AALBORG UNIVERSITET

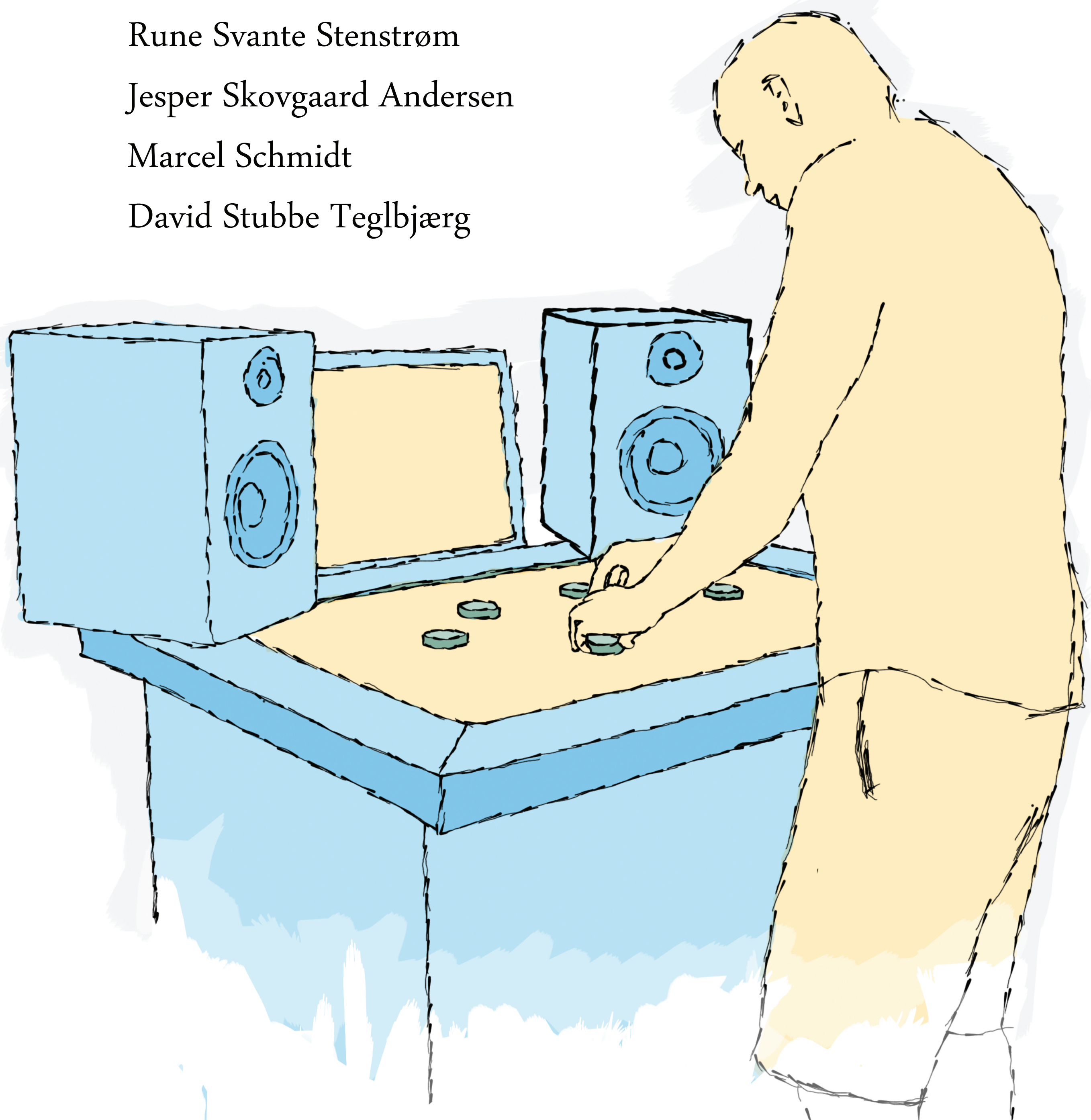
# Towards a more intuitive mixing process

Rune Svante Stenstrøm

Jesper Skovgaard Andersen

Marcel Schmidt

David Stubbe Teglbjærg



Spring 2012 / Supervised by Steven Gelineck

# Aalborg University Copenhagen

Aalborg University Copenhagen
Lautrupvang 4B,
2750 Ballerup, Denmark

Semester Coordinator:
Luis E. Bruni

Secretary:
Judi Stærk Poulsen

**Semester:**
6th Semester – Bachelor Thesis

**Title:**
Towards a more intuitive mixing process

**Project Period:**
February – May 2012

**Semester Theme:**
Integrated Design – Intelligent Systems and Computer Games

**Supervisor:**
Steven Gelineck

**Project group no.:**
607

**Members:**
Rune Svante Stenstrøm

______________________________

Jesper Skovgaard Andersen

______________________________

Marcel Schmidt

______________________________

David Stubbe Teglbjærg

______________________________

**Abstract**

Conferences such as The Conference for New Musical Interfaces for Musical Expression (NIME) continue to exhibit projects challenging the ingrained way we think about interaction with entities used for musical expression. However, some areas within music production, such as mixing and mastering, have been neglected. We believe that the post production phases of music production can benefit from the knowledge of human computer music interaction (HCMI).

This project proposes an alternate mixing console based on a tangible user interface, which is designed for a more intuitive interaction during the process of mixing music. The interface of our prototype makes use of a stage metaphor (representing physical space, such as a studio or a live stage) thereby allowing for easy control of sound sources. The stage has a listening point and panning of every channel present in the stage is determined by its position relative to the position of the Listening Point.

The prototype has been evaluated in terms of intuitive use and has been compared to a traditional mixer. Results indicate that the interaction with our prototype is more intuitive than interaction with a mixing console and suggest that a stage metaphor is better when designing for a more intuitive way of mixing music.

# Contents

# Chapter 1

# Introduction to problem area

With the introduction of the digital audio workstation (DAW) and the vast storage possibilities of the hard disk, the computer has become center for many musical processes such as compositing, synthesis, production, mixing and mastering. As the computer moved beyond the traditional desktop usage and into areas such as music production, so did the traditional WIMP-interaction (Window, Icon, Menu, and Pointing device) along with its limitations. The limited degrees of freedom in WIMP interaction have inspired many new researches within the area of Human-Computer Interaction (HCI) to come up with new ways of interaction [17].

Human-Computer Music Interaction (HCMI) is a sub area of HCI focusing specifically on the interaction between human and computer in a musical situation. Conferences such as The Conference for New Musical Interfaces for Musical Expression (NIME)[2] continue to exhibit projects challenging the ingrained way we think about interaction with entities used for musical expression. The majority of NIME research products therefore focuses on breaking or changing the barrier between interaction and musical expression. Most projects seem to only focus on new musical instruments, interfaces for composing music or interfaces for live performance.

Areas which seem to not have reached the full spotlight of HCMI are the post production phases in a musical production process, such as mixing and mastering. This might be so, because mixing and mastering are not commonly seen as artistic or creative as processes like composing or performing music. Mixing and mastering tend to be understood as the more technical processes of the overall musical production process. The literature of mixing proves this to be a misconception, by stating that the process an sound technician has to go through when mixing is just as subjective and creative as the process a musician goes through when composing and performing music [28] [25] [31].

We are interested in why there has not been done much research focusing on the post-production phases. We hypothesize how the post-production phases could benefit from the knowledge of HCMI. In terms of delimiting our focus area, we have chosen a specific area of interest, which is the mixing console used in the post production phase for mixing music.

Historically, the first 32-channel multi-track mixers emerged in the late 1960's. Figure 1.1 displays two Yamaha mixers. The Yamaha MG24/14FX 24-Channel mixer on the left was developed recently and is currently in stores and the Yamaha PM-1000 mixer on the right is from 1974. As one can see the interface design has not really undergone any major development.

Figure 1.1: Yamaha MG24/14FX 24-Channel (left) and Yamaha PM-1000 (right)

Considering the breadth of capabilities of a mixing console it is a straightforward device in terms of mapping. This simplicity stem from technical limitations at the time it was developed and the layout of the controls ended up mirroring an audio signals path through the board [20]. With few exceptions, every pot, slider, or other control serves only a single purpose, regardless of context. This console paradigm is in fact so deep rooted in modern music production that even when it became time to move the process of music production onto the computer platform, the mapping of a channel strip were once again chosen. Figure 1.2 shows the mixing windows of Logic Pro (on the left) and Protools (on the right) which are two of the most used DAWs in professional music production.

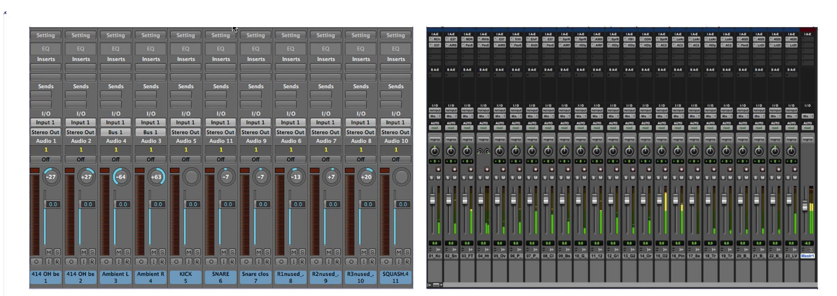


Figure 1.2: The mixer windows of the digital audio workstations Logic Pro (left image) and ProTools (right image)

In these software one actually uses the up and down motion of the mouse to control a virtual knob. Mappings like this caused us to hypothesize, that it might be time to benefit from the technological advances made since the 1970's together with using what has been learnt within the field of HCMI.

## 1.1 Preliminary Interview

To validate our initial hypothesis, we contacted professional sound technician, producer and professor at the Rhythmic Music Conservatory in Copenhagen, Morten Büchert for an interview concerning the current development of the mixer and the current paradigms within the area of mixing in general. The focus of this interview was on the needs and challenges, when it comes to the current and future ways of mixing music. A short overview of the interview guide is shown below:

**Briefing**

- Who we are and introduction to our project.
- Who is he?

**Interview**

- The purpose of mixing
  - What are the objectives and how are they fulfilled?
- What are his needs?
  - Are there any problems with the way the objectives are achieved?
- Target group
  - Who has these problems?

**Debriefing**

- Our work flow - Iterative process
- How we would like to structure further cooperation with him and his students

The interview served the purpose of an initial exploratory and motivational interview. The interview was conducted as an unstructured interview with traits from the semi-structured form of interviewing, as it was, research-vise, conducted on a somewhat clean slate. The interview focused on different aspects within the area of mixing music, which is a complex and multifaceted process. An exploratory interview therefore seemed appropriate [19]. The interview took mostly the form of a conversation, since this often can help with going in-depth with a particular topic [19]. The recorded, transcribed and codified interview can be found on the DVD under the folder *"Initial Interview"* and the complete methodology and interview guide can be found in appendix A.

Morten stresses that the interaction paradigm within music mixing has not changed the last 60 years:

*"this fader representation of volume, pan and eq. is really.. what 60 years old?"*

Furthermore he states that *"...we are really conservatory now in our thinking of approaching mixing: Volume, pan or effects, something like that"*. When trying to rethink a current paradigm you might experience that people who are used to this paradigm will have a hard time thinking out of the box. It is interesting whether this is one of the facts contributing to the preservation of the traditional mixing console.

In relation to the current state of the mixing console Morten expresses some concerns about the direction of development concerning mixing consoles:

*"So now we all have 265 audio channels to be able to record and process, but will it fit between to speakers. I dont know, I dont think so. It might be sonically better, but is it musically better? Does it feel better to listen too? That is one of the areas that Im most concerned about."*

It seems as the new developments in mixing are focusing on building mixers with more tracks and more technical controllability. A possible drawback of this direction is the creation of barriers between the sound technicians' idea for the mix and the process realizing the idea, thereby diminishing the influence of intuition. On this Morten comments:

*"Intuition is the most important thing in music. - a lot of computers still have a screen that you have to adhere to. You look at an eq. setting and its really "wow that looks crazy", but how does it sound? We are so focused on the visual side of things that its hard to discern audio from video or what you see from what you hear."*

The digital mixers and DAW mixers allow you to tweak parameters down to small floating point precision. This full control and insight can cause the sound technician to subconsciously rely on his visual senses when mixing instead of his auditory sense. The majority of the effects and devices, used when mixing, utilize the underlying mathematical process as representation. Morten expresses concern about this fact in relation to intuition when mixing:

*"Wouldnt it be better to be in an intuition style mixing instead of a mathematically correct mixing style? I think the music would be better, I hope so."*

Morten finds that many of these mathematical analogies compete with his auditory mental image. Most new consoles are digital and it seems that the manufacturers are trying to work on making mixers smaller while keeping the same functionality:

*"So the way they layout the console [red. digital console], they do a master channel section and then you have individual faders, but not the whole channel strip up at the same time. So you have to access the bass drum and turn a few knobs and then you have to access the vocals and turn on the same knobs. So there goes a significant amount of time before you know where you are at. You cannot look at the artist and work with two hands. You have to*

*access, turn, access, turn. So the response time is a lot slower"*

In above statement Morten is referring to the mixers used for live mixing mostly, but it seems that accessibility is a key factor within the area of mixing in general. There seems to be a paradox between the mixers getting more channels while decreasing in size. This paradox can have a negative impact on the accessibility and intuition when mixing.

Although, Morten is only one out of many sound technicians, he seems to initially validate our motivational hypothesis. Investigating other means of representing mixing parameters and new ways of interacting in the overall mixing process, can hopefully lead to different directions in the development of mixing equipment. Such work can hopefully break down the barriers in the current mixing paradigm and maybe allow sound technicians to rely more upon their intuition.

Based on this our main research question will be:

"*In what ways can the process of mixing benefit from a rethinking of the interaction paradigms used today?*"

# Chapter 2

# Gathering knowledge

The following chapter serves the purpose of gathering relevant knowledge about the research area. Before venturing further, one should have a basic understanding of the concepts of mixing. In section 2.1: *entering the realm of mixing* the reader will be given the most basic knowledge in order to follow the terminology and mixing concepts utilized in the rest of the report. Section 2.2: *intuitive use* aims at clarifying and elaborating on the meaning of intuition and explores how to design an interface that is intuitive to use. Section 2.3: *tabletop tangible user interfaces* will investigate the world of tangible interfaces and their advantages. Finally section 2.4: *future technology workshop* describes a workshop conducted at the Rhythmic Music Conservatory in Copenhagen with the aim of investigating sound technicians' own ideas about the future of mixing.

## 2.1 Entering the realm of mixing

This section will give an explanation of the process of mixing and a short overview of the mixing console and its functionality. By dissecting the process of mixing, it should also provide a better understanding of the different aspects involved in a mixing phase. This helps establishing requirements for an alternative system. The differences between the situations of live mixing and studio mixing will then be described, both in terms of a general level and in terms of interaction.

### 2.1.1 What is mixing?

Overall the process of mixing is just as subjective as the other phases of music production and there are as many definitions and takes on mixing as there is literature on the subject[28] [25] [31]. A definition of the mixing process and what it constitutes, is proposed by Izhaki. He defines mixing as a process in which multitrack material  whether recorded, sampled or synthesized  is balanced, treated and combined into a multichannel format, most commonly two-channel stereo [28]. Furthermore, Owsinski mentions that mixing, on a more abstract level, is about getting the energy of the song across to the listener. It is almost like a musician who picks up a guitar and tries to play: he may have the chart in front of him, but soon he has to go beyond the notes in order to get creative [25]. From this it becomes clear that mixing constitute a post processing phase of material in multichannel format, and on a less technical basis Owsinski seem to point in a direction that treats mixing as a sonic presentation that allow creativity and performance - as in composing music.

#### 2.1.1.1 The 6 elements of mixing

How do technicians then approach mixing? Most technicians rely upon their intuition, together with a mental picture of the final product before they begin to mix. According to [25] the foundation for mixing consists of figuring out the direction of the song. What is meant by this is finding the most important element and emphasize(s) it/them - everything should then be built on top of this/these element(s). There exists a coherent point of view that every piece of modern music has six elements to a great mix [25]: balance, panorama, frequency range, dimension, dynamics and interest. It is important to notice that the methods used to achieve the following goals are very subjective and can vary a lot. Despite, we will try to describe the elements in more general terms in order for the reader to get a better understanding of the areas on which the sound technician focus his/her attention when mixing.

**Balance**

By balancing the sound technician makes sure that all the audio tracks are properly presented i.e. making sure that instruments with same frequency and loudness do not fight for attention. If two instruments share the same frequency masking can occur. Masking is when two instruments share a frequency making them hard to distinguish from each other or sounding muddy. An example of this is when the kick and the bass share frequencies causing the bass line to disappear when a kick is present. In order to avoid this, the sound technician makes sure that every instrument more or less sits in its own frequency range.

### Panorama

Panorama is achieved by panning. Panning is about placement of sound elements in the sound field and allows the technician to separate the different sonic elements so that every element can be heard much more distinctly. If done correctly it can make the sound bigger, wider and deeper [25]. When mixing in stereo there are typically three areas that get the most attention - the center and the extreme right/left.

### Frequency range

The material used in a given song might have been recorded in different studios using different monitors, different producers or even different musicians. In order to smooth out the difference in sound elements, the technician can manipulate the frequency range. The tool used for this is the equalizer and the primary goals are to make an instrument sound clearer and more defined, together with making all the elements of the song fit together. When equalizing, the audio band is typically broken down into six distinct ranges - ranging from Sub-Bass - Bass - Low Mids - High Mids - Presence - Brilliance [25]. This process also contributes to the balance of the mix.

### Dimension

By adding effects to a mix the sound technician can achieve a somewhat bigger aural space and add excitement to the mix. By using effects he can also move a track or instrument back in the mix and thereby create a sense of dimension. Typical ways of achieving a sense of dimension is by using delay and reverb or any other modulated delay e.g. chorus, flanger, etc.

### Dynamics

The fifth main element is about creating dynamics to the soundscape and is controlled by the use of compression, limiting and gating. Controlling the dynamics is all about keeping the sound level even - lifting the level of soft passages and lowering the level of loud passages. Compression can, on the other hand, also be used as an effect in itself. It can radically change the sound of a track [25] and when set correctly it can make the track seem closer and thereby cause more aggression and excitement. If there is one major difference between the sound of a demo or semi-pro recording and a finished professional mix, its the use of compression [25].

### Interest

Finally the interest is concentrated around determining the direction of the song. According to Owsinski this element is the key to a great mix. This is the foundation of mixing - it is not about being technically correct - but more about passion and emotion. This is seen as the first step of mixing (or maybe even as something that comes before the mixing phase). As mentioned earlier the role of this element is to decide on the direction of the song i.e. figuring out what is important and what should be pursued. The function of interest can be seen as a bridge between the phase of composition and mixing [25]. Making sense of the mixing process is done by working out where to start - this should then be the most important section of your mix [31]. The artist behind the song will almost always have a specific interest and this should be pursued (or taken into account) throughout the process of mixing.

The above mentioned elements are the basic building blocks of a mix. In addition to these Izhaki mentions four distinct objectives of mixing [28]. The mood and emotional context should be reflected in the mix. The mix should be balanced; typically through the domains of frequency balance, stereo image balance and level balance. Then there is the aspect of defining how distinct and recognizable sounds should be and finally the interest of the mix. Clear thoughts on the interest of the mix should ensure a dynamic product. The vision can sometimes end up changing because of input from the surroundings i.e. producer and artist. Therefore some technicians will prefer to complete the job unattended by the producer/artist, but others will prefer the input from producer/artist.

### 2.1.2 The mixing console

The core tool used by all sound technicians when mixing is a multitrack mixing console. Whether this is physical or a software simulated entity it seem to be the chosen standard no matter the environment. This section will introduce the basic functionality of a mixing console.

The mixing console, or mixing desk, or simply mixer is the tool used for studio- and live mixing. There is a big variety of different mixers available all sharing the same purpose: combining signals from multiple sound sources, processes them to get a nice balance and quality, and output the resulting mix to a recorder, a sound system etc.

Mixers can be categorized into analog and digital mixers. The difference lies in the fact that analog mixers do not convert the audio signal to digital data, whereas digital mixers accept both analog and digital inputs and all input, analog or digital, is converted to digital data. The digital mixer also affects the digital data before it is converted back to an analog signal at the mixer output. Both types of mixers have advantages and drawbacks. Analog mixers tend to have all functions accessible at all time, whereas digital mixers are smaller and programmable.

Disadvantages for pure analog mixers are that the settings cannot be saved between sessions. A disadvantage of a digital mixer is the propagation delay, as well as the "layering" of functions per knob, which can make it confusing to operate. A nice trade-off is a "digitally controlled analog" mixer, where the control surface takes advantages of digital technology and the audio processing is kept analog. Whether digital or analog, mixers are usually classified based on their input and output (I/O) capabilities. For instance, a 32x8 mixer has 32 input channels and 8 output channels. One or more of the inputs are mixed to one output, which is known as a sub-mix. These outputs can send a mixed signal to a multi-track recording device. Almost all mixing consoles also provide a stereo output. Figure 2.1 shows the Mackie Onyx 2480 mixer and its sections.

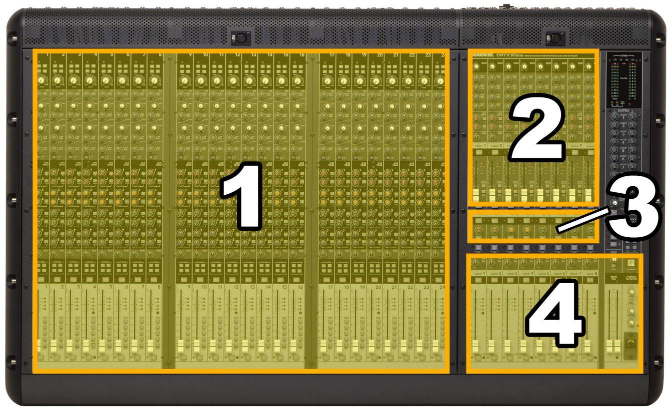



Figure 2.1: The Mackie Onyx 2480 mixer with its input channel strips (1), aux send master section (2), the monitor section (3) and output section (4)

Now that we have got an overview of the mixing console we will have a deeper look at the following sections: input channel strips, the aux send master section, the monitor section and the output section.

**(1)Input channel strips**

The vertical input channel strips are identical for each input. Even so the functions differ a bit from manufacturer to manufacturer. There are several basic elements which are the same on each mixing console. Each channel gets its input from a device patched to the back or top of the mixer. This section is called line level input and can be connected to synthesizers, microphones, computer audio interfaces etc. There are usually three options of connecting these devices to the mixer, XLR-input for microphones, 1/4"-input and a tape-input. Firstly you have the possibility to adjust and/or amplify the incoming signal. This is referred to as the gain knob and is usually located on the top of the channel strip. Gain can be used to adjust all incoming signals, so they have the same value in the channel fader section. This is also where the preamp is located. The preamp is an important part, because lost quality in the signal cannot be regained. Besides the final audio output of the mixer, each channel has a separate set of outputs connected to effect devices/processors (reverb, delay, distortion, etc.) i.e. effect sends.

The second option, offers the possibility to equalize (EQ) the signal after it has gone through the effect device separate from the original signal. The EQ is used to boost (increase) or cut (decrease) different frequencies for a few dB. These frequencies are usually categorized in high-, mid- and low-frequencies, as explained in the previous section. Often "EQ-In" and/or "low-cut" buttons can be found. The EQ-In just turns on the EQ or bypass it and the "low-cut" button cuts off frequencies lower than around 75Hz. It eliminates any noise and room rumble and is therefore mainly used to EQ vocals.

The last part of the input channel strip consists of the pan-pot, the solo and mute button and the channel fader. The pan pot, or "panoramic-potentiometer" is used to place the stereo sound in the stereo-image. The mute-button is used to silence the channel, whereas the solo-button sends the channels signal to a separate solo-bus. Soloing one channel usually mutes all the other channels.

#### (2)Aux send master section

As described earlier, every input channel strip has an "aux send" knob which controls the relative amount of signal being sent through the effect device. The "aux send" at the aux send master section controls the total amount of signal sent to the buses. Often the "aux send" can be soloed by a solo button in this section.

#### (3)Monitor section

The monitor section is used to control what signal and with which level it is send to the studio speakers, control room speakers and headphones. This section just changes the level of the speakers and does not change the level of the final output. This is what the output section is for.

#### (4)Output section

The output section contains a fader for each of the output group (buses) and one fader for the final stereo mix. This fader controls the level for each output group and a solo button makes it possible to solo each group.

What is worth noting is that a mixer includes three different types of controls: faders, knobs and buttons. Faders are typically used for controlling signal gain or amplitude. Knobs are typically used for controlling panning or amount of different frequency ranges. Buttons are used as on and off switches and are for example used to mute and sole tracks.

### 2.1.3 Studio and live mixing

Overall mixing can be divided in two; the situation of live mixing and the situation of a studio. In order to pin point the difference we conducted two contextual interviews focusing on how the interaction differs depending on the situation. The methodology will firstly be explained followed by a presentation and discussion of the findings.

#### Methodology

A contextual interview form is chosen since the focus is on interaction. This is done because the interviewee might not be aware of how he/she interacts when performing a task. Doing a strictly verbal interview could therefore generate limited or biased data. In a contextual interview the interviewee is asked to *demonstrate* how he/she solves a problem, instead of *explaining* how to do it [19]. Two observational sessions were done. One at the Rhythmic

Music Conservatory in Copenhagen and the other at the concert venue Vega, also in Copenhagen.

The interviewee is Morten Büchert who also participated in the preliminary interview. Each session is further divided in two subsection: Analogue approach and digital approach to mixing. The observations were recorded on video and the can be found on the attached DVD. During the observations one researcher controlled the camera, one took notes, while the remaining two asked questions when something needed to be elaborated. The video observations have afterwards been used for interaction analysis.

### Studio session

As explained, the studio session was split into two parts one with an analogue mixing approach and one with a digital mixing approach. The same mixing console was used during each session. Figure 2.2 shows Morten in front of the mixing console.

Figure 2.2: Morten in front of a mixer at the Conservatory in Copehagen, showing analogue and digital studio mixing

#### Digital approach to studio mixing

In the digital approach the mixing console is used to control Pro Tools [5] software and the music score is a typical rock number.

Before starting on the mix, Morten explains that he usually decides beforehand on which direction he will take the mix. This relates to the interest process in section 2.1.1. It also shows that he has a clear idea of how he thinks the mix should sound before he starts mixing.

After deciding on the direction he starts mixing. In the first step Morten concentrates on listening to the song and adjusting the levels and pan, by setting the volume-faders and pan-knobs. Here he uses the channel strip section of the mixer (see 2.1.2. The first step can be seen as the balancing and panorama process explained in section 2.1.1. After finishing this process he continues with mixing in Pro Tools. Here he starts by adding and adjusting equalizers to all the tracks. This part of the session relates to the frequency range process also explained in section 2.1.1. By doing this each instrument becomes more defined in the mix. They also get a fuller sound. He explains that, if he had continued he would have continued by tweaking dynamics and adding effects, such as delay and reverb. While explaining this he briefly demonstrates how he would add compression and reverb. These last steps relate to the dimension and dynamics processes explained in section 2.1.1. Morten's process of mixing corresponds well to descriptions found in mixing literature.

From the observation it is clear that it is only during the balancing and panorama processes that he actually uses the mixing console. In the frequency range, dynamics and dimension processes he only use WIMP interactions to control effect plug-ins on the DAW. For instance to control the equalizers, a curve is represented on the screen and with the mouse he tweaks this curve until reaching a satisfactory result. This curve representation relies a lot on the underlying mathematical representation. This confirms what Morten explained in the preliminary interview (see section 1.1), about the tools relying a lot on mathematical representations.

Before ending the digital approach, he quickly show how he would automate the mix. This process is not explained in section 2.1.1. In this process sound technicians can record changes to the different parameters through the mix. This can for example be changing the amount of reverb on a vocal during the track. Here Morten switch back to the console, but states that it is also possible to automate by using the mouse. Morten concludes that in the digital approach there is actually no need for the console, it can all be done by mouse and keyboard.

*"I wouldn't need a console, I could do everything by mouse...you saw me editing all the plug-ins in that way"*

This observation indicates that there are some general interaction problems in the digital approach. The fact that he uses WIMP interactions to control a mathematical representation, not related to how we perceive the sound, is definitely contributing to the creation of interaction barriers when mixing.

**Analogue approach to studio mixing**

In the *analogue* approach the sequence of the steps is the same. The first step is adjusting level and panorama to get a rough-mix. After this the frequency range is adjusted, but here Morten uses the equalizer on the mixing console instead. Controls for this are found on the channel strip and typically include four knobs controlling four different frequency bands (see 2.1.2). Asking Morten whether he imagine the equalizing curves used in the digital approach he answers *"no, just by ears"*. This indicates that a graphical representation is not necessary and maybe actually mislead the sound technicians from using their auditory sense.

In the dimension and dynamics process he uses hardware effect machines instead of software plug-ins. This process consists of connecting the effect rack via patch cables to the mixer. In this process Morten continuously switches back and forth between effect rack and mixing console. Compared to the digital approach, where it is just a matter of adding effects in the DAW by mouse and keyboard, one can see a big difference.

After completing both sessions, Morten concludes that there is a fundamental difference in mixing digital and analogue:

*"There is a fundamental difference in the way you attack the mix. I have never done this before, like comparing A and B. But I haven't looked at the screen at all [red. in the analogue approach]... I just oriented with my ears"*

This statement is also confirmed by our observation. During the analogue mix Morten is not using a computer monitor. In the digital approach he is influenced a lot by discrete numbers on the monitor. While mixing without a discrete representation of values he uses his intuition and auditory sense to a larger extend.

When asking Morten about the importance of intuition and how it differs in the two approaches he replies:

*"I'm a big fan of intuition... all important decisions were made in a really short time, and I would probably take the same decisions tomorrow... it is not so much about making this decision, it's more about intuitively reaching out for this knob and you just compress it...it's a gut feeling"*

In above statement he is referring mostly to the analogue approach since he did not use a knob for compression in the digital approach.
To sum up, there is a big difference between the two approaches. In the analogue session he relies mostly on the auditory sense whereas in the digital approach he also relies a lot on the graphical representations of the mix.
This observation indicates that it is a good idea to avoid discrete representations, because they tend to disrupt and disturb ones intuitive feel. Furthermore, an alternative approach should not make as much use of graphical representations as a DAW does.

## Live session

The live session was conducted at Vega. This time two different consoles were shown and the session can be split into two sections, similar to the studio mixing session. One session focusing on the analogue approach to live mixing and one focusing on a digital approach to live mixing. The digital approach was shown during a sound rehearsal with the band M83 [3]. Since only a digital mixer was in use during the live rehearsal, Morten explained the functionality of the analogue mixer without any audio output.

### Analogue approach

The first session focused on the analogue mixer. We have referred to mixing using a analogue mixer and mixing using an analogue approach. The difference between mixing on an analogue mixer and using an analogue approach on a digital mixer only has to do with the underlying processing of the audio (see 2.1.2. Despite, in a live situations the digital mixer allows for the use of effect plugins on a DAW. This means that one uses WIMP interaction more in live situations.

### Live session - digital mixer

The session with the digital mixer was, as explained, conducted during a sound rehearsal with the band M83. Morten served as sound technician for the warm up band, but we were also able to observe M83's sound technician. The digital mixer used is shown in figure 2.3:


Figure 2.3: Morten in front of the digital mixer at Vega during a sound rehearsal.

Live mixing more or less contains the same processes as in the studio mixing. There is though several differences. One of the most important differences is probably the fact, that you in a live situation have a band and a real sonic room which you relate to. This means that the interest phase is not as prominent as in the studio mix. This also influence the different processes of mixing explained in section 2.1.1. In terms of balancing and panorama there are much more limited possibilities. Firstly panning an instrument to the far right, for example, can cause the audience to the left not being able hear it. In terms of volume, then there is firstly laws about how loud you can play, secondly the sound has to be loud enough for the audience to be able to hear it. Since the audience contribute to the overall loudness in the room, the range where one can place audio tracks is - volume-wise - much lower than in studio mixing. Because of this a decibel measuring device is placed close to the mixer. This device records the overall loudness in the room and both Morten and the other sound technician paid a lot of attention to it. In relation to this Morten comments:

*"In the studio you can create whatever you want, here [live situation red.] you have to work with whatever you got"*

Live mixing furthermore relies a lot on quick on the spot changes to the mix. For example, if the musical score is about to go to a guitar solo, the sound technician has to be able to quickly adjust the different parameters of the mix so it suites the solo. This means that the automation phase explained in the studio session section is a lot different when live mixing. This also means, that accessibility of the interface is very important. The digital mixer has a lot fewer controls compared to a analogue. This is especially a problem when applying effects, since one knob control the same parameter for all the tracks. One has to press a button on the track to activate the knob for that specific track. Morten finds this very problematic:

*"One thing (a knob red.) doing several things is really dangerous, as you always have to think"*

The tendency to make digital mixers more compact seems to be problematic and as a result one has to assign several parameters to one knob. Another important difference between studio and live mixing is that there is much less time for doing a mix in a live situation. In a studio you theoretically have all the time you need, but in a live session you have a much shorter period of time. Because of this portability and stability of the equipment is of great importance. Morten also refers to this:

*"stability is key"*
*"stability over functionality, all the time"*

The observations at the Rhythmic Music Conservatory and at Vega illustrate clear differences in mixing. No matter environment the console should support intuitive use. There seem to be many related problems to this. In the studio environment intuition benefits the quality of the mix. It also helps the sound technicians in the process of transforming their ideas created in the interest phase to a sonic output. Another important aspect is how to lay out the controls. The live session proved that layering in controls is a very bad idea and this must be taken into account when designing for an alternative way of mixing music.

### 2.1.4 Initial requirements

If one is to generate other means of interaction it is important to pinpoint which functionalities this interaction should be able to control. The 6 elements of mixing are the most basic functionalities that an alternative tool for mixing should encompass. This means that one should be able to control the balance and panorama of the mix by controlling parameters such as panning and volume. The 6 elements also require the implementation of an equalizer for controlling frequency as well as basic effects such as reverb and compression. Moreover accessibility is an important factor that needs to be taking into account. This means that an alternative tool should not utilize layering of controls, but instead try to implement each control in an easy accessible manner. The four sections of the traditional mixer can also be used as inspiration for the functionality of a new mixing device. Portability and stability are also important factors, but since the outcome of this project will be some sort of prototype, the later three are not going to be taken into account.

How do one then create an alternative mixer that encompasses the above stated functionalities, but at the same time allow for more intuitive use than a traditional mixer? The rest of this chapter will try to investigate this and the focus can be summed up by the following research question:

*"How can we design an alternative mixer that allows for more intuitive use than a traditional mixer?"*

## 2.2 Intuitive use

### 2.2.1 Defining Intuitive Use

There are two big research centers when it comes to "intuitive use", one located in Australia leaded by Alethea Blackler and one in Germany leaded by Jörn Hurtienne. In the paper "Towards a unified view of intuitive interaction" they explain that even though different methodological approaches have been used their results are complementary[11]. This can also be seen in how related their definitions are. The following describes Blackler and colleagues definition of intuitive use:

*Intuitive use of products involves utilising knowledge gained through other experience(s). Therefore, products that people use intuitively are those with features they have encountered before. Intuitive interaction is fast and generally non-conscious, so people may be unable to explain how they made decisions during intuitive interaction.*

Hurtienne is a part of the Intuitive Use of User Interfaces (IUUI) research group, which was founded in 2005 with the aim of creating a tenable definition of the term "intuitive use" and providing tools and guidelines for designing interactive products that are intuitive to use. Group members have backgrounds in psychology, computer science, engineering, linguistics and industrial design. Their definition of intuitive use is:

*A technical system is intuitively usable if the users' unconscious application of prior knowledge leads to effective interaction.*

Both definitions focus on prior knowledge, which leads to unconscious interaction. Hurtienne further states that the interaction has to be effective, which means that the outcome should be the same as the users intended goal. If a system is intuitive it leads to effective interaction which is a classic usability measure. Even though the definition focuses on effectiveness Hurtienne and Blessing also suggest to test on efficiency and satisfaction, when trying to validate an approach for designing an interface that allows for intuitive use [16]. In general intuitive use should reduce the cognitive processing time and therefore IUUI emphasize "effortlessness" as a subjective criterion [11].

The suggestion of minimizing the cognitive load correlates well with the goal of intuitive interaction, but how do we create "intuitive use"? In order to get this question answered it is important to understand "intuitive use", by analysing what is meant by *prior knowledge* and *unconscious application*, as they both occur in the definition of "intuitive use".

According to Hurtienne **prior knowledge** may stem from four different sources which are:

- *innate* knowledge which is the first an lowest level and is acquired through the activation of genes or during the prenatal stage of development. Generally this is what reflexes or instinctive behaviour draw upon.
- *sensorimotor* is the second level and it consists of general knowledge, which is acquired very early in childhood and is from then on used continuously through interaction with the world. An example of this could be face recognition, gravity or concepts of speed and animation.
- knowledge specific to the *culture* an individual lives in. This means, what is known within the western group of cultures is not necessarily equivalent to the knowledge of people in eastern cultures.
- *expertise* is the most specific knowledge and is specialist knowledge acquired in one's profession and through hobbies which in the case of this research could be sound technology.

Across the *sensorimotor, culture* and *expertise* levels of knowledge Hurtienne also distinguish knowledge about tools. Tool knowledge seems to be an important reference when designing user interfaces. These tools differ from each knowledge group. Where a tool on the *sensorimotor* level could be stones used as weight, a tool at the *expertise* level could be complex software as a DAW. Hurtienne has coined this theory *continuum of knowledge* and an overview can be seen in figure 2.4. The further one rises towards the top level of the continuum, the higher the degree of specialization of knowledge is required, while the potential number of users - possessing the knowledge needed - will be decreasing.

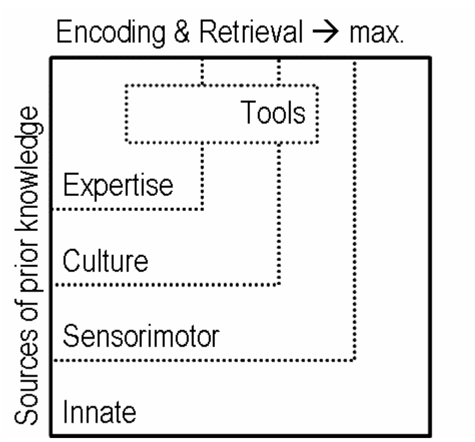



Figure 2.4: 2D illustration of the continuum of knowledge where the vertical axis represents the source of knowledge and horizontal axis represents the frequency of encoding and retrieval

How is the continuum of knowledge then *unconscious applied*? The more frequent encoding and retrieval of data has been in the past, the more likely it is that memorized knowledge is applied without awareness by the user [11]. Knowledge at the expertise level is acquired relatively late in life and is (during a life span) not as frequently used as knowledge from the culture or sensorimotor level. As learning theory suggests, knowledge from the lower levels of the continuum is therefore more likely to be applied unconsciously than knowledge from the upper levels [16]. Having this in mind, we can reason that it is more likely that interaction with an interface is intuitive, when knowledge from the lower level of the continuum is involved.

Hurtienne and Blessing[16] states that there are further advantages by limiting "intuitive interaction" to the lower levels of the knowledge continuum e.g. extremely frequent encoding and retrieval events lead to a higher robustness of information processing. In situations of high mental workload and stress a fall-back on lower stages of the knowledge continuum will occur. In terms of mixing this suggests that a mixer for live purpose should be more intuitive as the user might be stressed due to time constrains. An advantage of designing for low levels of knowledge is that subconscious processing of user interface elements in general means less work-load on the cognitive processing capacity. Thus more cognitive resources will be available for solving the working task at hand, instead of figuring out how a piece of technology works wasting time and mental effort. This is especially important in relation to the mixing process, as the creative part of the mixing can be disrupted when the user has to use mental efforts for technical issues

Aiming for an interface that requires lower levels of knowledge might be a good idea. As the innate knowledge is found to be too basic, we aim for the sensorimotor level. According to these theories, designing an interface where the interaction is based on physical laws and therefore categorized as sensorimotor knowledge should then be intuitive to use.

### 2.2.2 Designing for Inuitive Use

## Principle for designing for "intuitive use"

The IUUI group has not only focused on defining "intuitive use". They have also formulated as list of principles which can be used as a guideline when designing such an interface. The principles are explained in [11] and the following paragraphs explain the different principles.

#### Suitability for the task

This principle focuses on the fact that the user interface should only present the user with information related the successful completion of the task. Furthermore input and output should be appropriate for the task and only the necessary interaction steps for performing the task should be included.

#### Compatibility

Blackler divides compatibility into three levels: the user interface level, the level of the technical system, and the user-task level.

At the user interface level classical stimulus-response compatibility relates to the relationships of controls and the object they are controlling. This is an really important principle to consider as a greater degree of compatibility, is likely to result in faster learning and response times, fewer errors and a lower mental workload.

The technical system level focuses on two sides; the location and movement of display and controls and on the location and movement of parts of the technical system.

**Consistency**

Consistency refers to internal consistency and external consistency. Internal consistency refers to consistency within the system whereas external consistency refers to the consistency between the system and things outside the system. This for example refers to user knowledge, work domain and metaphors. It can also refer to other systems.

**Gestalt laws**

Gestalt laws are based on the basic principles of perception. An example is the law of similarity. This suggests that one could perceive similar objects to be the same and thereby assume that they have the same functionality. The IUUI group are far from the only ones suggesting that gestalt laws can be used for design. In "what is beautiful is usable"[33] Tractinsky measures a strong correlation between the perceived aesthetics and perceived usability of a system.

**Feedback**

The system must provide appropriate feedback after any user operation. This feedback must be immediate and self-evident so a user does not experience any uncertainties about the results of performed actions.

**Self descriptiveness**

This refer to the fact that it at any time should be obvious which step the user is in at any given time. It furthermore states that the interface should be designed so the user at any time knows which interactions he can perform or which actions are available.

**Affordances**

Every object should be based on affordance, so it immediately becomes clear how to interact with it. Physical objects are natural affordances, but virtual objects, like buttons, needs to invite the user to interact in a specific way i.e. pushing, scaling, moving etc.

## Image schemas

Another relevant theory suggested by IUUI is a concept called image schemas. In [16] Hurtienne and Blessing defines image schemas as *"abstract representations of recurring dynamic* *patterns of bodily interactions that structure the way we understand the world and thus are important building blocks for thinking"*. This means that we have certain structural conceptual ideas on how to describe concepts such as space and containment. These image schemas can be classified into the sensorimotor level of the knowledge continuum. As we are aiming to base the prototype in this level, image schemas can prove to be relevant for the design phase. Hurtienne and Blessing divides the image schemas into seven different groups:

| Group | Image Schemas |
|---|---|
| BASIC SCHEMAS | SUBSTANCE, OBJECT |
| SPACE | UP-DOWN, LEFT-RIGHT, NEAR-FAR, FRONT-BACK, CENTER-PERIPHERY , CONTACT, PATH, SCALE |
| CONTAINMENT | CONTAINER, IN-OUT, CONTENT, FULL-EMPTY, SURFACE |
| MULTIPLICITY | MERGING, COLLECTION, SPLITTING, PART-WHOLE, COUNT-MASS, LINK, MATCHING |
| PROCESS | ITERATION, CYCLE |
| FORCE | DIVERSION, COUNTERFORCE, RESTRAINT REMOVAL, RESISTANCE, ATTRACTION, COMPULSION, BLOCKAGE, BALANCE, MOMENTUM, ENABLEMENT |
| ATTRIBUTE | HEAVY-LIGHT, DARK-BRIGHT, BIG-SMALL, WARM-COLD, STRONG-WEAK, STRAIGHT, SMOOTH - ROUGH |

Figure 2.5: Image schemas organised into seven groups[16]

Apparently the most relevant groups are the two groups focusing on image schemas relating to how we describe and understand space and containment. This is due to the fact that elements in a mix are typically described using these metaphors. For example, a guitar can be placed to the left or to the right, it can be in front or in the back of the mix, its volume can be turned up and down and so forth. Furthermore the containment image-schematic metaphors can prove important as sound technicians seem to refer to the mix as a spatial container.

In their example of the UP-DOWN schema, Hurtienne and Blessing explain that *linguistic analysis points to the metaphorical extension of the UP-DOWN schema to conceptualize abstract domains like* [16] and they then list several domains. The most relevant is the quantity domain, where they state that UP normally maps to MORE and DOWN normally maps to LESS. This image schema is utilized by the traditional mixing console as well as a whole range of everyday devices.

Although image schemas describe human sensorimotor experiences with the physical world, their actual strength lies in their metaphorical extension for structuring abstract concepts[16]. Linguistic analysis have shown that image schemas can serve as source domains of countless metaphors. As language reflects thought, image schemas and their metaphorical extensions should also be working in non-linguistic reasoning. In fact there is growing evidence of this coming from the field of cognitive psychology [16]. Based on this idea, Hurtienne and Blessing have conducted a test where they try to evaluate whether applying these extended conceptual abstractions - based on the image-schemas - can help designers build better products that are more intuitive to use. One of the tests focus on the UP IS MORE, DOWN IS LESS conceptual extension of the image schema. In one part of test they have a slider implemented so it follows the conceptual extension and in another they have implemented it reversed. The participants were primed randomly with words like "more" or "less", and asked to adjust the sliders accordingly. Hurtienne and Blessing then measured the response time of the participants, with the intention of investigating whether the participants use less cognitive activity in either test. Although their results were not significant, they argue that there is still a correlation between the response time and the use of image schemas. They furthermore conducted the test with horizontal sliders and found that people tend to understand RIGHT AS MORE and LEFT AS LESS.

Based on this it is clear that the mixing console actually makes use of well known image-schemas, such as UP-IS-MORE, DOWN IS LESS and RIGHT-IS-MORE, LEFT-IS-LESS. The latter translates to a knob, showing that people tend to understand clockwise rotation as a metaphor for more and counter-clockwise as less. The interaction can therefore also be placed on the sensorimotor level of the knowledge continuum. If this is so, why do our observations show possible interaction barriers? The problem could be related to the external consistency of the system. Meaning that although it uses well known conceptual extensions of image schemas, the consistency between how one understands the output of the system and the controls used to produce it is problematic.

The first part of this chapter indicates that sound technicians refer to the mix as a container or a room in where one places elements. If this is the common conceptual understanding of a mix, one-to-one mapped controls using image schemas such as UP-IS-MORE, DOWN-IS-LESS might not be sufficient. The last part of this section will therefore investigate research that focus on utilizing the theory of image-schematic metaphors in a different ways. Hopefully this can inspire more appropriate metaphors for the manipulation of a mix.

In an article done by Sabin and Pardo [29] they do not particular mention image schema theory but one could argue that they unconsciously utilize the theory. They do this by representing an equalizer by common auditory adjectives in a 2-dimensional space. Furthermore every point in the space has a particular color that reflects probability of our perception of the adjective. The mapping of color and adjectives are red = warm, yellow = bright, black = dark and green = tinny. This illustrates a less mathematical way of representing an equalizer and could allow for "more intuitive use". Unfortunately they do not test whether the interface is more intuitive in any way than the standard equalizer normally used when mixing music. Although we are not focusing specifically on a equalizer, the idea of mapping the mathematical data to how we perceive the sound is an interesting and relevant idea.

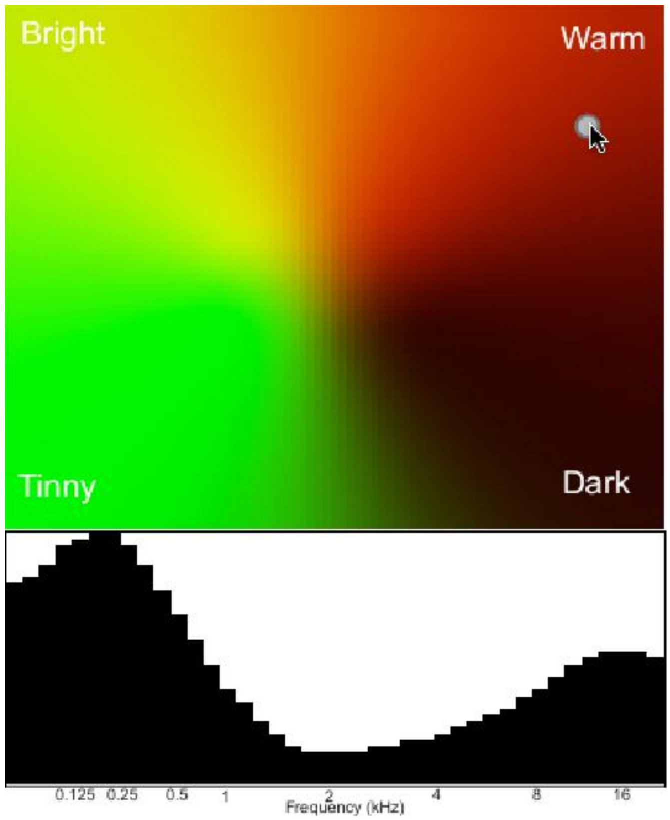



Figure 2.6: Screenshot of the 2Deq's interface

Another example of the usage of image schematics in musical interfaces is research done by Wilkie et. al. [37]. They developed the so called Harmony Space, a tool for playing, analysing and learning about harmony. This is a short study where three experienced musicians had to discuss some questions relating an excerpt of music. The data was then analysed to identify conceptual metaphors for the harmonic progression, the melodic movement etc. The result of this is that the Harmony space makes use of conceptual metaphors such as A CHORD IS A CONTAINER [FOR NOTES], HARMONIC PROGRESSION IS MOVEMENT ALONG A PATH and MODULATION IS MOVEMENT ALONG A PATH. Figure 2.7 is an example of how one of these metaphors is implemented an it illustrates how a harmonic progression can be shown as movement through the space following a path.

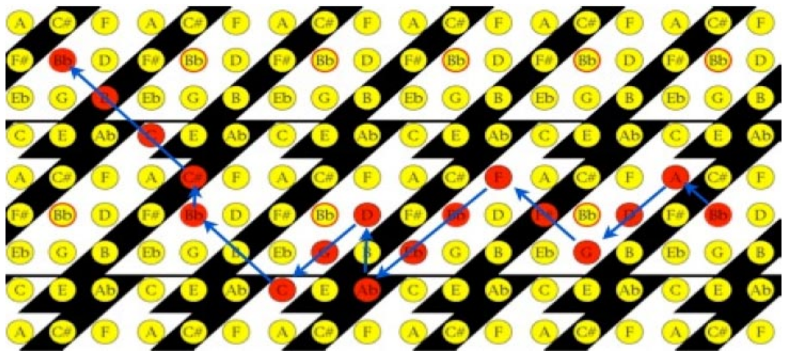



Figure 2.7: Screenshot of the interface of Harmony space

Wilkie et. al. concludes that musical interfaces could benefit from utilizing image-schemas. Their approach also illustrates the benefits of mapping conceptual extensions of image-schemas to our perception of sound and music.

The third, and most relevant related research, is an alternative mixer proposed by Jordá and Carrascal [12]. By analysing their interface it is clear that they use the image schemas SPACE and CONTAINMENT. They extend these by using a conceptual metaphor where tracks are placed in a musical space. The parameters such as volume and pan are changed depending on the location of listener and track. The prototype is a tabletop with multitouch and consists of a number of channels which can be dragged around the screen. The GUI, shown in figure 2.8, is divided into three sections - an inactive zone (at the top,(a)), the stage (in the middle,(b)) and a control zone (at the bottom,(c)). All channels are initially placed in the inactive zone and are therefore muted i.e. will not create any control data. When placed on the stage they become active and each channel's functions can then be accessed in the control zone.

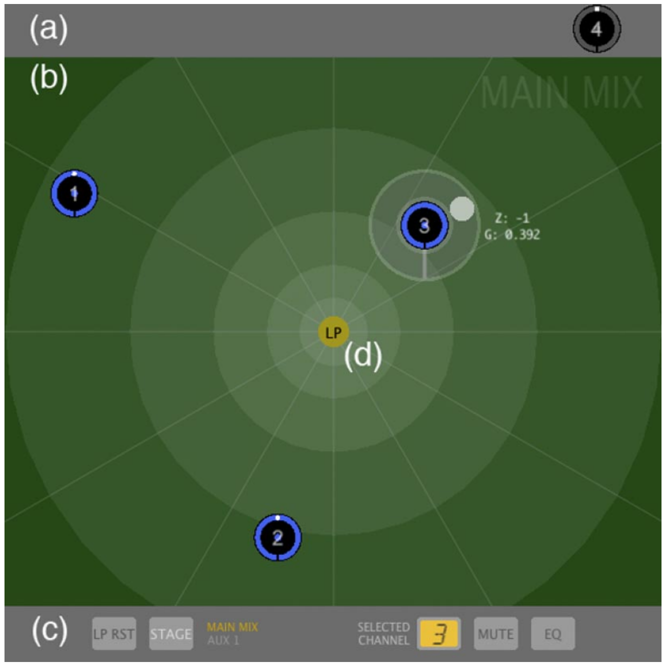



Figure 2.8: The GUI of Jordá and Carrascal's prototype

Jordá and Carrascal conclude that users show preference for the multitouch interface over a traditional mixing console. Furthermore they conclude that non-experts find the multi touch interface easier to learn and some users feel that it encourages creativity more than a standard mixer. However, they only test on 6 participants and only evaluate in relation to time efficiency and not how well the metaphor actually works. Their test can therefore only indicate tendencies and are not to be seen as significant results. Their system also has disadvantages: *The biggest drawback of the proposed system, as in any touchscreen, is the lack of tactile feedback. Traditional mixing consoles, with lots of physical controls, have a great advantage in this aspect.*[12].
Despite, their usage of image-schemas is very relevant to our project and the way Jordá and Carrascal uses them seems as a possible solution. We therefore find it interesting to continue their research, but instead we will focus on intuition.

The fact that they solely use a GUI for their implementation could be problematic in relation to the above discussed theory about intuitive use. A possible compensation for this disadvantage could be the use of a tangible user interface (TUI). This is due to the fact that TUI is said be "more intuitive", because they use physical objects. Since we learn how to manipulate physical objects throughout our lifetime [15] this interaction could be more easily placed on the sensorimotor level of the knowledge continuum.

Based on this, the next section will investigate TUIs and furthermore see how such an interface has been used in relation to music applications.

## 2.3 Tabletop tangible user interfaces

TUIs emphasize tangibility and materiality [17]. The conceptual structure is a physical embodiment of data and the purpose is to augment the real physical world by coupling digital information to everyday physical objects and environments. These objects can represent abstract concepts or real entities. Objects can be moved, turned around and all spatial changes can affect internal properties and relationships with neighbouring objects.
During the last decade we have seen several examples of such tabletops that make use of physical objects. These include ReacTable [13], Audiopad[26] and mixiTUI [27].

Reactable (figure 2.9) is a state of the art instrument, which seeks to be collaborative, intuitive and sonically challenging and interesting [18]. At the base it is a modular synthesizer, where tokens (tangibles) represent the sound generators, filters and modulators that usually constitutes a synthesizer. It enhances the interactivity between musicians who simultaneously use it. It furthermore displays in realtime all the information about controls and provides the user immediate visual feedback about his moves.

Figure 2.9: The Reactable

Audiopad[26] (figure 2.10) is a loop based tangible sequencer, using samples as sound basis. The Audiopad is operated by pucks, each representing a loop and by using modification pucks the user can navigate through options and menus, allowing addition of effects and changing loops.

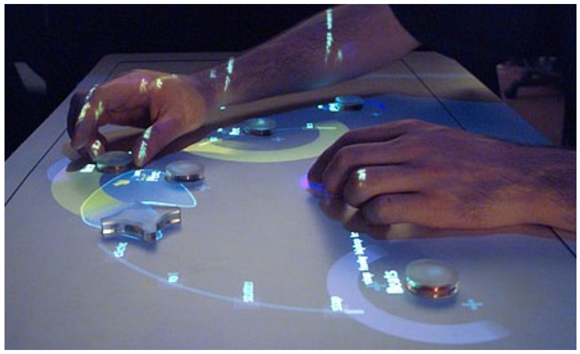


Figure 2.10: The Audiopad

MixiTUI [27] ( 2.11) is developed with the intention of live mixing and helps electronic musicians in their staging. It is a tangible sequencer that allows electronic musicians to import and perform their music live, while the audience experience visualizations of the music, created by the tangible interface. MixiTUI is operated by three categories of tokens - loop tokens, control tokens and effect tokens. The loop tokens are the sound producing part of mixiTUI, the control tokens allow for different motifs according to what part of the song is being played (verse, chorus, bridge ect.) and the effect tokens are coupled with digital sound processing effects.

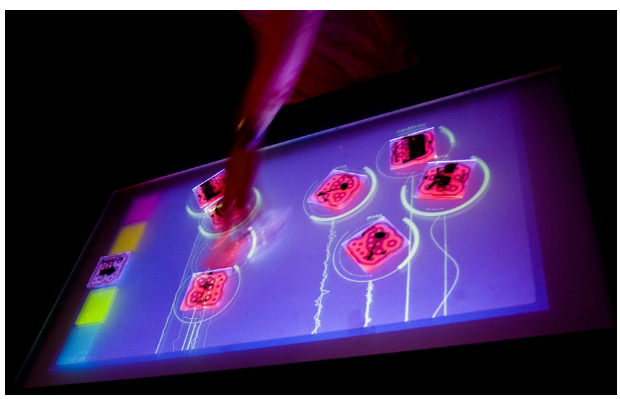


Figure 2.11: The mixiTUI

Reactable, Audiopad and mixiTUI have one thing in common; they are all a combination of a graphical user interface (GUI) and a tangible user interface (TUI). In the article "Emerging Frameworks for Tangible User Interfaces" Ullmer defines a TUI as something that embodies the digital information or functionality of the system in some sort of physical form [34]. All the parts of the underlying model, which is being manipulated by the interface, should be represented and manipulated by physical objects. The primary way a user interacts with such an interface is thus by spatially reconfiguring these physical objects i.e. physical placement,

translation, removal and rotation. A GUI on the other hand represents the underlying model of the system by graphics through a visual display, where the interaction is done by different input devices such as a mouse/keyboard or a touch screen.
This combination of GUI and TUI can result in a powerful combined interface and this combination will be investigated in the following section.

## Combining TUIs and GUIs

Cognition and memorization not only happens in the brain but also involves the physical world. Actions like counting using fingers, taking notes at a lecture and keeping a calendar are all examples of processes serving the purpose of minimizing cognitive load. A study by Norman and Zhang show that people perform better when information is represented externally by the system. This is because they do not have to memorize it internally[39]. TUIs take advantage of this by representing information in a physical way. This takes away cognitive load and therefore makes the interface more likely to feel "intuitive".

Norman and Zhang also note that the constrains created by physicality helps the user to determine which actions are possible and which are not. This again minimizes the cognitive load for the user and therefore is a step in the direction of "intuitive use". On the other hand these physical constraints can also limit the functionality. A good example of this is the BeatBearing[10] which facilitates an alternative way of controlling a drum machine, as seen in figure 2.12.

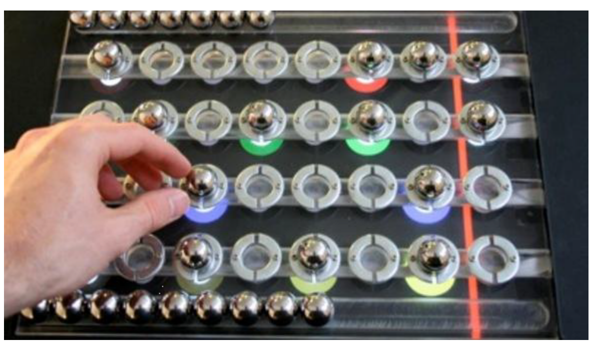


Figure 2.12: Picture of The BeatBearing in use

Although, the physical constraints make it easy to see which actions are possible, functions for this interface are limited, as it was designed and built for only one purpose. Ullmer states that TUIs can be more costly to develop since they often require non-standard components and sensors. They also tend to be more specialized than traditional interfaces. This makes them less economic viable than traditional interfaces as these later can serve a variety of purposes. As the mixing process consists of many sub processes, one could imagine that it could be beneficial to be able to change the design and functionality of the interface dynamically to fit a given sub process. Here a combination with GUI can be beneficial, as the GUI can easily be changed and switched. These are factors that make this type of systems a good choice for prototyping and therefore relevant for our research.

Another advantage of a combined GUI and TUI is the representation of complex concepts. For instance, when mixing music it is often required that a waveform is represented in different ways. This can be done much easier when using graphics rather than physical objects.

By combining a TUI and a GUI one can utilize the advantages of both. Developing a combined interface therefore seems as a good choice and could be used for many different purposes in relation to the process of mixing music.

All the discussed topics are based on theories, but how can they used in praxis? How could the mixing process benefit from these theories? To relate this theoretical knowledge to the mixing process and see what experts think, we conducted a workshop at the Rhythmic Music Conservatory in Copenhagen.

## 2.4 Future Technology Workshop

In section 2.2 we found that using a stage metaphor to represent a channel strip might be a reasonable alternative, but what do professionals actually think of this? And how do sound technicians meet the idea of a tangible based mixing device? Even more important do they actually see any limitations within the process of mixing and mixing console itself as it is now? Our initial interview with Morten Büchert indicates a need for rethinking the mixing console and its interactions, but is this a general opinion or is it more a matter of preference? Would requirements differ from the ones we have found already and how do sound technicians' ideas about the future of mixing look like?

These are some of the questions we tried to get answered by conducting a workshop at the Rhythmic Music Conservatory in Copenhagen. The design of the workshop were inspired by Vavoula and Sharples' method for designing new technologies and activities[35], and is referred as Future Technology Workshop (FTW). Four students in the field of sound technology and their teacher participated. In the following we will describe the different steps of the procedure together with a detailed description of our findings.

Figure 2.13: Conducting a Future Technology Workshop at the Rhythmic Music Conservatory in Copenhagen

### 2.4.1 Methodology

Usually a FTW consists of seven sessions with different goals, but we modified this framework in a flexible manner to make it more suitable for our research.

Our modified version of the workshop consisted of an introduction and three sessions, which are:

- A brainstorm session
- A modelling session
- Introduction the the realm of TUI's and fitting ideas to such devices

The brainstorm session served the purpose of thinking out of the box in the context of mixing music. By starting out with an unrestricted session like brainstorming, we hoped for the students to inspire each other and in general think out of the box, before restraining them. The important factor here is not to think about whether ideas are technologically possible or not, but only to generate ideas that are out of the box.

In the modelling session participants were asked to do low-fidelity prototypes of some of the ideas they developed i.e. paper based prototyping. This involved "building" or modelling one of the ideas discussed in the brainstorm session. In this phase they should still not focus on whether ideas are possible when it comes to technology, but more on the functionality of the prototypes. They were afterwards asked to present their ideas one at a time.

Finally we presented them to the idea of a TUI and the possible interactions with such an interface. This session served as a way of directing their thoughts and attention to the domain of the interface. This was done to convert their the ideas to fit an interface including tangibles and also get general feedback on the idea of a TUI.

During the workshop one researcher was writing down the ideas created during the brainstorm on a blackboard, while the other three researcher took notes. The whole workshop was audio recorded and can be found on the DVD in the folder *"Workshop at RMC"*. Furthermore can the complete procedure of the workshop can be found in appendix B.

### 2.4.2 Results

Before the workshop started, the participants were asked to fill out an initial questionnaire (see appendix C) in order to get a clear picture of their experience within the field of mixing music. The questionnaire shows that they all have at least 3-5 years of experience of both producing and mixing music. They all mix music professionally, but results are more split when it comes to questions about knowledge and interest in already existing and new music technology. All participants answered that they do to some extent have knowledge and interest in both existing and new music technology.

#### Brainstorm session

This session generated a lot of interesting ideas and several new ways of mixing music were proposed. One proposes the idea of a room were audio tracks can be represented by the objects within the room and the location of the objects will then define the mix. A concern about this idea is how would you change other parameters then volume and panning. Based on this one state hat a new system should give the user the ability to do *"everything you can do in a DAW"*. The idea about defining a room is based on the participants mental images of mixing music and one emphasize: *"that is your image, not everybody has the same image"*. On behalf of this they agreed that a new system should be able to be configured to individual users. This also led to the idea of downloading other user configurations. This would allow a user to download specific settings of his favourite artist (referred to as signature plug-ins).

Many times during the workshop the room metaphor was discussed and both a 2D and 3D interface was proposed. In the 2D interface pan and volume should be mapped to the X and Y axis and the 3D interface could make use of the Z axis to for instance represent frequency range. The advantage of a system using 3D is the possibility of controlling more parameters. A more concrete idea based on a 3D interface is based on glasses that create a 3D interface in front of your eyes allowing gesture based interactions.

In spite of the above mentioned ideas the students have some concerns regarding new technology. In general they have a hard time believing that such new ideas could be better than the environment they are currently using. This scepticism is mostly based on concerns about technical and precision wise limitations. Such concerns lead to their specific requirement that future products should have the exact same GUI as known from a Digital Audio Workstation (DAW). The only difference is would be, that the input should be through a touch screen.

One participant emphasizes the importance of mixing by only using the ears and as a practice he sometimes does a mix with the lights turned off. This let to a long discussion about how all the visual processes involved in the mixing process can ruin ones judgement regarding the sound, and one says *"the final user will not see the mix, but hear it"*. In order to comply with this a brain computer interface (BCI) is proposed. Ideally this would allow for total control by using only thoughts.

Several express a longing for the performance part of the mixing process. One participant tells about an old recording experience with his former band: *"Due to old recording equipment, where you had to do the mix in one take, band members all contributed to change parameters on the mixing console while recording, and this gave a performance feel to the session"*. Nowadays this way of recording is a rare occurrence and has been replaced by the infinite possibilities of the digital mixer and the DAW. The result is a more technical process rather than a creative process, using less devices. Based on this they propose a limit on how much "processing" you are allowed to do before a mix gets "destroyed". Because the systems for mixing music have become so complex, there arises a huge gab between having the idea and getting it out. The process often requires the user to go through multiple layers of software and one participant states that in order to solve this problem, *"there should be no layers in the interface"*. Some state that the problem is not so present when using analog mixing consoles, but *"even analog mixing consoles are too technical"*.

### Modeling session

This section will shortly present the specific prototypes developed, based on the ideas created during the brainstorm session.

One of the prototypes presented is based on the idea of having your ordinary DAW on a multitouch screen. This illustrates the before mentioned conservatism among some of the sound technicians.

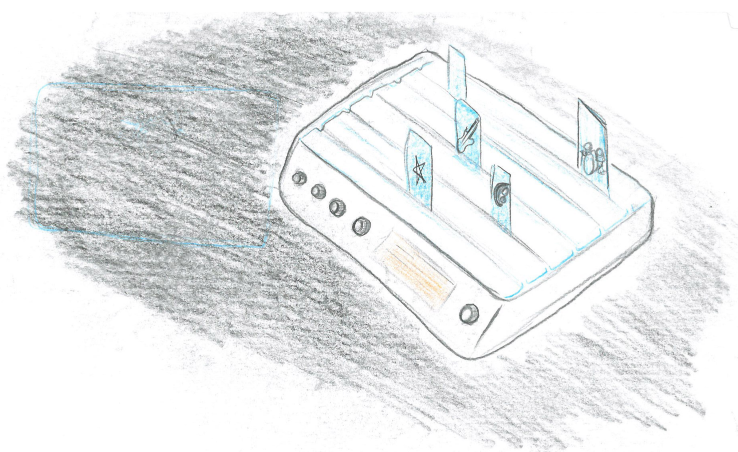


Figure 2.14: One of the drafts made during the workshop

The prototype seen in figure 2.14 is based on the idea of having a 2D interface but instead of a screen, you have a grid where you can place cards, representing the individual audio tracks. The X and Y position of these cards then represent the position of the audio tracks in the mix. In this system he also "implemented" some rotary switches which can be used to control different assignable parameters.

Another idea is based on a combination of using a 3D interface and being able to sculpt sound using your fingers. The participant explains that you have a window for each audio track where you have a "blob" with a lot of buttons representing the different parameters for this audio track. The user can then press a button and thereby control a specific parameter while sculpting the "blob". The participant explains that there should be another window where all the "blobs" representing the different audio tracks should be shown in relation to each other. This second window is based on the room idea in 3D discussed during the previous session.

**Fitting the ideas to a TUI**

When explaining the concept of a TUI and the way you interact with such an interface, the participants seem a little sceptic and "see" a lot of limitations within such an interface. Participants agree that a new system should be a complement to what they already have in a studio. This could for instance be a system where the user can try out some creative ideas in a fast and intuitive way and afterwards use all the traditional tools to edit this raw idea in a more technical manner. One participant state that he would prefer pure multitouch as *"you are only able to control two tangibles at a time, but with multitouch you can control all ten fingers"*. To this proposal some argue that it would be too hard to control something with all ten fingers.

### 2.4.3 Discussion

Some of the ideas discussed during the workshop are simply out of the scope of this research e.g. brain computer interface and 3D glasses. However these ideas are interesting and should be researched further. The idea of a brain computer interface arose in order to comply with the need for a fast way to get from idea to realization. This indicates that a future system should allow such directness.

The user should also be able to do everything that a DAW allows her/him to do and a system should be fully customizable in order to fit individual user preferences and/or mental images. On behalf of this they agreed that a system should be a compliment to their currently used technology.

In general there were a lot of concerns regarding a new system for mixing audio and it seems as the participants were kind of split. On one hand they liked the process of the system they use now i.e the mixing console, but on the other hand they also wish for a more intuitive system where creativeness and performance are more in focus. This can also been seen in some of the prototypes created by the participants. For instance the idea with the grid and the cards shows a new and simple way of mixing music.

A room metaphor were used several times in different contexts, which indicates that the stage metaphor might be a good alternative and we will therefore conduct further research on the stage metaphor.

When we introduced the participants to TUI technology they seemed a bit concerned and immediately found it very limited. However one of the requirements mentioned during the workshop evolved around customizability. Tangibles might therefore be a tool to expand possibilities for creating an analogy that fits the individual user's mental image. Some of the problems concerning natural mapping of different parameters on a 2D interface could also be solved with TUI's.

The participants were in general more conservative as expected. Even though new ideas were created, all participants seemed very pleased with the process as it is now. They feel comfortable with the mixing console and as one mentiones: *"it works"!*. This conservative mind set is also illustrated in the idea of implementing the traditional mixing console interface on a multitouch screen. This seems as the general way of thinking in the industry which is the part of the motivation for this research.

Even though the idea about a multitouch interface with tangible objects raised some concerns, it seems to be a good way to go since many of the ideas generated, could actually fit such a system. The participants also feel the need for already known input types, such as potentiometers, to get a more precise and technical way of controlling different parameters. Tangibles might actually suit this purpose of physicality and materiality. Further they might contribute to a more creative process with performance in focus.

Although the sound technicians seem very conservative about rethinking of the traditional mixer, they still have many ideas on how to make the process of mixing more intuitive. If we are to investigate whether an alternative system is more intuitive to use than a traditional mixer, using professionals could cause a lot of biases due to the fact that they are very used to using the traditional mixer.

## 2.5 System Requirements

During this chapter a lot of knowledge has been gained about how an alternative mixing console could be designed, to make the process of mixing music more intuitive. We will now sum up what we have gained and based on this knowledge formulate system requirements.

In section 2.1 *entering the realm of mixing* the goals of the mixing process were analyzed and we found that the tools used to achieve these goals include volume, panning, equalizing and effects. Therefore **our system should be able to adjust volume, panorama and frequency as well as give the possibility to add effects**.

We also found that the process of mixing constitutes multitrack material that is treated and combined into multichannel format. This means that **our system should be able to facilitate input of multiple tracks, which can be separately adjusted**.

The user interface should only present the user with information related to the successful completion of the task and only include the necessary interaction steps for performing the task. Further an alternate mixer **should not utilize layering of controls, but in-**

**stead try to implement each control so it is easy accessible**.

In section 2.2 *intuitive use* we saw that interaction with an interface is intuitive, when knowledge from the lower level of the knowledge continuum is involved. If a system is intuitive it leads to effective interaction and should reduce cognitive processing. We aim at the sensory motor level since situations of high mental workload and stress can result in a fall-back on lower stages of the knowledge continuum.

By using image schemas we have certain structural conceptual ideas on how to describe concepts such as space and containment and their actual strength lies in their metaphorical extension for structuring abstract concepts.

As the goal of this project is to research a more intuitive interaction during the process of mixing music, **our system should meet theses design principles**.

Section *tabletop tangible user interfaces* touch upon the fact that people perform better the more information are represented externally by a system, because they do not have to memorize it internally. TUIs take advantage of this by representing information in a physical way. This takes away cognitive load and therefore makes the interface more likely to feel "intuitive". The constrains created by physicality helps the user to determine which actions are possible and which are not. This again minimizes the cognitive load for the user and therefore is a step in the direction of "intuitive use". TUI's are likely to feel intuitive and natural because users have learned to manipulate physical objects through their lifetime.

TUIs in combination with GUIs are good for prototyping and usually can facilitate more functionality then purely tangible interfaces. Based on this **our system will make use of tangibles in combination with a GUI**.

Jordá and Carrascal's proposed stage metaphor showed positive results in their evaluation and in section *future technology workshop* several sound technicians propose similar ideas. By the use of a stage metaphor we can represent the mix in a more visual way. Presenting parameters such as panning and volume in this way seem to be more intuitive, than the representation used today and **our system will be based on this stage metaphor**.

# Chapter 3

# Design & Implementation

Until now we have focused on knowledge gathering, as we analysed the mixing process, looked at the benefits of combining TUIs and GUIs, introduced theories about "intuitive use" and conducted a workshop with experts. All this knowledge culminated in the requirements discussed in the previous section. This chapter will focus on how these requirements can be used for the design and afterwards how this design can be implemented as a prototype.

In section 3.1 we will first investigate how we can meet the proposed requirements. We will hereafter look at the aesthetics and why they are important. Lastly we will see how the GUI could facilitate the design principles for "intuitive use" introduced in section 2.2.2.

In section 3.2 we give a system overview of the final prototype and will furthermore explain the implemented hard- and software.

## 3.1 Design

### 3.1.1 Designing for functionality

This section will discuss how the system requirements can be woven into the design.
The the most straight requirement is that **our system will make use of tangibles in combination with a GUI**. We will therefore implement a multitouch interface which is also able to handle tangibles. As another requirement suggests, **our system should be able to facilitate input of multiple tracks, which can be separately adjusted**. Here we will use the just mentioned tangibles, where each tangible represents one track. A total of six tangibles seem to be sufficient for an prototype made for testing, as the most simple mixes use only one track per band-member. Each of this tracks has a set of parameters like low-,mid- and high-frequency as well as other effect parameters. Each of these parameters will be accessible through a menu, which is visualised at the bottom of the screen and can be controlled via touch 3.1. Through this menu we will have each function accessible without submenues and therefore meet the requirement that **our system should not utilize layering of controls, but instead try to implement each control so it is easy accessible**.

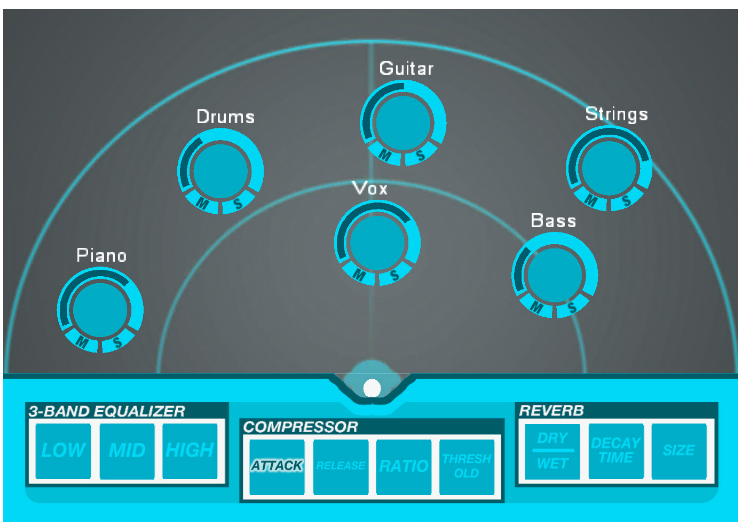



Figure 3.1: The graphical user interface

The parameters of each track will be adjustable through two different ways. The three frequency-bands as well as effect parameters will be adjustable, by selecting the parameter in the menu and then rotating the tangible. Volume and panorama is location depended, which means we will use the stage metaphor here. This means the volume and panorama will be adjusted relative to the position of the tangible and the listening point. This listening point is in the lower middle of the screen and can be seen as a microphone. If the tangible moves away from the listener point the volume of the track will be decreased, if it moves closer, the volume increases. The panorama works similar. If the tangible will be located to the left of the listening point, the track is panned to the left, if placed to the right, the track will be panned to the right. By doing so we will meet the following requirements: **our system should be able to adjust volume, panorama and frequency as well as give the possibility to adjust effects** and **our system will be based on this stage metaphor**.

The requirement that **our system should meet the design principles for intuitive use** will be elaborated in a later section, where we will go through each of the principles and explain how they will be applied.

### 3.1.2 Designing for aesthetic

As intuition can be classified as a part of usability, it is important to take also the aesthetic design into account as this contributes to the usability as mentioned in section 2.2.2. Therefore some basic aesthetic design choices had to be done, even though this is not the focus of this research.

Figure 3.2: Color Scheme

To not distract the user, by flashy colors, from listening to the audio, the overall color scheme was chosen to be different saturations of blue, as blue is the color which evokes the mood of calmness and comfort [36]. To present important states of the track like solo and mute, the colors green and red were chosen as they will give a easily perceptible contrast to the user, as seen in figure 3.3. Furthermore the color red is usually used to indicate off/inactive/no sound whereas green symbolizes a state of activity.

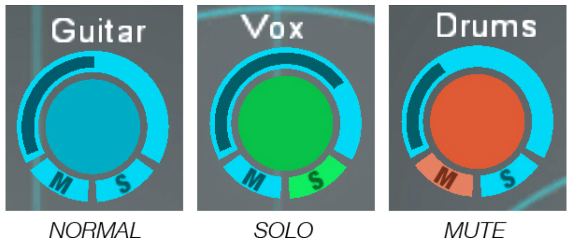



Figure 3.3: The three different states of the track

Besides the colors, the arrangement of the UI-elements are important to make the interface easy to use as well as nice to look at. Therefore we will us the most relevant Principles of Gestalt Theory in Visual Screen Design proposed by Chang [14].
The law of balance/symmetry is especially important for our interface, as the users goal is to make a balanced audio mix. Therefore the complete UI will be symmetric, otherwise it would be very hard to make a balanced mix. Another law which will be applied is the law of closure, which says, that our mind will close gaps and complete unfinished forms. This can be seen at the virtual "frame" around the tangibles. Even though it is split into solo-button, mute-button and slider, it is perceived as one object (see figure 3.3). The common fate principle adds to this perception of one object, as all the just mentioned elements move at the same time in the same direction, when the tangible will be relocated. The law of proximity states that objects placed near each other appear to be a group. This will be used in the menu, where the buttons for each effect are placed closer to each other (see figure 3.4).

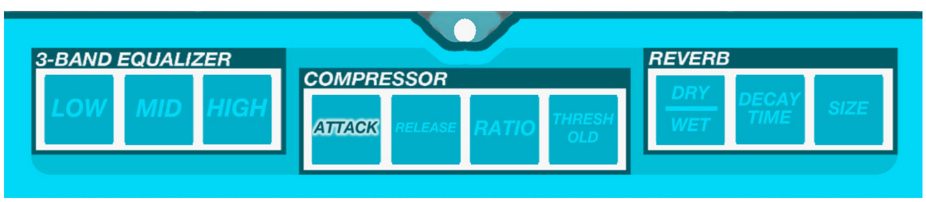



Figure 3.4: The menu based on the law of proximity

### 3.1.3 Designing for intuitive use

All the gestalt laws mentioned above are just one of the principles for designing "intuitive use". The other six principles used in the design process were described in section 2.2.2. The following paragraphs will explain how these principles have been utilized in the design process.

#### Suitability for the task

As mentioned, this means the user interface should only present information related to the successful completion of the tasks. Therefore it is essential to know what the task is and what is needed to fulfil the task. As seen earlier in the report, this was done by studying literature regarding to the mixing process (section 2.1.1) and observing a actual user while performing the task (section 2.1.3). By doing this we gained knowledge over the necessary tasks while mixing music, like level adjustment and equalizing. Like mentioned earlier, this and other tasks, like adding compression or reverb will be accessible through the menu, as seen in figure 3.4. There will be no unnecessary buttons, or buttons which can not be accessed during the current state of the system and it therefore only presents information related to the successful completion of the task.

### Compatibility

Based on its definition described in section 2.2.2, compatibility is close related to mapping, as mapping is defined as relationship between controls and their effects in the world. Norman [24] states *"good mapping is natural mapping"*, which means taking advantage of physical analogies and cultural standards. In the mapping of our interface we took advantage of the stage metaphor. Everybody has experienced that a sound emitting object moving away from the listener decreases in volume. The same happens with the panorama. The object which emits a sound to our left, is perceived louder at the left ear. Therefore in our interface the volume is mapped to the distance between tangible and listener and the panorama is mapped to the angle between tangible and listener. for the mapping of rotation of the tangible the physical analogy of a potentiometer is used. This is done because the user prefers some already known input types, as discovered in the workshop.

### Consistency

The consistency principle will be used in different ways in our design. For instance all tracks will behave in the same way. If moved away from the listener, the volume decreases, if moved closer to the listener the volume increases. This behaviour and all the others will be the same for all tracks, which means, if the user understands how to control one track, he/she will also be able to control the other tracks. The just described consistency is referred as *internal consistency*, as it is consistent within the system.
On the other hand *external consistency* is where we take advantage of the stage metaphor, as everybody (except deaf) experience the propagation of sound daily and can transfer this knowledge to the interface of our prototype.

### Feedback

Even though the primary feedback for a mixing interface should be auditory feedback, in some cases it is necessary to give the user visual feedback. For instance the buttons in the menu will give no auditory or tactile feedback when pressed via touch and therefore need visual feedback by changing the buttons appearance, as seen in figure 3.5.

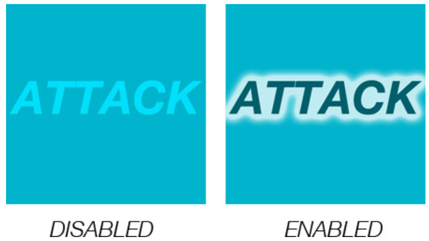



Figure 3.5: Buttons have different appearances, when in different states

The same will happens when using the multi-touch on the screen. As there is no tactile feedback if the user actually pressed the screen, there will be a yellow dot giving visual feedback that the user has pressed the screen.

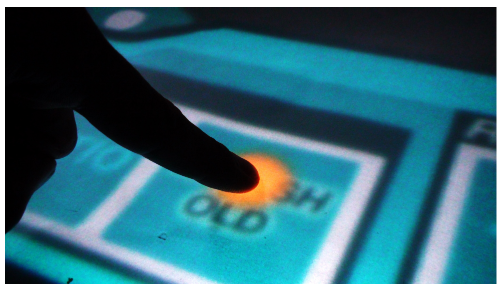


Figure 3.6: A yellow dot indicates where the screen is pressed

Furthermore all effect-parameters will be represented as visual sliders and when changing these parameters, by turning the tangible, the visual sliders will represent the state of the parameter (see figure 3.5). A graphical "frame" around the physical tangibles will give the user additional visual feedback, indicating that changes in the physical world are also applied to the virtual world (see figure 3.7). This is a very important part to connect TUI and GUI.

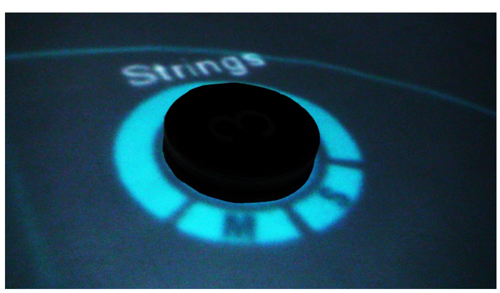



Figure 3.7: The "frame" around the tangible

#### Self descriptiveness

It should be obvious to the user in which interaction step he/she is in, where he/she is within the sequence of interactions, which actions can be taken and how they can be performed. By labelling the buttons with LOW, MID, HIGH aso., the user will immediately see which action he can take. Furthermore the buttons will change there appearance as described above, and therefore convey in which step the user is in. To convey the metaphor to the user a icon, which looks like a human with big ears (listening point) will be placed in the lower middle of the screen. Furthermore we will draw rings around this listening point, to make it more obvious how the space is related to the listening point.

#### Affordance

Another important thing, when designing interfaces is affordance. Norman defines affordance or perceived affordance, in the field of HCI, as *"the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used"*[24].
For this reason the tangibles will be flat and round. The flat shape gives the user not many possibilities how to place the tangible, as it has just two sides which it can "stand" on. A cube, for instance, could be placed in six different ways.
The round shape was chosen, because it takes advantage of users experience, that round things are usually used to rotate, hence can be turned.

Now, after all fundamental design choices were made based on the system requirements, the system can be implemented and afterwards evaluated. How this was done will be explained in the following sections.

## 3.2 Implementation

This chapter will give a detailed overview of our system and describe the hardware and software implementation in detail. Depending on the readers background some sections might seem quite technical and redundant. Such sections primarily serve the purpose of letting others replicate our set up, and the reader is encouraged to have this in mind when reading the chapter.

### 3.2.1 System overview

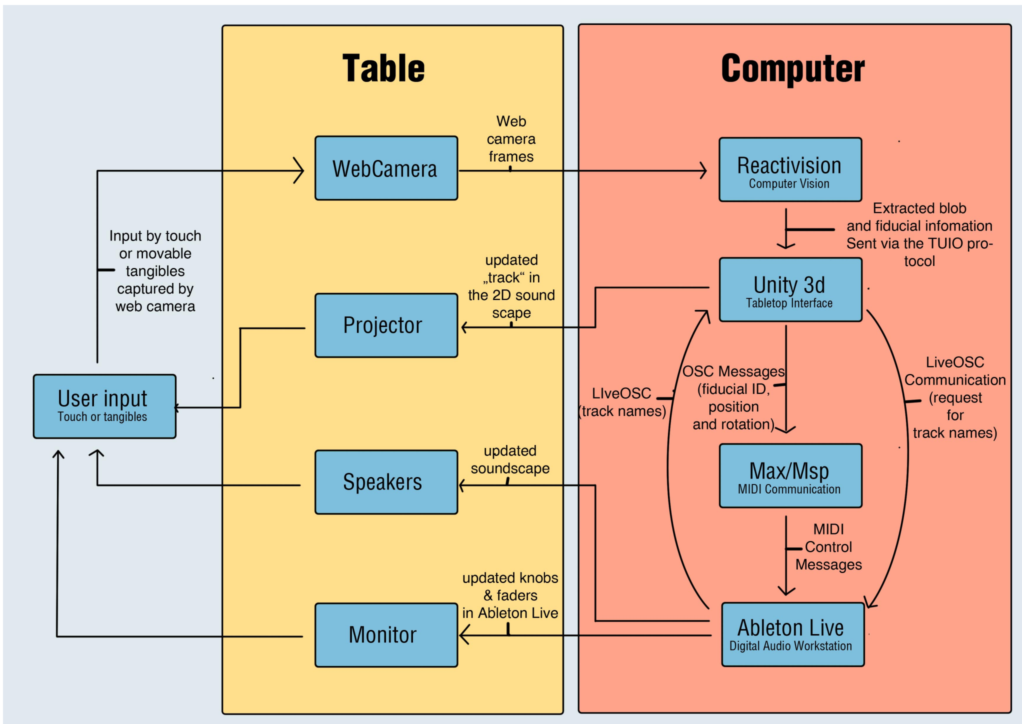



Figure 3.8: Overview of the system

Figure 3.8 displays an overview of the whole system as well as the flow of information. The system can be divided into three subsystems: The user, the table and the computer. These subsystems operate within the system environment (displayed in gray on figure 3.8). The user provides input to the system and the system reacts by providing auditory and/or visual feedback. The system takes two types of input: touch finger input and manipulation of round wooden tangibles. In relating to the touch-input, the system encompasses single touch and drag, two finger touch and drag, double touch and pinch scale. The touch input is primarily used for toggle buttons and therefore currently only utilizes single touch input. The system additionally registers x and y coordinates and angle of rotation of the tangibles. The input system is implemented using the image processing software Reactivision [6]. Reactavision processes web camera input. The processed information is hereafter transmitted to Unity 3d [9] via the User Datagram Protocol (UDP) using the TUIO packet wrapper [7]. Unity 3d uses the information from Reactivision to change the visuals of the interface accordingly. The interface is sent to a projector that, from below, beams it on the tabletop surface. Unity additionally sends information about the position and angle of the tangibles to the node programming framework Max/Msp [4]. This in turn converts the information to MIDI control data. This data is sent to the digital audio workstation Ableton Live [1] altering parameters such as volume, panning, equalizer and more. Unity 3d furthermore communicates with Ableton Live via the OSC protocol. This is used for retrieving additional information, such as track names. The interface of Ableton Live is shown on a computer monitor and the audio from Ableton Live is sent in stereo to two speakers. As mentioned in section 3.1, the system encompasses two types of feedback; visual feedback and auditory feedback. The first is primarily used to indicate changes in system state whereas the other are used to give feedback about the changes in the audio mix. The following sections will explain the system more in detail and are intended for readers who wants to know specific implementation details.

### 3.2.2 Hardware implementation

This part will give a small overview of what the multi-touch table consists of. There are usually three different methods to build an optical based multi-touch table: Frustrated Total Internal Reflection (FTIR), Diffused Illumination (DI) and Diffused Surface Illumination (DSI)[30]. As stated earlier, our table should be able to handle multi-touch as well as tangibles. Therefore the Diffused Illumination Model is used, as it is able to handle both. To be more concrete, rear DI was used, which means that the infrared light sources are placed in the box and illuminate the screen from below. Figure 3.9 gives an overview of the hardware, including (1)Table, (2)Projector, (3)PC Monitor, (4)Screen, (5)Tangibles, (6)PC, (7)Mirror, (8)Speakers, (9)IR LEDs, (10)Camera.

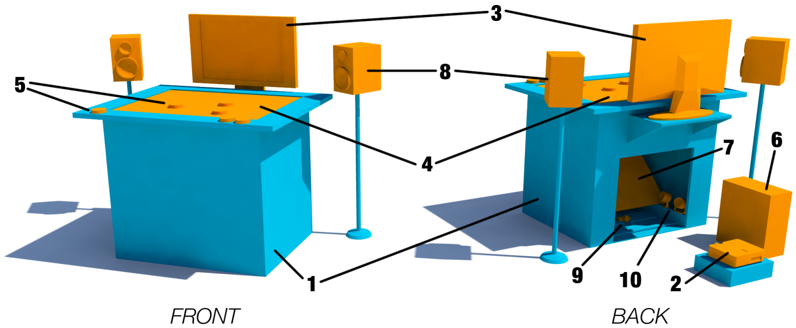



Figure 3.9: Hardware overview

#### The Table

The box for the table is mainly made out of particle-board and measures 100cm in length, 65cm in depth and 90cm in height. The inside of the box is covered with white fleece blankets, which reflect and diffuse the infrared light emitted by the four 48LED infrared illuminators. This is essential to get an even screen illumination. The multi-touch screen is placed on top of the box. It has a height of 61cm and width of 87cm. The screen is made out of an 8mm thick transparent acrylic sheet, which is frosted on the bottom side. This makes it less reflective and therefore gives a more even illumination. The acrylic sheet itself is still to transparent and absorbs to little light from the projector, therefore a layer of wax paper is put on top of the acrylic sheet, the so called rear projection film. This gives the screen the final layer and gives a bright enough picture.

#### The Camera

The camera is placed in bottom of the table, facing upwards. It should sense the fiducials and fingers, but not the projection on the screen. Therefore this information has to be separated in terms of their light spectrum. This is done by removing the infrared light filter of the Camera enabling the it to sense infrared light. Furthermore an infrared band pass filter, made out of the overexposed, processed photo-negatives, was added to "block" the visible light emitted from the projector. The camera used in our table is a PlayStation®Eye Camera which has a resolution of 640px * 480px and runs between 50-75 frames per second.

#### The Projector

The used projector is a View Sonic®PJ750 with a resolution of 1024px * 768px and a brightness of 2300 ANSI Lumen. As the projection angle is too small, the projector is placed outside the table and the projection is reflected by a mirror onto the screen.

#### The Fiducials

As mentioned above, the screen-width is 87cm and the camera-resolution is 640px * 480px. This means if the camera would be perfectly aligned to the screen, the camera would sense 1,36mm of screen per pixel. As the camera is not perfectly aligned, due to distortion, we got a good tracking when the smallest element to detect was around 5mm. Therefore the fiducials with 5,6cm in diameter are used to achieve a good tracking result.

### 3.2.3 Software implementation

This section describes the software implementation. It does so by explaining how each of the programs has been used in the implementation of the prototype. These are Reactivision, Unity 3d, Max/Msp.

### Reactivision

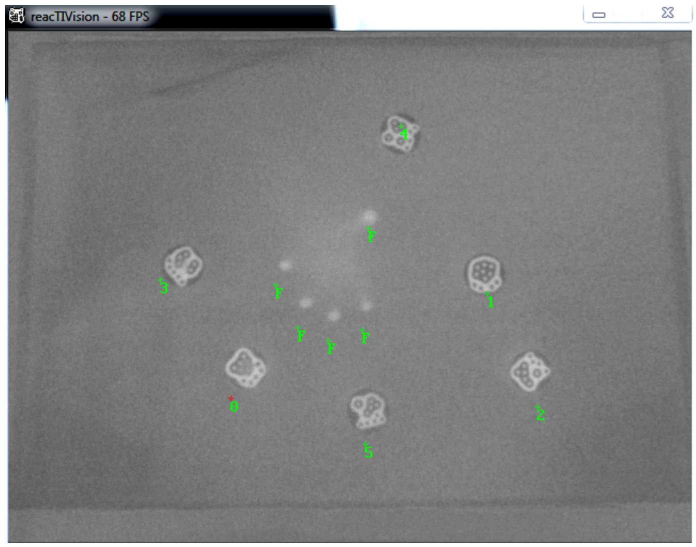



Figure 3.10: Screenshot of Reactivision running

Reactivision is an opensource image processing program originally developed Martin Kaltenbrunner and Ross Bencina for the Reactable project [6]. It detects the Cartesian coordinates and angle of rotation of fiducials (small figures shown in figure 3.10). It labels each fiducial with an ID. Fiducial ID, Cartesian coordinates and angle of rotation is transmitted to a specified UDP port via the TUIO protocol.

### Unity 3d

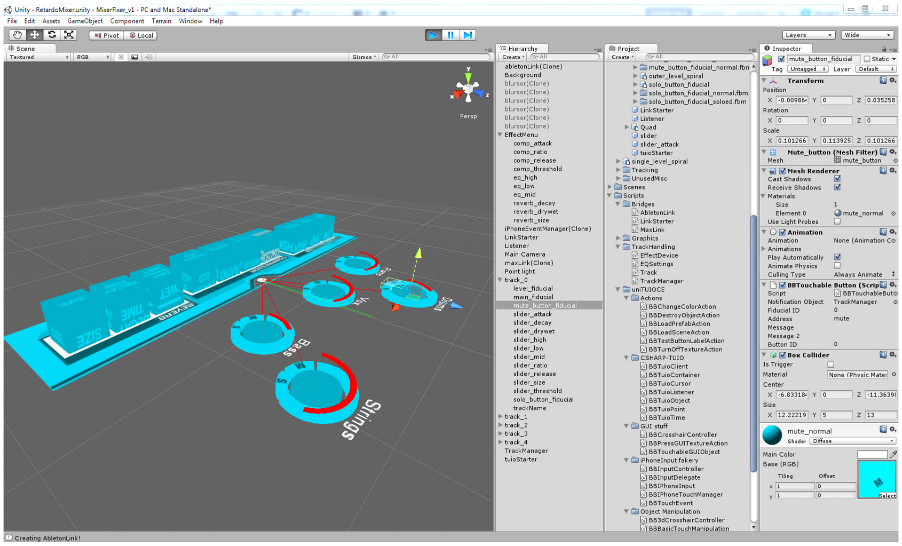



Figure 3.11: Screenshot of Unity 3d running

The table's interface is implemented in Unity 3d. Unity is a game engine usually used for 3d games. The interface program utilizes the uniTUIO library [8], which in turn uses the C# TUIO library, which again makes use of the open sound control (OSC) protocol. The OSC protocol is an extension of the UDP framework and it is commonly used in audio / media applications. The uniTUIO library and Reactivsion more or less provides an out of the box multi-touch implementation in Unity. We had to implement the fiducial communication ourselves, since there does not exist a library encompassing both. The uniTUIO library was therefore extended to incorporate this functionality.
Besides the TUIO communication, the most important scripts and objects in the interface program are:

- Fiducial Controller
- Fiducial game object
- Trackmanager
- Ableton Link
- Max link
- Listener

Each time a fiducal is placed on the tabletop Reactivision then tracks the fiducial figure and sends information carrying the ID and other information of the new fiducial. The TUIO implementation calls a specific function which notify the Trackmanager that a new fiducial is added. The TrackManager then instantiates a new fiducial game object in the Unity scene and assigns an ID corresponding to the ID of the fiducial figure. This game object has a Fiducial Controller script added to it. This script handles the individual fiducial's information and functionality. It does so by creating a link between the game object's Fiducial Controller and the TUIO implementation. Through this link the Fiducial Controller gathers all the needed information, such as x and y coordinates and angle of rotation. As explained in the design section (see section 3.1, The Fiducial game objects displayed on the tabletop as a menu which follows the tangible as the user moves it across the table. It does so by using the x and y coordinates received from Reactivision. Another game object representing the listener, is also present on the scene. The program calculates the relative angle between the fiducial and the listener. This angle is sent to Max/Msp by using the MaxLink instance and is later used for panning a track in Ableton Live. If the user rotates the tangible the Fiducial Controller script calculates the angle difference and transmits this as well. This is later used for adding effects in Ableton Live. The menu around the tangible also displays a mute and solo button aswell as the track name of the corresponding Ableton Live track. The mute and solo buttons work by communicating with Ableton Live via OSC using the LiveOSC communication library. As soon as one of these buttons is pressed by the user, Reactivision tracks the finger, which is transmitted in the same fashion as the fiducial information. Unity checks on each button if the coordinates of the tracked finger collides with one of these buttons. If this is true, it sends a OSC message in the format "/live/mute (track id)" or "/live/solo (track id)" to Ableton Live. Ableton receives this message and handles it accordingly. The tracks names from Ableton Live are received by Unity and these are displayed above the menu giving the user an overview of which tracks are currently on the table.

### Max/Msp

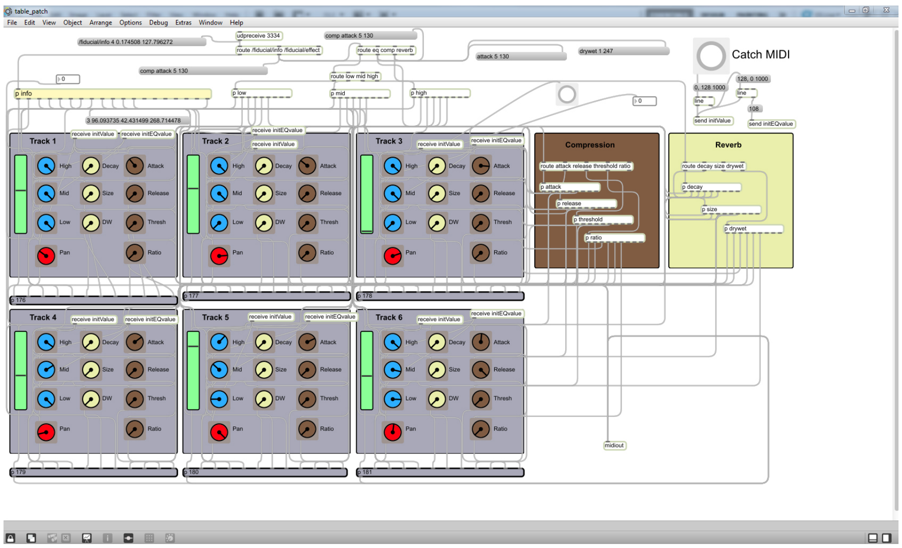



Figure 3.12: Screenshot of Max/MSP running

The Max/Msp implementation consists of one main patch (shown in figure 3.12) and several sub patches. It basically receives OSC messages from Unity through the "udpreceive" object. It then routes the messages and extracts the usable information. The OSC messages are formulated as such:

- /fiducial/(message group) (parameters)

An example of a message is

- fiducial/info (fiducial id) (x-coordinate) (y-coordinate) (relative angle to listener)

A described in the design section (see 3.1) each received fiducial ID is mapped to volume, pan and equalizing controls. These are converted to MIDI information in the sub patches and sent to Ableton Live through the "midiout" object. The final aim is to also implement the LiveOSC communication as MIDI communication. This will enable the prototype to be used with any DAW that utilities the MIDI communication protocol.

This ends this part of report which have described the main considerations of the design and how it has been implemented. The next chapter will focus on evaluating this prototype.

# Chapter 4

# Evaluation

This part of report serves the purpose of evaluating different aspects of the prototype. These areas include usability, intuitive use and areas of use.

Section 4.1 presents a usability test of the prototype. Section 4.1.2 follows up by focusing on the most prominent discovered errors and the corrections of these. This test serves the purpose of eliminating possible software bugs or usability issues, since these can influence the other types of evaluation.

Section 4.2 describes a test focusing on evaluating how intuitive use the prototype is, compared to a mixing console. This section includes a discussion on how one can test intuitive use as well as a presentation of the methodology and results of test.

The test focusing on intuitive use does not generate knowledge about which area of music production and performance could benefit from our research and prototype. The results from a focus group with primary focus on this specifically, will therefore be presented and discussed in section 4.3.

In section 4.4 the results from the tests are discussed in terms of validity and relevance.

## 4.1 Use-case testing - preparing for evaluation

### 4.1.1 Methodology

A user based usability test is found most suitable, since it involves representative users performing representative tasks [19] in opposition to expert based and automated use case tests. It is important to note that users are testing the interface, but are not being tested themselves. Since we have a functional prototype and high level choices have been made at this point, the foundation will be a summative test (also known as high fidelity prototyping). The goal is however not only to evaluate effectiveness of specific design choices, but also to incorporate more informal elements, which is based on communication between test moderator and participant. This will culminate in a combination of formative and summative test methods.

Metrics used in usability testing usually include measurement of task performance and time performance. Furthermore the methods utilized as part of usability testing include surveys to measure user satisfaction[19]. Since we are not focusing on time performance the primary data consist of task performance gathered through observations as well as oral feedback. The test subjects were therefore encouraged to think out loud while trying the prototype.

Preparation-wise, a task list and a script is formulated in order to ensure that all participants is given the same instructions and information. Examples of task's include placing tracks specific places in the mix, moving tracks around in the mix and adding effects to specific tracks. The script includes an introduction to our study and an introduction informing the participant about the overall structure of the test. Initially, before starting on the task list, the participants are invited to explore the table by themselves for approximately 5-10 minutes.

The test will be conducted in an apartment located on Nørrebro, Copenhagen. The participants are selected between friends - all with more than 5 years of experience in music and/or music production. Six will participate - two semi-professional hip-hop producers, three semi-professional electronic musicians and a professional producer mainly producing within the disco pop genre. According to [19] 7+-2 participants are usually enough when use case testing and is seen as a part of the HCI lore.
The fact that the test participants are among friends could introduce biases in their oral feedback, as friends can tend to hold back negative feedback. The observations are therefore seen as primary data from the use case test. An advance of inviting friends, on the other hand, is that they all are willing to participate over a longer period of time.

### 4.1.2 Results

Each participant allocated 5-10 minutes for trying the table, approximately 15 minutes for use-case testing and 5 minutes on debriefing. The whole use-case test ran for about 2-3 hours. Below is a short list of the most important findings:

- Receiving track names from Ableton Live does not work properly

- The tracks sometimes remained active on the table although the participant removed the tangible
- Volume does not map appropriately to the y-axis as the participants all seem to arrange the tangibles very close to the listener.
- Removing a soloed track does not cause it to unmute.
- The position of the effect sliders does not always correspond to amount of effect added in Ableton Live
- Each time the program is restarted one have to pick up the MIDI mapping. This means that the MIDI mapping is not working until all the mapped parameters in Max/Msp have had at, least once, the same value as the controlled parameters in Ableton Live
- Clipping occasionally occurred when the participant moved the tangibles to close to the listener point

All the listed bugs are fixed and this should ensure less issues in the following tests. The majority of the bugs were caused by small errors in the software of the interface program and the technical corrections of these will not be elaborated.

The mapping of the volume has been tried solved by logarithmically mapping the y-axis to volume, instead of simply using a linear function.

The MIDI mapping issues has been corrected by implementing an auto-pickup function in Max/Msp. This function initially sweeps all the controls Max/Msp through all the possible values for each control and thereby catching the MIDI mappings.

Lastly, the maximum input for volume has been decreased and this has eliminated the possibility of clipping.

Besides these issues, there was an overall problem with the calibration of the table. This meant that the tracking worked very poorly. It is crucial that this is working properly for the the following tests. We must therefore expect to allocate time of calibrating before conducted the next tests.

## 4.2 Test of intuitive use

This test serves the purpose of evaluating how intuitive to use the stage metaphor is compared with a traditional mixer.
As the stage metaphor only utilizes volume and panning, the effects menu, solo and mute buttons will be disabled during the test. To meet the simplicity of the metaphor, the mixing device to be used will only encompass a slider for volume and a knob for panning. We furthermore hope that this test will indicate how embedded conventional image schemes, such as UP IS MORE  DOWN IS LESS, are in our perceptional understanding, compared to physical analogies such as that of the stage metaphor.

### 4.2.1 Testing intuition

Intuition is a commonly misunderstood and misused concept. In section 2.2 we tried to clarify the concepts of intuition and intuitive use and used the following definition to describe intuitive to use system:

*"A technical system is intuitively usable if the users' unconscious application of prior knowledge leads to effective interaction"* [11].

Based on this we hypothesize that our prototype is intuitive to use if the users will be able to control the system by subconsciously utilizing their prior knowledge. The hypothesis to be tested in this test is therefore:

*Is our table prototype more intuitive to use than the traditional mixing console?*

This hypothesis alludes to the fact, that it is crucial to baseline how much prior knowledge the test participants have and how much knowledge is required in order to use the prototype. In order for the participants to utilize their prior knowledge they must not have any current knowledge about the system. Initially we had the idea of involving professionals i.e. sound technicians, but since this test group has extensive knowledge on controlling a traditional mixer it will not be appropriate to involve them in a comparative study like this.

Prior knowledge is an important variable factor and it is therefore important that we are able to control - to a certain extend - how much prior knowledge the test participants have. We strive for providing participants with more or less the same amount of prior knowledge and we realize that due to the complexity of how people acquire and store knowledge, this will only be possible to a limited extend. Despite, we find that it will be possible to control the level of knowledge to a sufficient extend, enabling us to generally conclude on how intuitive to use the prototype is compared with the traditional mixing console.
In order to eliminate possible interference caused by possible unknown prior knowledge of the test participants, we choose users without any prior music experience. In doing so we are, to a larger extend, able to provide them with the right amount of prior knowledge. Furthermore the knowledge continuum states that if novices find a new system to be intuitive to use, so will experts within the field (see section 2.2).

In this case, an important aspect of prior knowledge is constituted by image schemas embedded in our daily lives (see section 2.2). These image schemas can influence the test. For example, the LEFT-IS-LESS, RIGHT IS MORE image schema is commonly used on devices such as ovens, radiators and stereo music systems. These all commonly utilizes above mentioned image scheme in form of rotary controls. On the other hand, inter-aural level difference and inter-aural time difference creates the sensation of location and movement of sound sources [21]. This sensation falls under the categorization of prior knowledge and since it is a biologically rooted phenomenon every participant posses it. This sensation relates to the stage metaphor and can be seen as a more subconscious process, where as the image schema metaphors are learned through linguistics and motor skills. Both of these aspects of prior knowledge are possibly going to influence the consistency of the prior knowledge of each participants. This must therefore be taken into account when discussing the test results.

### 4.2.2 Methodology

The test procedure consists of three parts: A briefing section, a test section and a debriefing section. Below are the structure and purpose of each section described.

#### Briefing section

The briefing section serves the purpose of firstly introduce the test to the participants and secondly to provide them with the appropriate amount of knowledge. As explained, it is crucial that each participant is given the same level of knowledge and that this knowledge serves as prior knowledge when trying the prototype. A short manuscript has therefore been produced. This makes sure that even though the person conducting the test varies, the test participant will still receive the same level of knowledge. It is clear, that the nature of the person conducting the test is also influencing the outcome of the test; general guidelines for how to present the manuscript has therefore been created. It can be found in appendix D. The participants are asked to fill out an initial questionnaire base-lining their experience with producing music and mixing music. The participants are recruited based on the stratification variable that they do not have any experience with either, but the questionnaire serves the purpose of double checking that the participants are within the correct target group.

#### Prototype test

This part of the test follows the form of a within group test design. The tasks given in the test assert such a level of cognitive load that there could be big difference from test participant to test participant. The test design chosen is therefore a within group design. This is also due to the fact between group design requires more participants and is therefore also more resource consuming [19]. The test consist of two conditions where one is the prototype which half of the participant will start with and the other is a traditional mixing console which the other half will start with.

The prototype test consists of two parts: one where the test participants are given a number of tasks to be performed on the table prototype and one where the participants are to perform the same tasks on a traditional mixing console. During both tests the researcher will note down how well the participants perform the task. After each part the participant will be asked to fill out a questionnaire. The questionnaire is developed by [23] and can be found in appendix G. The questionnaire asks the participants to rate (by the use of a likert scale) how much he or she agrees with the following statements, where 1 = Fully disagree, 2 = Mainly disagree, 3 = Neutral, 4 = Mainly agree and 5 = Fully agree:

- I could use the system without thinking about it
- I achieved what I wanted to achieve with the system.

- The way the system worked was immediately clear to me.
- I could interact with the system in a way that seemed familiar to me.
- No problems occurred when I used the system
- The system was not complicated to use.
- I was able to achieve my goals in the way I had imaged to.
- The system was easy to use from the start
- It was always clear to me what I had to do to use the system.
- The process of using the system went smoothly
- I barely had to concentrate on using the system
- The system helped me to completely achieve my goals
- How the system is used was clear to me straight away
- I automatically did the right thing to achieve my goals

The questions are derived from Hurtienne's definition of intuition use (see 2.2.1). Hurtienne and Naumann operationalized this definition resulting in five different criteria which a user interface must implement in order for it to be intuitive to use. These are:

1. Low subjective mental workload
2. High perceived achievement of goals
3. Low perceived effort of learning
4. High familiarity
5. Low perceived error rate

The above 14 questions each fit into one of these five groups. When analysing the results afterwards one will be able to see how well the interfaces scored in each of the these categories.
**Debriefing**

The debriefing section will encompass a short open discussion with the participant on how well he understood the two devices. Since the questionnaire are only testing if the participants *feel* that the system is intuitive to use, this should validate that they also understood the functionality of the system.

### Sample size

The lowest aim is to conduct two initial pilot tests and 20 final tests. The initial tests should serve the purpose of pinpointing error sources and biases. See the next section (section 4.2.3) for a description of the error sources found and corrections made as a result of the pilot test.

### Test setup

The test is to be conducted in an apartment in Nørrebro, Copenhagen. Each time a test is conducted, every group member, except the one conducting the test, will leave the room. This is considered important since the participants will be asked to conduct tasks and a room filled with people could cause the participant to feel performance anxiety.

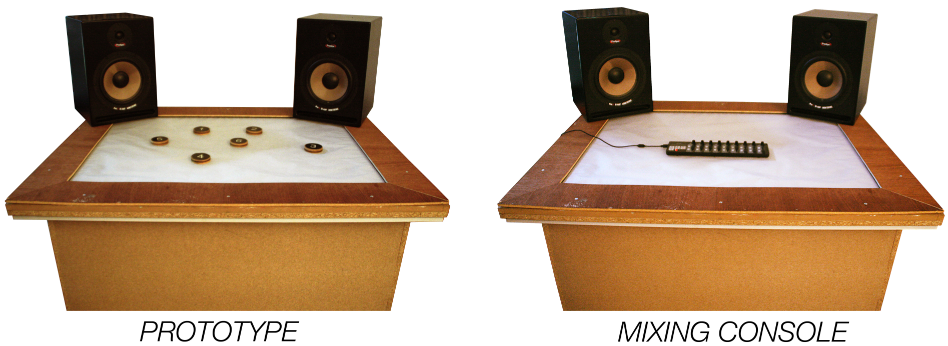



Figure 4.1: The test setup

The test setup consists of the prototype table, a MIDI controller and one set of speakers (see figure 4.1). The chosen MIDI controller is a Korg NanoKontrol. It suits our purpose as it has only faders for volume and knobs for panning and not many irrelevant buttons, which could distract the user and therefore influence the intuition. (see figure 4.2). The fact that it encompasses only volume fader and pan knob, which matches the functionality encompassed in the room metaphor.


Figure 4.2: The Korg NanoKontrol MIDI controller

The same Ableton Live project is mapped to either the MIDI controller as mixing console or the prototype table. The participants will in both tests be controlling six tracks and each track will have the same audio clips during every test.

### 4.2.3 Pilot test

The pilot test was conducted with 2 participants; one trying the prototype first and one starting with the mixer. Both participants were asked to think out loud if anything during the test was unclear. The investigator present furthermore focused on observing any confusions or misunderstandings.

The first finding concerned the questionnaire. In the questions we refer to the goals the participants are to achieve. Since the participants were given specific predetermined task they could not understand what goals in the questionnaire referred to. This has to be corrected before the final test. Another finding was in relation to the tasks that the participants were asked to do. The formulations of the tasks seem to have a big influence in how the participants perform the task. Consider these two formulations:

- Place track 4 in the middle of mix
- Place track 4 with no panning and half volume

Both questions seem to describe the same placement of an instrument in the mix. The first participant was given tasks of the first form where the other participant was given tasks of the second form. There seem to be a big different between how well they performed on the table and the traditional mixing console. The first participant understood the table very well, whereas the second had a better understanding of the traditional mixer. How one formulates the task list is therefore very important to consider. From the interview, workshop and observations with professional sound technicians we find the placement terminology to be most relevant. This is due to the fact, that one of the barriers of the traditional mixer is that it does not represent the room or the stage of the mix. This way of formulating the tasks therefore seems to do a better representation of the how sound technicians think about the mix.

A shortcoming of the pilot test was that both participants tried the table first. This was done, as testing the formulation of the tasks were found more important than generating an initial idea of how the test differs depending on which device is tested first. If one had chosen to vary both elements, it would have been very hard to establish causal links.

### 4.2.4 Test results

As explained in section 4.2.2, the test was conducted on 23 test participants. As the first 2 participants were used for pilot testing, the results from them will therefore not be considered when analysing the data. This gives us a sample size of 21 test subjects, and therefore a degree of freedom (df) of 20. The test participants of the final test were aged between 23 and 28, where 6 were female and 15 male. None of them had any experience in music production and music mixing.
As explained in section 4.2.2, the 14 questions in the questionnaire can be categorized into

subscales. To analyse how the prototype and mixing console performed in each of these subscales, the mean of each subscale has been calculated (see figure 4.3).

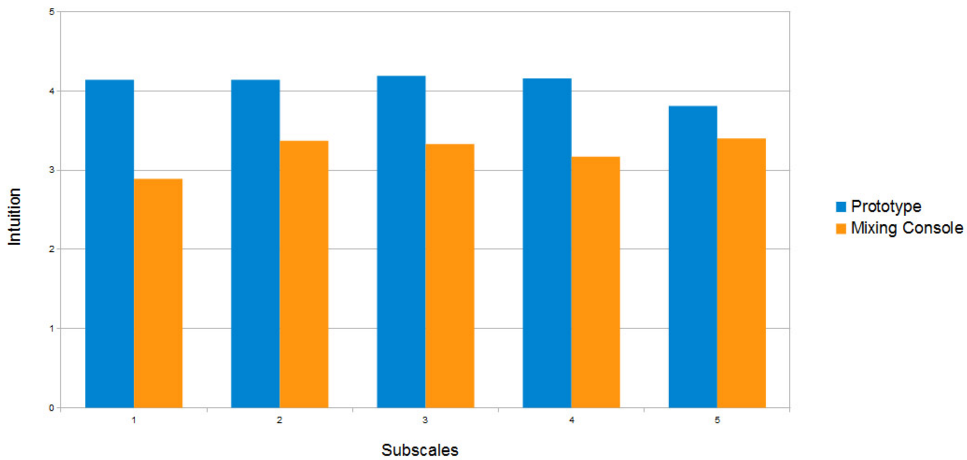



Figure 4.3: The mean of each subscale: (1)subjective mental workload, (2)perceived achievements of goals, (3) perceived effort of learning, (4)familiarity, (5)perceived error rate

This illustrates that the prototype had a higher value in all subscales, even though the difference between the means varies from subscale to subscale.
To get the overall intuition per person, the mean of five subscales per person were calculated, as seen in table 4.2.4.

| | Participant No. | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| **Mixing Console** | 4,47 | 3,6 | 2,53 | 2,87 | 2,33 | 3,3 | 2,87 | 4,57 | 3,5 | 3,57 | 3,43 | 4,43 | 3,77 | 3,13 | 1,3 | 3,6 | 3,63 | 2,23 | 2,93 | 3 | 2,83 |
| **Prototype** | 3,77 | 4,3 | 3,97 | 4,47 | 4,23 | 5 | 4,27 | 4,17 | 4,33 | 4,83 | 2,67 | 2,33 | 4,1 | 4,4 | 4,27 | 4,13 | 4,07 | 4,23 | 4,03 | 4,03 | 4,27 |

The mean of all the participants is 3,233 with a standard deviation of 0,784 for the mixing console. For the prototype the mean is 4,089 with a standard deviation of 0,596. (see figure 4.4).
To test if there is a significant difference between these means, a paired one-tailed t-test was applied to the data seen in table 4.2.4. Paired was chosen because the test was conducted as an within-group test. One-tailed because the test hypothesis indicates the direction of the difference: *"more intuitive"*.
We get a $t$ value of 3.477, which is higher then the $t$ value for the specific degree of freedom (df=20) at the 90% confidence interval. This specific value can be found in t distribution table [22].

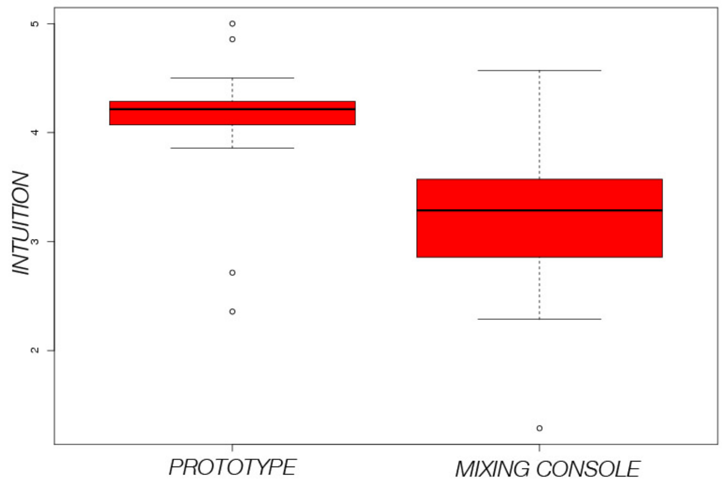



Figure 4.4: Intuition on Mixing Console vs. Prototype

The paired-samples t test suggests that the difference between the two means is significant, and therefore the interaction with the prototype was more intuitive then the interaction with the mixing console *(t(20)=3.477, p <0.10)*.

## 4.3 Focus Group on Areas of Use

### 4.3.1 Methodology

As supplement to the intuitive to use test and in relation to future perspectives, a focus group interview is conducted. The interview aims to get an insight into which areas the prototype could benefit as well as possible improvements. Furthermore, it focuses on future perspectives in relation to our research. The interview will have the form of an unstructured focus group interview. We are not able to interview sound technicians, but rather semi-professional musicians and producers. The interview does therefore not focus specifically on studio mixing, but rather encourages a discussion about possible improvements as well as all the areas where a system like ours can be used. This implies trying to figure out whether or not participants feel that they can use a system like this in their own musical processes. The overall questions, serving more as guidelines for the discussion, are of the form:

- In which situations do you think a system like this would fit?

The first part of the interview should help us understand their immediate thoughts on which area of music production this system naturally fits. The following question will focus more on their own possible use of the system:

- Could you imagine yourself using a system like this?

This question should serve the purpose of getting a discussion going on whether the feel their music production phase could be benefit from a system like this. It should naturally lead to the next question:

- Are you missing any features in order for you to use this system in your own music production process?

Ideally, this question will provide some knowledge about which possible improvement can be made for the prototype.

The same group of people participating in the use-case test participated in the focus group (see section 4.1). The focus group was lead by one researcher and two other researchers took notes during the interview.

### 4.3.2 Results

Before presenting the results, it is important to notice that the test participants had a background in producing and performing music and their answers therefore possibly tend to focus on how it could be used in their own area.

The participants more or less agreed that the prototype would be best suited as a live performance tool of some kind. There was not a consensus though in which live situations specifically, the prototype would fit. One tended towards the prototype being used as a live performance DJ-tool. Another advocated that it would fit best as a tool for live recoding mixes in a studio. This means that it could serve as some sort of automation tool for when recording automations in the mix. The other participants agreed that it could be used for live and studio mixing, but they were leaning towards live mixing. They argued that the visuals could then be used as a part of the stage show. This is an interesting idea. The Reactable, described in section 2.3, is in fact used by artists such as Bjørk as not only an instrument, but also as a part the visuals. One could imagine incorporating the sound technician in the visuals of the stage show when live mixing.

In relation to their own music production process, they all agreed that they could see themselves using it. Whether this would be serious usage or just playing around is hard to tell, since, as alluded to in the earlier conducted workshop (see section 2.4, many musicians tend to be conservative about their preferences for musical tools. Several of them additionally stated that one requirement had to be fulfilled if they had to use it; namely, the fact that it had to be smaller. This led to a discussion about how it would fit an iPad. They all agreed that it could be perfect for a tablet. The fact that you would not be able to utilize the tangibles did not seem to really concern them.

Another nice suggestion was that it could be used as a collaborative tool when bands discuss their mixes. This could also be translated to the situation where an sound technician and the artist have to agree on how the mix should sound (see section 2.1.1).

In terms of whether they found the prototype to be missing features, different suggestions were made. Most of the participants had a hard time relating the tangibles to the tracks in Ableton and suggested either color coding the tangible or using other means of visual mapping or representation of what is playing. One could argue that this could be influenced by the fact that they did not bring their own tracks and did not set the tracks up themselves. If they were using the prototype in their own process they could map it their to favourite DAW and they would also know the mix.

Another missing feature mentioned is gain input control. Furthermore two participants argued that they would like to have main mix and volume control. This is also an important aspect of mixing. These two functions should be implemented in future implementations.

Concerns about the automation feature were also expressed. The main concern where that a desired automation path could collide with other tangibles. One solution could be adding a feature which enables tracks to still be present even though one removes the tangibles from the table, but this could also cause confusion when one adds the tangible after the automation is completed.

Only one expressed that he had a hard time relating to the stage metaphor, but he expressed later that after he got to know it, he really liked it.

All participants stressed that they would like to be add effects for each track instead of having all effects added at once. This has also been a concern of our own, since as it is now, it poorly relates to how effects are normally added when mixing. On an off buttons on each effect could be a quick, but not necessary sufficient solution to this problem. Another possible solution could be a drag and drop feature, where the user could drag the desired effect to the track on which he wants the effect to be added. These suggestion will be discussion further in the future perspectives section (see section 6).

Overall the participants liked the prototype. Many of them expressed that they were thrilled about the fact that new research were done in this area. Many suggestions were made which allude to a lot of possible improvements. These improvement will be further discussion in section 6.

## 4.4 Discussion

Although the results show our prototype to be more intuitive one must consider how the nature of the briefing influences the test outcome. Our pilot test show that task formulations have an influence on one part of the external constancy of the system i.e. the balance between how much knowledge the user has and how much knowledge is required to use the prototype. The knowledge provided for the participants could overwrite their usage of prior knowledge stemming from the low end of the knowledge continuum. Despite, the fact that the participants were provided knowledge about both the stage metaphor and the mixing console can still be used as argumentation for the table prototype being more intuitive. How "low level" this knowledge is - in relation to the knowledge continuum - is however hard to specify.

Pin pointing how much prior knowledge each participant had before testing is complicated by the fact that the content written in the manuscript to be insufficient in terms of understanding the concepts. The concepts were therefore further explained in such situations. This was done in order for the participants to have the required knowledge to operate both systems. This could have caused some of the participants to have a higher level of prior knowledge than other participants and thereby influence their rating of intuitive use. On the other hand the difference between how intuitive the participants find the devices will arguably be evened out by the fact that participants had the same extra knowledge before trying either device.

Participants that started out by testing a traditional mixer, initially tried to turn up the volume on the prototype by moving the tangibles away from the listening point. This could be because they used their prior knowledge acquired through testing on the traditional mixer first. It could also be the influence, explained in section 4.2.2, of prior knowledge stemming from everyday devices. Whether or not, it shows how much the UP-IS-MORE, DOWN-IS-LESS metaphor is embedded in their intuitive understanding of a prototype. The participants starting with the prototype, on the other hand, had problems understanding that volume and panning are relative to the listening point.

In general the participants found it easier to achieve specific tasks on the prototype compared to a traditional mixer. As explained in the pilot test section (section 4.2.3) the nature of the task formulation is definitely influencing this, but the formulation is found most appropriate in relation to the terminology of sound technicians. Despite, these formulations are only based on our observations and interviews with sound technicians and could therefore still be formulated wrongly. An approach could therefore be to get sound technicians to formulate these questions for another validation of how intuitive the prototype is to use.

It is important to note that the test on how intuitively usable the prototype is only evaluates the stage metaphor/tangible objects and not the effects menu. Results from this test do therefore not provide us any knowledge about how intuitive to use these menus are. The focus group interview on the other hand gave us some initial indications. These allude to the fact that our chosen method is not the best way of implementing effects. By default each channel has all effects added automatically and these cannot be turned off. The only way to remove effects from each channel is by turning down effect parameters. Possible future solutions to this are discussed in the future perspective section 6.

The focus group provided us knowledge on which areas of music production could benefit from the prototype. This knowledge is though limited due to the fact that it is knowledge generated from only a single focus group interview. Despite, this knowledge can still serve as basis for an interesting discussion. The fact that semi-professional musicians like the prototype and thinks it would be well suited for live performance, indicates that the findings in this project should not only be seen as studio mixing specific. A modified version of the interface could prove very usable in live performance situations. In regards to studio mixing the participants suggest that the prototype could be a good tool for collaboration if more people are in the studio at ones.

The process of mixing is much more visible on our prototype than on a mixer. This can still inspire sound technicians to rely too much on their visual sense, but the focus group indicated that there also could be advantages of the visuals in live situations. The visual part of the interface has not been the focus of this research, but this finding indicates that the visual design plays a bigger role than we have given it.

The participants all mention that our prototype would easily be integrated into their own musical process. As the prototype encompasses another way of mixing this could allow for a different musical expression. Whether sound technicians share this point of view is unclear, but if this is the case then this different means of expression could be an indication of the barriers haven been broken.

# Chapter 5

# Conclusion

The direction for the current development of mixing consoles is problematic. It is moving away from analogue mixing towards digital mixers that rely too much on the visual sense and the current mixing tools rely too much on an underlying mathematical representation. Alternative interfaces for mixing can therefore help in making sound technicians rely more on their intuition and their auditory sense. Despite, sound technicians can be very conservative about maintaining the traditional mixer as the tool used in their process.

Interfaces are intuitive to use when the user can subconsciously use his prior knowledge to operate the system. The more the knowledge required to use the system relies on knowledge learned through childhood the more likely it is that the interface is found to be intuitive. Using conceptual extensions of image-schematic metaphors can be a powerful tool when designing interfaces for intuitive use. Furthermore, TUIs can help in making interfaces rely on this acquired knowledge in the fact that the utilized physical objects which the user can manipulate. By combining a GUI and a TUI one can make use of the advantages of both, making the system tangible and dynamic.

In this project we propose an alternative mixer, which is based on a tabletop TUI and make use of a stage metaphor

Three tests were used to evaluate this prototype; a use-case test, a quantitative test evaluating how intuitive to use the prototype is compared to a traditional mixer and a focus group with semi-professional musicians. The test results suggest that a stage metaphor is better when designing for a more intuitive way of mixing music. However, biases question the validity of the test results.

Further our results suggest that this prototype could be suited for live mixing, live performance, collaborative mixing or live recording when mixing in a studio.

# Chapter 6

# Future perspectives

This research was a very general research about rethinking the mixing process. As explained through the report there are different areas that approach the process in different ways. One way to go from here could therefore be to use the knowledge gained in this research and further develop the prototype with focus on a specific area such as live or studio mixing.

It was original planned that part of the evaluation would be interviewing sound technicians in order to evaluate the system and get there opinion on how this kind of interface could be used in the mixing process. This was not done, since they cancelled the test. In a future study this is a crucial evaluation and could be relevant to do in order to get new knowledge about the direction future research should pursue.

A possible way of doing this test could be setting the prototype in a professional studio and letting sound technicians use it over a longer period of time. Iterations could then be made with focus of improving the functionality of the prototype. The iteration steps could be letting sound technicians use the prototype, have them report back about issues or improvement to be made and this could then be corrected before the next iteration.

Another aspect of the prototype that needs further research is how to apply effects. Since the focus group showed possible issues with how the effects were implemented. Future work on finding new ways of representing equalizing, reverb etc. could be very interesting. More specifically future research should focus on finding new ways of representing the underlying system models for the audio effects used when mixing music. One way could be using a system like the 2DEQ, mentioned in section 2.2.2. Another way could be using some metaphors for the patch cables used in the analogue approach to mixing.

In relation to effects, the prototype lacks functionality similar to aux or sends channels. These are commonly used in mixing and would possibly be missed if sound technicians had to use it professionally. Jordá and Carrascal's prototype described in section 2.2.2 makes use of sub spaces where one, additionally to the mixing state, also has a spaces for the effects. This is though an problematic approach when using a TUI. The fact that the tangibles can not be moved by the system makes it not very suitable for automation, saving mixes and sub rooms. If one could implement the table so the system also can manipulate the tangible objects, it would open up for a lot of very interesting solutions to the just mentioned problems. Although one has to be careful when layering menus. One interesting implementation could be the possibility of switching between different views. In this way the user could switch to an equalizing mode where the tangibles would control different equalizing parameters. If automatic movement of the tangible could be implemented they could move automatically to the right location when switching mode.

In terms of visuals, implementing the graphics dynamically so they respond to the sound would really improve the prototype. One could imaging having the frequencies of the tracks presented in the background. By doing so one could more easily spot masking of tracks. This could on the other hand also be a too visual way of doing it.

Finally a system that allows for easier midi-mapping of the tangibles could transform the prototype to a very interesting midi-controller. As it is now one can map the x and y coordinates as well as the angle of the tangibles to almost all parameters in DAW's. This means that it is not limited to only controlling volume, pan and different effects. One could imaging controlling envelope on a hi-hat or controlling the keys on a piano. This would broaden the use of the system and allow for full configuration to the individual needs.

Another way to go in future research is to investigate how other kind of technologies could benefit the mixing process. TUI is only one way of doing it and the workshop in section 2.4 showed many other interesting ideas. The idea about sculpting sound is very interesting. This would be a very physical way of mixing and such a system could also be very intuitive to use. This is just one example, but we suggest future work could focus on mapping gestures to the controls of a mixer. This could for example be "throwing" effects at instruments or other more active ways of mixing.

One concrete example of a technology that was considered in the early phases of this research is virtual environments. Wozniewski[38] has created a framework and a programming library for manipulating sound in a virtual environment. The reason why this research could be relevant for future work is that the framework he suggests could be used for mixing audio or other artistic purposes. This differentiate his work from many other researches and tools in this area which mainly focuses on realism. A future study could be start by letting people listen to music in headphones afterwards ask them to draw or explain how they mentally perceives the music. Do they all, for example, imagine some sort of sonic space or are there other ways of imaging sound? Data from such a research and the framework by Wozniewski could then be used to create a virtual environment where one could enter the mix. This could be basis for an alternative mixer that might be extremely intuitive to use. This was chosen not to do in the start of the project due to time constrains.

Although this research did not focus on the commercial market. But there could also be a great potential in designing a system for novices who want to try mixing. This could for example be a multi-touch version of our interface for tablets. This idea could also be extended to educational purposes where one could have a split screen between mixer and our interface. In this way, one could see how the traditional mixer is manipulated by using our prototype.

Lastly, one problem experienced during this research is that, when investigating different design approaches for interfaces allowing for intuitive use, theories suggest that one utilizes the prior knowledge of the user. This can though limit the potential of new technologies and completely new ways of interaction. This is a problem that not only concerns the mixing process but all other areas where systems could benefit from a new representation of the underlying system model. It could therefore be very relevant to investigate how to introduce new interfaces with completely new forms of interaction in a way that allow for intuitive use. A motivational focus in this research was initially on how the expression of a mix alternates when one uses other means of interacting when mixing. We hoped to investigate whether more intuitive ways of mixing could allow for alternative expressions of the mix, but since so little research had been done, starting on a lower level was necessary. Despite, this is still a very interesting direction that future work could pursue.

# Bibliography


[1] Ableton live 8. http://www.ableton.com/live-8.

[2] International conference on new interfaces for musical expression. http://www.nime.org/.

[3] M83's official website. http://ilovem83.com//.

[4] Max/msp - cycling74. http://cycling74.com/.

[5] Pro tools. http://www.avid.com/us/products/family/pro-tools.

[6] Reactivision - a toolkit for tangible multi-touch surfaces. http://reactivision.sourceforge.net/.

[7] Tuio software. http://www.tuio.org/?software.

[8] unituio library. http://www.xtuio.com/index.php/projectsmain/utuiomain.

[9] Unity 3d - game enigne. http://unity3d.com/.

[10] P. Bennett and S. O'Modhrain. *The BeatBearing: a Tangible Rhythm Sequencer*. Proceedings of NordiCHI 2008: 5th Nordic Conference on Computer-Human Interaction (electronic proceedings), 2008.

[11] A. L. Blackler and J Hurtienne. *Towards a unified view of intuitive interaction: definitions, models and tools across the world*. Zentrum Mensch-Machine-Systeme, 2007.

[12] J. P. Carrascal and S. Jorda. *Multitouch Interface for Audio Mixing*. University of Oslo and Norwegian Academy of Music.

[13] J. P. Carrascal and S. Jorda. *The Reactable: Tangible and Tabletop Music Performance*. Proceedings of the 2010 ACM SIGCHI 28th ACM Conference on Human Factors in Computing Systems, 2010.

[14] D. Chang, L. Dooley, and J. E. Tuovinen. *Gestalt theory in visual screen design: a new look at an old subject*. Australian Computer Society, Inc., 2002.

[15] C. Huang. *Not just intuitive: examining the basic manipulation of tangible user interfaces*. Association for Computing Machinery, 2004.

[16] J. Hurtienne and L. Blessing. *Design for intuitive use - testing image schema theory for user interface design.* International Conference on Engineering Design ICED07, 2007.

[17] S. Jordà, C. F. Julià, and D. Gallardo. *Interactive surfaces and tangibles.* Association for Computing Machinery, 2010.

[18] S. Jorda, M. Kaltenbrunner, and R. Geiger, G.and Bencina. *THE REACTABLE.* Universitat Pompeu Fabra, 2005.

[19] et al Lazar, J. *Research Methods in Human-Computer Interaction.* Wiley, 2010.

[20] N. Liebmann, M. Nagara, J. Spiewla, and E. Zolkosky. *Cuebert: A New Mixing Board Concept for Musical Theatre.* University of Technology, Sydney, 2010.

[21] Garath Loy. *Musimathics - The Mathematical Foundations of music - Volume 1.* The MIT Press, 2006.

[22] W. Mendenhall, R. J. Beaver, and B. M. Beaver. *Introduction to Probability and Statistics.* Cengage Learning, 2012.

[23] A. Naumann and J. Hurtienne. *Benchmarks for Intuitive Interaction with Mobile Devices.* Association for Computing Machinery, 2010.

[24] D. A. Norman. *The Design of Everyday Things.* Basic Books, 2002.

[25] B. Owsinski and M. O'Brien. *The Mixing Engineer's Handbook.* Mix Books, 1999.

[26] J. Patten, B. Recht, and H. Ishii. *Interaction Techniques for Musical Performance with Tabletop Tangible Interfaces.* Association for Computing Machinery, 2006.

[27] E. W. Pedersen and K. Hornbk. *mixiTUI: A Tangible Sequencer for Electronic Live Performances.* Association for Computing Machinery, 2009.

[28] Izhaki R. *Mixing Audio: Concepts, Practices and Tools.* Focal Press, 2008.

[29] A. T. Sabin and B. Pardo. *2DEQ: an intuitive audio equalizer.* Association for Computing Machinery, 2009.

[30] Johannes Schöning, P. Brandl, F. Daiber, F. Echtler, O. Hilliges, J. Hook, M. Löchtefeld, N. Motamedi, L. Muller, P. Olivier, T. Roth, and U. von Zadow. *Multi-Touch Surfaces: A Technical Guide.* Technical University of Munich, 2008.

[31] M. Senior. *Mixing Secrets for the Small Studio.* Elsevier Science, 2011.

[32] H. Sharp, Y. Rogers, and J. Preece. *Interaction Design: Beyond Human-Computer Interaction 3. edition.* Wiley, 2011.

[33] N. Tractinsky, A. S. Katz, and D. Ikar. *What is beautiful is usable.* British Computer Society, 2000.

[34] B. Ullmer and H. Ishii. *Emerging frameworks for tangible user interfaces.* IBM Corp., 2000.

[35] G. N. Vavoula and M. Sharples. *Future technology workshop: A collaborative method for the design of new learning technologies and activities.* International Society of the Learning Sciences, Inc.; Springer Science + Business Media, 2007.

[36] L. B. Wexner. *The degree to which colors (hues) are associated with mood-tones.* American Psychological Association, 1954.

[37] K. Wilkie, S. Holland, and P. Mulholland. *Evaluating musical software using conceptual metaphors.* British Computer Society, 2009.

[38] M. Wozniewski. *A framework for interactive three-dimensional sound and spatial audio.* McGill University, Department of Electrical and Computer Engineering, 2006.

[39] J. Zhang and D. A. Norman. *Representations in distributed cognitive tasks.* Elsevier, 1994.

# Appendix A

# Interview guide for the interview with Morten Büchert

**First interview with Morten Büchert** Place: Staalvand caf Date: 23-02-2012

**Methodology** The interview was conducted at Staalvand caf in Brønshøj. Interviewee was Morten Büchert, a sound technician, producer and assistant professor at the Rhythmic Music Conservatory in Copenhagen. Present were also Steven Gelineck, the supervisor of this project as well as each member of the project group. Only one from the project group asked the questions.

The interview was conducted in the very initial state of the project. In this state of the project we prioritized to go in-depth and generating rich data. The interview served the purpose of being an initial exploratory interview. The interview can therefore be categorized as an unstructured interview with traits from the semi-structured form of interviews. This was chosen as the interview was conducted, research-vise, on a somewhat clean slate. Furthermore the interview focused on different aspects within the area of mixing music, which is a complex and multifaceted process. An exploratory interview therefore seemed appropriate[19]. Since the interview was conducted very initial in the project phase, it took mostly the form of a conversation, since this often can help with going in-depth with a particular topic[32][19].

Exploratory interviews often focus on the needs and challenges presented in a situation, rather than specific functionality or design issues [Research methods in HCI, p. 181]. Our focus were as well on the needs and challenges, but since this phase of the project also focused on defining exactly what constitutes the process of mixing, we also incorporated a few questions concerning design and functionality. The structure of the questions was open-ended, because of our limited knowledge within the area. The structure of the interview was defined as three parts. The interview started with an introduction or briefing, where each of the attendants introduced themselves. Hereafter after one from the project group briefly introduced the project area. The research was kept brief as a detailed description could cause the interviewee to provide answer according to the project groups aims and interests. This introduction also served the purpose of recording the knowledge and expertise of Morten.

After the briefing, the interview started by the interviewer asked a broad question about mixing. From here each time the interviewer felt that Morten was explaining the current subject a new question was posed.
The interview ended with a short debriefing as well as planning of further interviews with Morten and his students.
The main focus areas of the interview were:

- The process of mixing
- The purpose of mixing
- Needs
- Target group

Below, you will find the guidelines of the interview listed. These were operationalized, based on the topics above:
**Interview guidelines**

**Briefing**

Who are we?

- Names
- A little about ourselves and our interests
- Our interest and motivation
  - Interest in music and sound
  - Specifically we are interested in the area of mixing
  - A lot of people seem to have problems in this area
  - Mixing are frequently referred to as "magic" or "mysterious"
- Briefly about the research area
  - Interaction and music technology
  - Mixing
  - Prototype development

Who is he?

- His work/profession
- His knowledge within the area
- Which kind of genre is he used to?

**Interview**

The purpose of mixing

- What do you see as common objectives of mixing?
- What do you see as a good mix?
- What do you try to fulfill in the process of mixing?
- What are you trying to achieve?
- What kind of goals do you have?

What are his needs?

- Reference the idea for a system he had at the workshop?(referring to a prior meeting with a one of the members from the group)
- Would you like to be able to achieve these objectives by other means?
- Why? What is ”wrong” with the traditional way of doing it?
- How would the new way of doing it improve the process of achieving the objectives?
- Is this common?
- Are these common in the process of achieving these objectives?

Who is the target group?

- Who have these problems? (Students or professionals?)
- Problems in the process of mixing
- Tools?

**Debriefing**

How we would like to structure further cooperation

- Thank you
- Our work flow - Iterative process
- Additional interviews with Morten and his students

# Appendix B

# Procedure for the workshop at RMC

# Workshop at the Rytmiske Konservatorie
**24-04-12 : 13:30-15:30**

**The workshop will be conducted with 5 students studying sound technology at the ”Rytmiske Konservatie” in Copenhagen. We will record the workshop with video cameras and document the process by taking pictures.**

## Introduction to the workshop:
(10 minutes)

**Who we are**?
*We are Medialogy students currently writing our bachelor project about new ways of mixing music in a studio environment*
**What is this workshop for?**
This workshop serves the purpose of rethinking the mixing process as well as the tools used in the process.
**Workshop structure**
The workshop is divided in X sessions. One session of 15 minutes, followed by two (or three) session of 40, 10 and 20 minutes respectively. There will be a break between sessions at one point.

## Brainstorm session:
(15 minutes)

This session serves the purpose of thinking out of the box about the mixing process. We try this by not thinking about whether our ideas are technologically possible. The outcome will be a number of new ideas about how to mix music.

**Guidelines:**
We ask an initial question that inspires "science-fictional" discussion about future systems for mixing. We then direct the discussion so they do not focus on whether their ideas are implementable.
**Initial question**
*Imagine you are far in the future. The tools and systems you typically use in a music production phase has changed and evolved. Imagine that you have any kind of gadgets or props you may need to help you in the mixing process. How do you think the mixing process be like? How would the tools you use be like and how would they be used?*

## Modeling session (version A)
(40 minutes)

**Introduction**
In this session they should try to do low-fidelity prototypes of their ideas. This means that they should try do "build" or model one of the ideas discussed in the last session. We are still imagining us being far in the future, so they should still not focus on the technology, but more on the functionality of the prototypes.

**Guidelines**
We have brought different kinds of props they can use to build it. They should not worry too much

about the aesthetics or whether it looks exactly as they picture it. As long as it sufficiently demonstrates their ideas it is fine.

**Props**

- Pencils
- Paper
- Scissors
- Tangibles
- Maybe more

## Modeling session (version B)

(40 minutes)

**Introduction**

In this section they should try to do low-fidelity prototypes of their ideas. Since we are working with tabletop technology we will ask them to "fit" their idea on to a multi-touch TUI. This means that you control an interface by manipulating physical objects. These objects then represent digital information in the program. In our case with have these round wooden objects. They represent x and y position on the screen as well as an angle. The interface also has full working multi-touch control. This means that you can move, press, scale, rotate objects exactly as you are doing on your iPhone or Android phones. Try to "model" one of the ideas we discussed in the last session so it fits on the TUI. Do not focus on the technology, but more on the functionality of the prototypes.

*(Above demonstrated on the table (not showing our prototype) and it should be refined.*

**Props**

- Pencils
- Paper
- Scissors
- Tangibles

## Break

(10 minutes)

**Presentation of the models**
(10 minutes)

**Introduction**
Here each group demonstrates their model for the other group. This includes an explanation or demonstration of how the system should be used and maybe why it is a good way of doing it.

**Fitting the ideas to a TUI**
(20 minutes)
*(This is only if we choose the A modeling session)*

**Introduction**
This session serves the purpose of fitting the future ideas to a known technology. We have in this case chosen a multi-touch TUI. We firstly them a simple explanation of the interface and how you interact with it. We then discuss how their prototypes could be made to fit such an interface.

# Appendix C

# Initial questionnaire for the workshop at RMC

# Spørgeskema

Workshop – Det Rymiske Konsevertorie 2012

Dette spørgeskema handler om din erfaring med musik produktion og mixing. Hvert spørgsmål kræver kun ét svar. Læs venligst spørgsmålene ordenligt igennem før du svarer. På forhånd tak for din besvarelse.

**Er du mand eller kvinde?** ______________
**Hvor gammel er du?** _____________

**Hvor mange års erfaring har du med musik produktion? (sæt ét kryds)**

| Under 3 års erfaring | Mellem 3-5 års erfaring | Mellem 5-7 års erfaring | Mellem 7-10 års erfaring | Over 10 års erfaring |
|---|---|---|---|---|
| | | | | |

**Hvor mange års erfaring har du med mixing af music? (sæt ét kryds)**

| Under 3 års erfaring | Mellem 3-5 års erfaring | Mellem 5-7 års erfaring | Mellem 7-10 års erfaring | Over 10 års erfaring |
|---|---|---|---|---|
| | | | | |

**Mixer du eller har du mixet professionelt?**
Ja ______________
Nej ___________

***VEND SIDEN →***

*Følgende spørgsmål handler om din viden og interesse inden for musik teknologi. Med musik teknologi mener vi f.eks. Midi teknologien, DAW's såsom Logic, Protools, Ableton Live og Reason, Hardware eller Software (samplere, effects, instrumenter) osv.*

**Hvilke af disse sætninger passer bedst på dig? (sæt ét kryds)**

Jeg har *ingen* viden om musik teknologi___________

Jeg har *lidt* viden om musik teknologi__________

Jeg har *meget* viden om musik teknologi _________

Jeg har *utroligt meget* viden om musik teknologi __________

**Hvilke af disse sætninger passer bedst på dig? (sæt ét kryds)**

Jeg har *ingen* interesse for musik teknologi___________

Jeg har *lidt* interesse for musik teknologi__________

Jeg har *meget* interesse for musik teknologi_________

Jeg har *utroligt meget* interesse for musik teknologi__________

**Hvilke af disse sætninger passer bedst på dig? (sæt ét kryds)**

Jeg har *ingen* viden om *ny eller kommende* musik teknologi ___________

Jeg har *lidt* viden om *ny eller kommende* musik teknologi __________

Jeg har *meget* viden *ny eller kommende* musik teknologi _________

Jeg har *utroligt meget* viden *ny eller kommende* musik teknologi __________

**Hvilke af disse sætninger passer bedst på dig? (sæt ét kryds)**

Jeg har *ingen* interesse for *ny eller kommende* musik teknologi ___________

Jeg har *lidt* interesse for *ny eller kommende* musik teknologi __________

Jeg har *meget* interesse for *ny eller kommende* musik teknologi _________

Jeg har *utroligt meget* interesse for *ny eller kommende* musik teknologi __________

# Appendix D

# Test procedure for the comparison of intuitive use test

The participant will get the following explanation about the mixing process before the test:

*A music mix consist of a number of individual tracks which for instance could be one track with guitar, one with bass and one with vocal. In the case of this test the mix will consist of six tracks which you will be working by solving ten task given by the researcher(me). When an sound technician sits in a music studio and mixes a song one of part of that process is to spatially place the different tracks in the mix. If you imagine a band on a stage the different musician is place differently according to the listener. For instance is the vocalist often right in front often the listener and also closer than the other musicians. The guitarist could then be to the right of the listener and the bass to the left and bit farther away than the vocalist. It is often this which the sound technician tries to simulate. For instance when he/she is placing the track with the vocalist close and in front of the mix it means that the volume is higher compared to the other tracks since it is closer to the listener and that there comes an equal intensity of sound out of both of the listeners(yours) speakers in a stereo setup. The track with the guitarist would be placed to the right and farther away from the listener than the vocalist by making the volume lower than for the vocalist and pan the sound so that there comes a higher intensity of sound out of the right speaker than of the left.*

The participant will then be asked if he/she understand the concept and if not it will the parts that are not understood be explained again until they are. When the participant understands the concept of the process the test will begin.

# Appendix E

# Initial questionnaire for the comparison of intuitive use test

**What is your gender?** _____________ Male/Female
**How old are you?** ____________

**Do you have any experience with *music production*? (Mark with an X)**

| No | Yes |
| --- | --- |
| | |

**If yes what are your experience? :**

**Do you have any experience with *mixing music*? (Mark with an X)**

| No | Yes |
| --- | --- |
| | |

**If yes what are your experience?:**

# Appendix F

# Tasks for the comparison of intuitive use test

1. Place track 4 in the middle of the mix.

2. Place track 1 to the left and far away in the mix.

3 .Place track 6 to the right in the mix and close to the listener.

4. Place track 2, 3 and 5 somewhere in the mix.

5. Move track 4 and 1 closer to each other in the mix.

6. Switch places of track 2 and 5 in the mix.

7. Try to place all the tracks in the mix so they have the same volume.

8. Move track 5 and 6 farther away from each where track 6 ends up being closer to the listener..

9. Try remove track 4 from the mix so it does not have any volume.

10. Place It again in the mix so it is close to track 5 but a little closer to the listener.

# Appendix G

# Questionnaire about the system for the comparison of intuitive use test

## To which extend do you agree with the follow statements:

| Statement | Fully disagree | Mainly disagree | Neutral | Mainly agree | Fully agree |
|---|---|---|---|---|---|
| I could use the system without thinking about it | | | | | |
| I achieved what I was asked to achieve with the system. | | | | | |
| The way the system worked was immediately clear to me. | | | | | |
| I could interact with the system in a way that seemed familiar to me. | | | | | |
| No problems occurred when I used the system | | | | | |
| The system was not complicated to use. | | | | | |
| I was able to fulfill the tasks in the way I had imaged to. | | | | | |
| The system was easy to use from the start | | | | | |
| It was always clear to me what I had to do to use the system. | | | | | |
| The process of using the system went smoothly | | | | | |
| I barely had to concentrate on using the system | | | | | |
| The system helped me to completely fulfill the tasks | | | | | |
| How the system is used was clear to me straight away | | | | | |
| I automatically did the right thing to fulfill the tasks | | | | | |

## Appendix H

# Final Test Results

# General Information

| Information | Participant No. | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| Gender (Male/Female) | M | M | M | M | M | F | F | M | F | M | M | F | M | F | M | M | M | M | M | F | M | M | M |
| Age | 25 | 44 | 26 | 25 | 23 | 25 | 26 | 26 | 26 | 25 | 27 | 23 | 24 | 26 | 24 | 26 | 25 | 24 | 24 | 26 | 28 | 23 | 27 |
| Music Production Experience (Yes/No) | N | M | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N |
| Music Mixing Eperience (Yes/No) | N | M | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N |

# Mixing Console

| Statement | Rating of Participant No.* | | | | | | | | | | | | | | | | | | | | | | | Mean/Subscale |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | |
| 1) I could use the system without thinking about it | 2 | 4 | 5 | 3 | 1 | 2 | 1 | 4 | 3 | 4 | 2 | 3 | 2 | 4 | 4 | 3 | 2 | 3 | 4 | 3 | 2 | 4 | 3 | |
| 2) I achieved what I wanted to achieve with the system. | 3 | 4 | 4 | 5 | 3 | 4 | 2 | 3 | 2 | 5 | 3 | 4 | 2 | 4 | 3 | 4 | 1 | 4 | 2 | 2 | 4 | 3 | 3 | |
| 3) The way the system worked was immediately clear to me. | 4 | 5 | 5 | 4 | 1 | 4 | 1 | 3 | 4 | 4 | 4 | 4 | 4 | 5 | 4 | 4 | 1 | 2 | 4 | 1 | 4 | 4 | 2 | |
| 4) I could interact with the system in a way that seemed familiar to me. | 4 | 5 | 5 | 4 | 3 | 2 | 2 | 4 | 4 | 5 | 3 | 4 | 5 | 5 | 3 | 4 | 2 | 4 | 3 | 3 | 2 | 3 | 4 | |
| 5) No problems occurred when I used the system | 3 | 3 | 5 | 4 | 2 | 4 | 3 | 4 | 4 | 5 | 5 | 4 | 4 | 5 | 5 | 5 | 2 | 4 | 3 | 2 | 4 | 3 | 3 | |
| 6) The system was not complicated to use. | 5 | 5 | 5 | 3 | 3 | 2 | 2 | 4 | 2 | 5 | 4 | 4 | 4 | 5 | 5 | 2 | 2 | 3 | 4 | 2 | 4 | 2 | 4 | |
| 7) I was able to achieve my goals in the way I had imaged to. | 4 | 5 | 4 | 4 | 4 | 4 | 3 | 4 | 3 | 5 | 3 | 4 | 3 | 4 | 3 | 4 | 1 | 4 | 4 | 4 | 2 | 3 | 2 | |
| 8) The system was easy to use from the start | 3 | 3 | 5 | 3 | 4 | 2 | 2 | 2 | 3 | 4 | 5 | 3 | 4 | 4 | 4 | 2 | 1 | 4 | 3 | 2 | 3 | 3 | 3 | |
| 9) It was always clear to me what I had to do to use the system. | 3 | 4 | 5 | 4 | 2 | 3 | 2 | 2 | 2 | 5 | 3 | 4 | 4 | 4 | 3 | 3 | 1 | 3 | 4 | 2 | 3 | 2 | 4 | |
| 10) The process of using the system went smoothly | 4 | 4 | 5 | 4 | 2 | 2 | 3 | 3 | 2 | 4 | 4 | 3 | 3 | 4 | 4 | 3 | 1 | 4 | 4 | 1 | 2 | 3 | 2 | |
| 11) I barely had to concentrate on using the system | 2 | 5 | 2 | 2 | 1 | 2 | 1 | 3 | 1 | 4 | 2 | 3 | 2 | 5 | 4 | 2 | 1 | 3 | 3 | 3 | 2 | 2 | 1 | |
| 12) The system helped me to completely achieve my goals | 4 | 4 | 4 | 4 | 4 | 4 | 3 | 5 | 4 | 5 | 3 | 4 | 3 | 4 | 3 | 2 | 1 | 4 | 4 | 4 | 1 | 4 | 3 | |
| 13) How the system is used was clear to me straight away | 2 | 4 | 5 | 3 | 2 | 3 | 4 | 3 | 4 | 5 | 4 | 3 | 5 | 5 | 4 | 3 | 1 | 5 | 5 | 2 | 4 | 2 | 3 | |
| 14)I automatically did the right thing to achieve my goals | 3 | 5 | 3 | 3 | 4 | 2 | 3 | 2 | 2 | 4 | 3 | 3 | 3 | 4 | 3 | 2 | 1 | 3 | 4 | 1 | 4 | 4 | 3 | |
| | | | | | | | | | | | | | | | | | | | | | | | | |
| Mean of subjective mental workload (statement: 1,6,11) | | | 4,00 | 2,67 | 1,67 | 2,00 | 1,33 | 3,67 | 2,00 | 4,33 | 2,67 | 3,33 | 2,67 | 4,67 | 4,33 | 2,33 | 1,67 | 3,00 | 3,67 | 2,67 | 2,67 | 2,67 | 2,67 | **2,89** |
| Mean of perceived achievements of goals (statement: 2,7,12) | | | 4,00 | 4,33 | 3,67 | 4,00 | 2,67 | 4,00 | 3,00 | 5,00 | 3,00 | 4,00 | 2,67 | 4,00 | 3,00 | 3,33 | 1,00 | 4,00 | 3,33 | 3,33 | 2,33 | 3,33 | 2,67 | **3,37** |
| Mean of perceived effort of learning (statement: 3,8,13) | | | 5,00 | 3,33 | 2,33 | 3,00 | 2,33 | 2,67 | 3,67 | 4,33 | 4,33 | 3,33 | 4,33 | 4,67 | 4,00 | 3,00 | 1,00 | 3,67 | 4,00 | 1,67 | 3,67 | 3,00 | 2,67 | **3,33** |
| Mean of familiarity (statement: 4,9,14) | | | 4,33 | 3,67 | 3,00 | 2,33 | 2,33 | 2,67 | 2,67 | 4,67 | 3,00 | 3,67 | 4,00 | 4,33 | 3,00 | 3,00 | 1,33 | 3,33 | 3,67 | 2,00 | 3,00 | 3,00 | 3,67 | **3,17** |
| Mean of perceived error rate (statement: 5,10) | | | 5,00 | 4,00 | 2,00 | 3,00 | 3,00 | 3,50 | 3,00 | 4,50 | 4,50 | 3,50 | 3,50 | 4,50 | 4,50 | 4,00 | 1,50 | 4,00 | 3,50 | 1,50 | 3,00 | 3,00 | 2,50 | **3,40** |
| | | | | | | | | | | | | | | | | | | | | | | | | |
| Mean of subscals | | | **4,47** | **3,60** | **2,53** | **2,87** | **2,33** | **3,30** | **2,87** | **4,57** | **3,50** | **3,57** | **3,43** | **4,43** | **3,77** | **3,13** | **1,30** | **3,60** | **3,63** | **2,23** | **2,93** | **3,00** | **2,83** | |
| | | | | | | | | | | | | | | | | | | | | | | | | |
| Mean of all participants | | | | | | | | | | | | | | | | | | | | | | | **3,233** | |
| Standard Deviation | | | | | | | | | | | | | | | | | | | | | | | **0,784** | |

*) 1=Fully disagree, 2=Mainly disagree, 3=Neutral, 4=Mainly agree, 5=Fully agree

# Prototype

| Statement | Rating of Participant No.* | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Statement | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | Mean/Subscale |
| **1)I could use the system without thinking about it** | 5 | 3 | 4 | 4 | 3 | 4 | 4 | 5 | 4 | 2 | 5 | 5 | 2 | 2 | 5 | 5 | 4 | 4 | 5 | 4 | 4 | 5 | 4 | |
| **2)I achieved what I wanted to achieve with the system.** | 5 | 2 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 5 | 4 | 5 | 4 | 3 | 3 | 5 | 4 | 5 | 3 | 3 | 4 | 3 | 5 | |
| **3)The way the system worked was immediately clear to me.** | 5 | 3 | 4 | 5 | 4 | 5 | 4 | 5 | 4 | 4 | 5 | 5 | 2 | 2 | 4 | 4 | 5 | 5 | 4 | 4 | 3 | 4 | 3 | |
| **4)I could interact with the system in a way that seemed familiar to me.** | 4 | 3 | 4 | 4 | 4 | 4 | 4 | 5 | 4 | 4 | 4 | 5 | 4 | 2 | 4 | 4 | 4 | 4 | 4 | 5 | 4 | 4 | 4 | |
| **5)No problems occurred when I used the system** | 5 | 3 | 2 | 4 | 4 | 3 | 3 | 5 | 5 | 4 | 4 | 4 | 1 | 2 | 4 | 4 | 4 | 3 | 3 | 4 | 5 | 3 | 5 | |
| **6)The system was not complicated to use.** | 5 | 4 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 4 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 4 | 3 | |
| **7)I was able to achieve my goals in the way I had imaged to.** | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 5 | 4 | 5 | 3 | 5 | 3 | 3 | 3 | 5 | 4 | 4 | 4 | 5 | 4 | 5 | 4 | |
| **8)The system was easy to use from the start** | 5 | 4 | 5 | 5 | 4 | 5 | 4 | 5 | 3 | 4 | 5 | 5 | 2 | 4 | 5 | 4 | 5 | 4 | 5 | 4 | 3 | 4 | 5 | |
| **9)It was always clear to me what I had to do to use the system.** | 5 | 4 | 5 | 5 | 4 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 4 | 2 | 5 | 4 | 5 | 4 | 4 | 5 | 3 | 4 | 4 | |
| **10)The process of using the system went smoothly** | 5 | 3 | 3 | 5 | 3 | 5 | 4 | 5 | 5 | 5 | 4 | 5 | 3 | 2 | 3 | 4 | 4 | 5 | 3 | 3 | 4 | 4 | 5 | |
| **11)I barely had to concentrate on using the system** | 4 | 5 | 4 | 4 | 4 | 5 | 5 | 5 | 3 | 3 | 5 | 5 | 2 | 2 | 4 | 4 | 4 | 3 | 5 | 4 | 5 | 3 | 4 | |
| **12)The system helped me to completely achieve my goals** | 4 | 3 | 4 | 4 | 4 | 5 | 5 | 5 | 4 | 5 | 3 | 4 | 3 | 3 | 3 | 5 | 4 | 4 | 4 | 5 | 5 | 5 | 4 | |
| **13)How the system is used was clear to me straight away** | 5 | 3 | 4 | 4 | 4 | 5 | 5 | 5 | 4 | 4 | 5 | 5 | 2 | 2 | 5 | 5 | 4 | 5 | 4 | 4 | 4 | 4 | 5 | |
| **14)I automatically did the right thing to achieve my goals** | 4 | 4 | 2 | 4 | 5 | 4 | 4 | 5 | 4 | 4 | 4 | 5 | 2 | 2 | 5 | 4 | 4 | 3 | 5 | 5 | 5 | 5 | 4 | |
| | | | | | | | | | | | | | | | | | | | | | | | | |
| **Mean of subjective mental workload (statement: 1,6,11)** | | | 4,33 | 4,00 | 4,00 | 4,67 | 4,67 | 5,00 | 4,00 | 3,00 | 5,00 | 5,00 | 2,67 | 2,00 | 4,67 | 4,67 | 4,33 | 4,00 | 5,00 | 4,33 | 4,00 | 4,00 | 3,67 | **4,14** |
| **Mean of perceived achievements of goals (statement: 2,7,12)** | | | 4,00 | 4,00 | 4,00 | 4,33 | 4,67 | 5,00 | 4,33 | 5,00 | 3,33 | 4,67 | 3,33 | 3,00 | 3,00 | 5,00 | 4,00 | 4,33 | 3,67 | 4,33 | 4,33 | 4,33 | 4,33 | **4,14** |
| **Mean of perceived effort of learning (statement: 3,8,13)** | | | 4,33 | 4,67 | 4,00 | 5,00 | 4,33 | 5,00 | 3,67 | 4,00 | 5,00 | 5,00 | 2,00 | 2,67 | 4,67 | 4,33 | 4,67 | 4,67 | 4,33 | 4,00 | 3,33 | 4,00 | 4,33 | **4,19** |
| **Mean of familiarity (statement: 4,9,14)** | | | 3,67 | 4,33 | 4,33 | 4,33 | 4,00 | 5,00 | 4,33 | 4,33 | 4,33 | 5,00 | 3,33 | 2,00 | 4,67 | 4,00 | 4,33 | 3,67 | 4,33 | 5,00 | 4,00 | 4,33 | 4,00 | **4,16** |
| **Mean of perceived error rate (statement: 5,10)** | | | 2,50 | 4,50 | 3,50 | 4,00 | 3,50 | 5,00 | 5,00 | 4,50 | 4,00 | 4,50 | 2,00 | 2,00 | 3,50 | 4,00 | 4,00 | 4,00 | 3,00 | 3,50 | 4,50 | 3,50 | 5,00 | **3,81** |
| | | | | | | | | | | | | | | | | | | | | | | | | |
| **Mean of subscals** | | | **3,77** | **4,30** | **3,97** | **4,47** | **4,23** | **5,00** | **4,27** | **4,17** | **4,33** | **4,83** | **2,67** | **2,33** | **4,10** | **4,40** | **4,27** | **4,13** | **4,07** | **4,23** | **4,03** | **4,03** | **4,27** | |
| | | | | | | | | | | | | | | | | | | | | | | | | |
| **Mean of all participants** | | | | | | | | | | | | | | | | | | | | | | | **4,089** | |
| **Standard Deviation** | | | | | | | | | | | | | | | | | | | | | | | **0,596** | |

**Paired-samples one-tailed t-test:** **0,001195**

*) 1=Fully disagree, 2=Mainly disagree, 3=Neutral, 4=Mainly agree, 5=Fully agree