capable of eliciting not only psychological mood changes (e.g. anxiolytic effect of relaxing music), but also physiological changes that mirror the changes in mood. Authors have described the link between emotion and music proposing important theories such as Cooke’s melodic cues theory [31], stating that melodic intervals and melodic patterns have perceptive characteristics capable of evoke specific emotions. Other authors (Kivy [32], Langer [33], Meyer [34] and Juslin & Västfjäll [35]) have proposed in the
same way theories that are of importance for the understanding of how music can elicit emotional responses.

**Research approach**

For designing group improvisation systems as outlined above we adopt a stepwise approach, sometimes referred to as research-through-design [36, 37]. Research-through-design is an approach towards design research where design relevant, scientific knowledge is generated by iteratively designing, building and experimentally testing experiential prototypes. These prototypes can serve both as a vehicle that serves the experiment, and as the object at which the research is aimed. In subsequent iterations typically both the design research question, as well as the design itself increment in fidelity and validity.

As said, one of the drivers for our research is the interest in whether we can design a setting in which people can express themselves while being part of a computational, self-directing, co-creating organism. A nested question is of course on what such a setting should encompass. We approach these questions by designing experienceable design concepts of interactive, intelligent musical instruments or music modifiers that are specifically designed for group music improvisation. While there are certainly more abstract and methodological aspects regarding the facilitation of feedback processes in distributed environments, this paper elaborates on taking theory to a more practical level, towards working prototypes and demonstrators—particularly physical, expressive interactive products. These products may be aimed at musicians, entire music groups or even the audience. Important is the combination of rather theoretical research and practicing it in the form of “experienceable” prototypes.

As the second step of the approach, findings derived from designing and testing the prototypes in realistic settings are taken into another iteration both at research and design level. This second iteration aims more at the involved semantics and explicates them for the general design of systems, slowly detaching from the original use case and thus resulting in more generalizable outcomes, which again is part of testing and analysis. A final third iteration is then used to develop the findings further into information products for dissemination, publishing and sharing with the design community.

In this paper, only the first step will be in the spotlight: designing a system (EMjam) for group improvisation with additional musical layers that correspond to the group’s performance and the players’ individual emotions (state of bodily arousal as measured from skin conductivity sensors). Our attention in this first iteration goes to getting insight in: (1) the design process when designing for systems; (2) physiologically sensing emotional arousal for a musical mapping, and the validity of this mapping; (3) expressivity and richness versus control, and; (4) the participants’ perspectives on the resulting design (EMjam). In other words, in this iteration we aim to generate insight in both the process leading towards the artifact, as well as about the artifact itself. We come back to this in the Discussion at the end of this paper. The second and third iterations are planned in the near future. In the first-iteration work we describe in this paper we build on the notion that music improvisation requires musicians to communicate their intentions to each other and to express their emotions to the audience on different layers. Even though the process of making music is individual, in a group improvisation is built together by musicians influencing each other through communication of intentions and expression of emotions. We explain this in the following, illustrating three such layers.
The sound generated (first layer) during group improvisation is influenced by the way musicians express and communicate with each other and the audience (second layer). In the same way, the emotions (third layer) felt during improvisation are a result of the overall experience, influencing how musicians express and communicate. The layers of expression and emotion represent a new and challenging area for the design of musical interfaces.

The first layer (see Fig. 1, the layer labeled “sound”) is represented by the music itself. Improvisation has a “hidden communication” [38] where musicians set what they want listeners to hear. Individual interaction with their instruments gives as result an individual generated sound that builds the group improvisation. The second layer is represented by the communication and the bodily expressivity (see Fig. 1, the layer labeled “expression”) of musicians towards the band and the audience. Here, musicians agree when to start, stop or change aspects of the music, e.g., the volume. By using body language such as movements, facial expressions, reactions and signs, they can show their intentions. In addition to this, emotions are an important component in music improvisation, as musicians’ in-the-moment experiences are reflected into sounds. Hence, the third layer is settled at an emotional level (see Fig. 1, the layer labeled “emotion”). Musicians experience the music they make at a point that they could be “surprised or aroused by an any musical line exposed to the ears” [39]. Therefore, expressivity is influenced by the emotional state of musicians and it will also influence improvisation. Other external influences like the engagement of the audience with the band might also have an effect on the response of musicians.

All three layers together form a process where musicians play together giving shape to improvisation. The emotions evoked during improvisation have an effect on the way musicians communicate their intentions and express their emotions, which will in consequence affect the sound generated by every band member.

Led by this three-layered approach to interaction during music improvisation we applied the following process approach: As a first exploration, the focus in on understanding whether activities on the third layer (emotion) can actually influence the improvisation. Experiences are felt by all participants of music performance, both
Designing for Group Music Improvisation

musicians and audience members; however, the ways to communicate these experiences are comparatively limited, depending on the setting of the performance. On the one hand, musicians will, for instance, not always communicate their feelings corresponding to the experiences, as that might counter the desired mood of the music (“flute player enjoying the performing of a sad piece of music”). On the other hand, audiences cannot always communicate their personal emotions, as they are sometimes far away from the musicians or the musicians are hidden out of sight (as in many classical theatres). Within the group of musicians, there are mostly possibilities to submit more or less subtle social cues between players that convey a sense of the perceived experience and traces of emotions. While these can indirectly contribute to the overall performance by influencing the individual players, it is not yet easily possible to directly use emotions and map them to sound or music.

The approach has led the central concept of this paper: EMjam. Different components of emotions can be used to control an instrument in the context of music improvisation, which opens new possibilities for a new musical aesthetics and expression. Could we use emotional arousal to generate a ‘second layer’ of improvisation? An underlying sound generated and controlled by the band as a system, where one instrument receives input from every musician. As mentioned before in the models of emotions, arousal represents an important component of emotions and its variation during time can be monitored. In the same way, during music improvisation, different emotional states can be elicited. Regardless their quality (pleasure or displeasure), emotional states may vary in intensity or arousal and can be observed on musicians expressions during improvisation. EMjam uses variations of musicians’ emotional arousal during performance to control the digital instrument, to give musicians an extra level of expression that adds to the overall experience of improvisation. EMjam is discussed in the next section.

**EMjam**

**Concept**

In the context of band improvisation, musicians play their own instruments (for instance, guitar, bass guitar, and drums) and by doing so create the main layer of music (see Fig. 2, main layer shown in blue); emotional arousal during the performance can be used to control a digital musical instrument creating a second layer of music (see Fig. 2; shown in red). This second layer is independent from the improvisation itself, meaning that sound coming from the traditional instruments is not affected directly by this layer (as a sound effect would do), but the second layer creates a new input. Additionally, the second layer is defined by the interaction between band members, in a way that the emotional experience of musicians while playing define the sound generated. In this sense, musicians’ arousal can be reflected into a second layer of sound and improvisation itself represents the interface for playing. In this way, expressivity of the band is reinforced by an individual input based on their emotional arousal as a group. Since this concept builds on basic human bodily reactions that can be picked by fairly standard sensors, it could potentially be extended to a large group of musicians or even members of the audience. In this paper, however, we will limit the scope to three musicians, thereby not only reducing the technical complexity, but also the cognitive complexity of listening and potentially reacting to the output of the EMjam music instrument.
Fig. 2. Two layers of improvisation. Individual musicians’ interaction with their instruments results in the main layer of improvisation (blue). Individual musicians’ arousal results in an underlying improvisation (red). The performance of the group during improvisation provides the interface for the control and generation of the underlying improvisation.

**Design overview**

EMjam consists of a set if three sensors, a signal controller and a digital musical instrument (Fig. 4).

Fig. 3. Overview of the main components of EMjam. Variations in arousal coming from the sensors are sent to the signal controller holding an Arduino microcontroller and controls for the keynote, volume and tempo. Here, measurements are converted from analog to digital and sent to digital instrument on the computer.
Designing for Group Music Improvisation

On the physical side, three wristbands (Fig. 3, left) hold skin conductivity sensors that each measure individual changes in skin conductivity, which is mapped to “arousal level”. The sensors are connected to the signal controller containing the electronics (Fig. 3, middle) and holding knobs for the control of keynote and volume, and a tapping surface for the control of tempo. Measurements from the sensors are received by an Arduino microcontroller inside the signal controller and then sent to a connected external computer through USB. The digital musical instrument (Fig. 3, right) then converts variations in measured arousal into MIDI messages to control specific parameters in relation to rhythm, melody, and harmony of an instrument setup (see below for the digital musical instrument).

**Fig. 4. EMjam.** Iterative development process of EMjam (top). Final prototype (bottom): Signal controller (bottom-left) for the control of keynote, volume and tempo. Wristbands (bottom-right) holding skin conductivity sensors to measure changes in arousal.

**The sensors**

The sensors of EMjam are designed to fit comfortably around the player’s wrist and to be worn for an extended time (30 minutes to 1 hour). In this interface, arousal is measured through Electro Dermal Activity (EDA). EDA refers to the changes of the skin’s electrical properties when activated by various brain circuits [40]. When experiencing emotional arousal, cognitive workload or physical exertion, the brain sends signals to the skin to increase the level of sweating. Even though sweat is not felt on the skin, the pores begin to fill, increasing electrical conductivity of the skin. The most common measure of EDA is skin conductivity, which captures changes on emotional arousal. It increases when a person is more aroused: engaged, stressed or excited; and tends to stay low or drop when a person is less aroused: disengaged, bored or calm [40]. Even though different sensors were considered (EEG, EMG and ECG) to measure
different components of emotion, we chose to use skin conductivity to measure musicians’ EDA to control EMjam for the following reasons:

- possibility of measuring arousal with a non-invasive and comfortable sensor, to allow musicians to move freely on stage while improvising, without constraining physical gestures;
- linear relationship between measurements and arousal (high measurement – high arousal) [32];
- availability of components to build the sensors cost-efficiently and without large production overhead,

To control the instrument, three sensors were used (one for every musician) and connected to the signal controller for the control of keynote, volume and tempo.

![Schematics of the circuit including the sensors and the signal controller.](image)

**Fig. 5.** Schematics of the circuit including the sensors and the signal controller.

**The signal controller**

On this part, a wooden case holds the circuit and the Arduino microcontroller that receives input from the skin conductivity sensors, the keynote and volume knobs and the tapping surface controlling the tempo. To process measurements coming from the circuit connected to Arduino, a standard Firmata [41] sketch is uploaded to Arduino so the microcontroller can communicate with other software on a high level. Firmata is used to reduce the complexity of programming in several different environments and languages. The input from the sensors, knobs and tapping surface is converted from analog to the digital domain and finally sent through USB communication to a connected computer for its processing in the digital instrument.

**The digital musical instrument.**

The measurements transmitted from the Arduino are processed within the Pure Data [42] environment, where digital readings of skin conductivity and changes in volume, keynote and tempo coming from Arduino (in the range of 0.0 to 1.0), are then converted into MIDI messages (in the range of 0 to 127). Additionally, it is possible to monitor the readings of the sensors on the PureData sketch in real-time, which is useful for development and testing. Then, the MIDI messages generated are sent further to an Ableton Live¹ instance, where every MIDI message corresponding to a sensor is mapped to different parameters of a digital musical instrument, allowing musicians to control

¹ Ableton Live: http://ableton.com
Designing for Group Music Improvisation

them in accordance with their arousal level. Ableton Live is a powerful Digital Audio Workstation based on loops and different tracks in which loops can be placed and played. Since a few years, Ableton also allows for more sequencer-type usage, involving digital audio instruments or samplers, as well as a large range of digital audio effects to be applied to tracks and instruments. The feature set, availability, learning curve and ease of use made Ableton Live a good choice for the sound generation in the context of this project.

Fig. 6. Physical interface (left side) and digital musical instrument (right side)

Parameters of the digital musical instrument that is, a selection of digital instruments and effects within Ableton Live, are mapped in a linear way to each sensor, using the following criteria derived from literature [43, 44]:

- **Start from scratch:** the instrument should be designed such that knowledge about pre-existing instruments will not limit or constrain the design.
- **Instrument autonomy:** how much will the instrument do without musicians input?
- **Flexibility:** do as much as possible with few knobs as possible.
- **Interface optimization:** in this case “more is not more”. The number of controllers needed to control a setup reaches a peak at one point, beyond which it becomes noise.
- **Control over important aspects of the track.**
- **Map sensors in a linear way** (values from 0 to 127) to different parameters that are related to rhythm, melody and harmony.
- **Open:** to make it random yet controllable.

Based on these criteria, the resulting digital musical instrument consists of the following MIDI Effect Racks or compilation of sound effects (Fig. 7 and 8):  

**Rhythm: Generating Notes.**

The function of the random MIDI generator is to control the rhythm of the triggered notes. The generator includes a few effects: one Arpeggiator to affect subdivisions of the notes, two Velocity controllers to add accents and cut off the notes and a Note Length controller that determines the duration of sounds generated by this component.

**Harmonics: Using a Major Scale.**

This rack defines the musical scale in which the instrument will trigger notes. It includes a Random effect that triggers a note randomly, a C Major scale which constraints the pitches of notes that will be played to the C major scale and a Chord effect that creates cascades of chords of the notes triggered. This rack is important to align the harmonics of the generated music with the rest of the improvised music played by the
three musicians. If such a component would not be in place, musical expression would be very limited or results in unwanted disharmonic noise.

**Sound: Using an Synthesizer Instrument.**

This Rack assigns an instrument to the notes played and it is a combination of a *Muted 1 Pure* instrument, which gives definition to the notes and an *All Alone Pad* that gives Echo and ambience to the setup. This instrument will create sounds that blend in well, and will not attract too much attention – and at the same time, will not cause listeners fatigue, which is important if such a setup should be used in longer music sessions.

**Fig. 7. Digital musical instrument.** Sensor measurements coming from the Arduino microcontroller are received in the PureData software, where they are converted into MIDI messages and then sent to Ableton Live to control the digital musical instrument.

This digital musical instrument setup is only one example of a mapping of arousal measurements to sound. This setup was created with specific criteria in mind (as shown above), however, other configurations derived from a different balance or set of criteria will result in a different sound. In the same way, other digital musical instruments can be designed to tailor the resulting sound output better to more specific music styles. In the following, the selection and configuration of sound generators and sound modifiers will be described by showing the mapping from skin conductivity to musical parameters.

**Mapping of measures to the digital musical instrument**

Individual effects that are included into an Effect Rack have several parameters that can be controlled and help to define the sound. For instance, changes on the synced rate (see Fig. 8) parameter of the *Arpeggiator* will define how many notes will be triggered in a compass. Parameters are arranged into the categories based on basic musical concepts of rhythm, melody and harmony, which are then assigned to every sensor. In this sense,
Designing for Group Music Improvisation

every musician can control parameters on a defined category, which assures the richness of the performed improvisation.

The skin conductivity sensors were mapped to Ableton Live parameters in the following way:

- Sensor 1 (worn by musician A): rhythm.
- Sensor 2 (worn by musician B): melody.
- Sensor 3 (worn by musician C): harmony.
- Potentiometer 1 (in the signal controller): control of keynote.
- Potentiometer 2 (in the signal controller): control of general volume.
- Button (in the signal controller): tap to control tempo.

![Fig. 8. Mapping of sensors, tapping surface, keynote and volume knobs to the digital musical instrument.](image)

Variations in arousal are mapped in a linear way to Ableton Live, meaning that an increase or decrease in individual arousal will have a proportional effect on the parameters that are assigned to each wristband. In addition to this, the signal controller offers musicians the opportunity to control the keynote in which improvisation will take place and the volume of the instrument.

By tapping on the tempo controller, musicians can control the tempo of the sound generated on the instrument, giving them the opportunity to adapt the instrument to their own tempo instead of trying to follow the one settled of the instrument.

**Evaluation**

**Setup**

The final prototype including the sensors, the signal controller and the digital instrument, was tested with different groups of musicians. The main goal of this study was to detect patterns in the way EMjam could be used in relation to the second layer of improvisation, in order to get insight in the way musicians react and interact with this second layer.

The study consisted of an improvisation session with three musicians provided with the instruments: a djembe, a bass guitar and an acoustic guitar or a keyboard. Participants were presented to EMjam and the functioning of the prototype was explained. For the purpose of this study, the wristbands holding the sensors were specifically assigned to musicians to have control on the variables; however, they can be assigned arbitrarily to each musician. Each participant received a wristband connected to the signal controller, assigned in the following order:
Ostos Rios et al.

- Djembe: wristband controlling the rhythm.
- Bass guitar: wristband controlling the harmony.
- Guitar/keyboard: wristband controlling the melody.

Participants were not informed beforehand which of the parameters they were assigned, to identify if they were able to detect which part of the instrument they were controlling. After the improvisation session, a short discussion provided information about the impressions of participants on the use of EMjam during musical improvisation. The sessions were video recorded for its subsequent analysis through observation, making annotations on subjects related to the engagement of musicians with EMjam and different reactions towards it, the effect of EMjam on the interaction with other musicians and the general effect of the second layer of sound on the improvisation session.

Fig. 9. Evaluation of EMjam during group improvisation.

Participants

During this study, three groups of musicians performed in an improvisation session using EMjam:

- Group 1 and 3: Consisted of three musicians playing a djembe, bass guitar and keyboard (group 1) or acoustic guitar (group 3), who had never played or made improvisation together; yet they had done improvisation in other occasions mainly based on rock and blues structures with their own bands, so they had a bit of experience in this activity.

- Group 2: Consisted of three participants playing djembe, bass guitar and acoustic guitar, that had played together before, with less experience in improvisation. They had practiced before sequences based on rock structures.

Results

In the first group it was observed that the use of EMjam influenced the way in which musicians built up improvisation as well as the way they interacted with each other. During improvisation, participants were paying close attention to the sound generated by EMjam, noticing changes mainly in tempo: "at a time, it suddenly became faster". The sound provided them with a base that was used as inspiration to build a new improvisation. After improvisation gained structure, the sound became an ambience layer giving musicians a space to play further. We noticed that, while trying to keep attention to the generated sound from EMjam, participants were less conscious of what other
Designing for Group Music Improvisation

musicians were playing. When they were able to find a line to improvise, attention could be paid to each other. In this sense, responding to the sound generated requests attention from musicians, influencing the communication between them.

In the second group it could be observed that EMjam was used as a background layer for the improvisation; musicians were playing their own structure and next to it, the EMjam generated sound was audible: "I heard something in between, like water drops". As they had played together before, it was noticeable that they had already prepared a basic line to play, which left entirely the sound from the digital musical instrument as a parallel second layer of improvisation. Musicians paid more attention to each other during improvisation instead of the sound being generated, trying at times to adapt their own volume and tempo to the generated sound. However, they took a moment to explore the sounds without doing any improvisation in order to understand the changes. Two of the participants did deep breathing and jumping, which led to changes in the generated sound and showing surprise: "the pitch is higher now!". During the discussion, one of the participants provided specific feedback on the sound generated by EMjam, assuring that "you know I am pumped up because the pitch is high". Concerning this, hearing the changes on the digital instrument gave musicians awareness of their emotional arousal, as they were able to hear the changes coming from EMjam. In this sense, a more active generated sound would be an indication that the emotional arousal of the group is increasing, as mentioned by one of the participants: "it makes me a lot more aware of my personal well being and also of others".

In the third group it could be observed that EMjam provided a starting point for musicians to build their improvisation. This group took a moment to hear and understand the characteristics of the sound generated by EMjam, to progressively start improvising themselves on it. Noticeably, the group explored the possibilities of the use of the keynote knob and the tempo pad. By changing the configuration of EMjam, the resulting generated sound provided the musicians with a different composition that helped them to build new sets of improvisation on different keynotes. After musicians settled a line for improvisation they were able to keep adding music individually, and in consequence, the sound from EMjam was relieved to the second layer. In an attempt to have more control over EMjam, one of the participants tried to relax by controlling his breathing, which affected the sound generated.

The manipulation of the signal controller was done mainly by the bass player, but it was not exclusive and the other musicians tried at moments to explore the controls. Similarly to group 1, the engagement of this group shifted from being fully focused in EMjam while trying to propose ideas for improvisation, towards paying attention to each other and the group improvisation. During the discussion, one of the participants assessed the interaction with EMjam, mentioning that "when we were playing and not paying attention, but hearing in between, I reacted better to the sound". This shows that focusing entirely on the digital instrument increases the difficulty of the improvisation: "if I try to play with it (EMjam), it becomes difficult". The group found a way to interact with EMjam, building up improvisation based on the sound generated mainly as an accompaniment. To do so, one of the participants found that "it is stronger if you can hear the sound in the background", using the sound as inspiration.

In general, EMjam intrigued musicians, as the sound generated was a result of their own performance. Musicians had not much control on the individual input of the instrument besides the modification of the keynote, tempo and general volume. For participants, it was not clear which individual parameter they were controlling (rhythm,
Ostos Rios et al.

melody or harmony) with the wristband they were assigned. However, they were able to notice changes in the sound generated by EMjam, mentioning that "it is really like you feel it changing". Changes in rhythm and pitch were specially more obvious, and were associated with changes in emotional arousal. In this sense, by making an effort to relax or being more active, the sound from EMjam offered them awareness on the emotional arousal of the group.

Discussion

In the discussion, we touch upon the design process when designing for systems, sensing emotional arousal and validity of mapping, expressivity and richness versus control, and, the participants’ perspectives on the EMjam musical instrument concept.

The design process when designing for systems

During the design process we experienced that several complexities arise when designing for multiple users using multiple technologies at the same time, especially in a context that capitalizes on expressiveness and unpredictability. These complexities differ from those we know of traditional design challenges where single users use single products. A mechanism through which designers can get grip on these complexities is that he places himself into the system context to be designed—in this case, music improvisation in groups. Experiencing the design/research context from a first person perspective helps empathize with it and puts the arising challenges in perspective.

Furthermore, we experienced that designing for the band as a group required an approach where the whole structure of musical feedback loops in group improvisation settings [2] needed to be studied, rather than merely designing another interactive instrument. The tradition in this is long, varied and rich. Whereas most other instances in our group music improvisation system studies follow the path of breaking into the first feedback loop (i.e., intervening in the feedback loop originating from the interaction between the musician and the instrument directly interacted with), this feedback loop remains untouched in the EMjam concept. Instead, another layer corresponding to the emotional arousal of the group of musicians as a whole, is added at the collective level—as an additional, albeit disembodied, musical actor. This triggers the interesting thought that systems such as EMjam allow for a new kind of improvisation: one where even single musicians would never play alone.

Measuring and mapping emotional arousal to (generative) music

Secondly we should discuss the use of emotional components in the field of music improvisation, which represents a new feedback loop within music performance. It relates the emotions felt by musicians at certain points of the session to the sounds generated, offering a new way of expression. Arousal refers to the intensity of an emotion and valence to its quality (pleasure - displeasure), and both together are components of emotion. We used the arousal component as part of musicians’ emotional state to control EMjam, in order to monitor the variations on the intensity of the emotions felt by them during musical improvisation. It is however not entirely clear how well EDA measurements and skin conductivity can be attributed solely to emotions, as they are bodily reactions that can also be contributed to other factors. Among the advantages of using emotional arousal it can be mentioned that: the construction of the sensors can be made with commonly available components and measurements provide a good reference of skin conductivity, which is associated with emotional arousal. On the other hand, disadvantages are that: skin conductivity is dependent on many external factors that may make measurements unreliable to be used as an indication of emotional arousal. For
Designing for Group Music Improvisation

equivalent, if a musician is a heavy smoker, leads a sedentary life, has excessive sweating or dry skin, physiological measurements may very well not represent his emotional state. It is evident that, while the mapping technically works and also generates reasonable additional soundscapes, further research has to be carried to validate this.

Separating the emotional arousal components of the individual musicians was instrumental in establishing a relationship between individual band members and the sound generated by EMjam: different levels of emotional arousal measured from different members of the group influenced different parameters or aspects of the music generated by EMjam. The resulting expressivity is rather large given these independent components: every musician has a different EDA response in overall intensity, onset, and duration. The generated sound accurately responds to a continuously changing mix of mapped arousal levels, and thus creates a rich addition to the music produced by the traditional instruments alone.

Richness versus control

Whereas we intend to enrich music improvisation through the inclusion of additional, emotion-based, generative musical layers, many of the participating musicians reported that they wanted control over their own expression, in both the traditional instrument and the newly added EMjam instrument. We recognize this from other studies within our overall research into group music improvisation systems. Knowing this, the relation between these two instruments (the traditional one and EMjam), but also how the novel instrument can blend into the overall resulting music, becomes even more important than anticipated. Especially since physiological processes are challenging to control, perhaps even impossible when you are fully focused on the act of playing music. Currently, the sound generated through EMjam goes from slow to fast patterns, indicating an increase in the instruments activity. These patterns, and especially changes in these patterns, provide feedback to musicians about (1) changes in arousal and (2) the control they have on the second-layer instrument. Slow patterns in the generated sound result from lower levels of measured arousal, showing less control of a musician on the instrument. On the other hand, fast patterns in the generated sound result from higher levels of arousal, widening the possibilities for the control of the instrument. Changes in arousal are noticeable over time, meaning that the provided feedback about the control of the instrument will take place in longer periods of time.

Additional considerations stemming from user feedback

A final point for discussion is that the digital design of the second-layer instrument, i.e., the mapping of sensor values to sound and the creation of musical patterns, determines specific sound qualities that make EMjam suitable for a limited set of musical styles, rather than to all styles. This stylistic tailoring is important as improvisation is largely directed by responding to what you hear, so orienting towards a specific style influences the way musicians will respond to the generated musical output. In order for systems such as EMjam to be genuinely usable it should be flexible enough to adjust its musical qualities to the overall musical style.

Connected to this is the quality of the generated sound. Currently, only linear mappings are being used with a tonal rhythmic pattern being generated. This often has little connection with the overall piece of music, which actually has to evolve around the generated sound in a style-guiding way. In the future, more sophisticated generative music techniques could be employed to enhance the experience of both the audience and the musicians. Improvisation, in the end, is also about creating unexpected musical
experiences, which needs to be supported by similarly diverse and sometimes ambiguous generative music.

**Conclusions and Future Work**

The work presented in this paper demonstrated the feasibility of controlling sound generators through the measuring and mapping of individual arousal to a group music improvisation system in Ableton Live, creating a second layer of improvisation defined by the band as a group. The advantage of the design is that it provides musicians with a starting point from which they can build up their own improvisation, helping them break the ice at the beginning of the session. It can also be used as a background music, giving musicians an indication of their emotional arousal. The use of EMjam can be though in the future in scenarios such as the practice and teaching of musical improvisation for young students, as a motivator for a more involved and participatory improvisation. Other possible scenario is the inclusion of the audience in the modification of the generated sound by considering the audience as a separate node of the EMjam system.

We should note however that due to the narrow scope of the experiment (in terms of number of participants, selection of musical instruments available, and music style), the evaluation of EMjam is indicative rather than decisive. For future research there are various aspects that need attention: The calibration of the system needs to set up to a range where musicians can play, making recalibration possible over time, possibly on-the-go using learning algorithms. Additionally, the musician’s interaction with EMjam can be improved by developing the overall interface further, adding sensors controlled by physical gestures (e.g. accelerometers) to control parameters such as volume, keynote and tempo. New features such as the control of light or visuals can be also included to reinforce musicians’ expressivity. Using skin conductivity gave an effective reaction on the digital musical instrument; nevertheless, it would be interesting to use heartbeat sensors. Changes on skin conductivity can be noticed in a long period of time; by using frequency of heartbeats as input for the system, changes on sounds would be noticed faster and external aspects such as temperature will have less influence. For our purposes, emotional arousal was mapped in a linear way to Ableton Live, meaning that changes on measurements had a proportional effect on the parameters assigned. Different ways for mapping measurements could include logarithmic and exponential functions, to emphasize different parts of the arousal spectrum. Finally, the scope of the project can be expanded towards the inclusion of the audience in the control of the instrument, by finding ways map their reactions or engagement on the performance to the digital instrument.

We wish to work towards a set of demonstrators showing the potential and value of expressive interactions in a distributed world—and then take it beyond the current boundaries of design and technology.

**References**