Stim4Sound: A Diversity Computing device helps to alleviate the double empathy problem

MASTER THESIS BY THU NGUYEN

Supervisor: Dr. JELLE VAN DIJK
Internal expert: Dr. ir. RANDY KLAASSEN
Committee chair: Dr. ir. DENNIS REIDSMA

Interaction Technology, Faculty Electrical Engineering,
Mathematics and Computer Science
University of Twente
The Netherlands
May, 2021

Acknowledgement

This thesis project has not been possible without the guidance of my supervisor, Jelle van Dijk, who has always been motivated and has dedicated his time and effort to guide me weekly in the past eight months. I am grateful to have the opportunity to comprehend his philosophy and perceptions of technology for humans. These beliefs have not only contributed a major part in shaping my vision to develop the Stim4Sound application but also given me a much more profound view of applied technology.

I would also like to express my gratitude to my beloved partner, Viktor, who has always been helping and supporting me whenever I struggled the most during the project, especially with the technical and writing challenges. Furthermore, I would like to thank the participants who voluntarily participated in the workshops with enthusiasm and my friends who gave me great advice for the project.

Finally, I am very grateful to receive the love and care from my family, especially my father, who has been giving me the mental support and wise words throughout the journey.

Abstract

Stimming is a repetitive behavior exhibited by autistic people, characterized best by sudden movements, such as flapping, rocking, kicking, and others. It has been extensively studied in clinical and social science literature, and its functions have only recently begun to be understood. One important role of stimming is stress-relieving. Stimming also eases excessive sensory overload, as experienced by autistic people.

However, stimming is still not well received as a normative activity among the general population. Specifically, it is perceived as abnormal behavior because of the odd movements and mannerisms associated with stimming. This stigmatization negatively affects people with autism because they often feel prejudiced and forced to suppress this essential activity.

While traditional clinical treatments of stimming have approached this behavior by applying suppressive therapies, recent studies have shown that these therapies are not very effective and can even be torturous for people with autism. The modern clinical view is that stimming should be encouraged, and its acceptance promoted in the general population. More concretely, different approaches need to encourage stimming and help neurotypical people understand the meaning of this essential function from the perspective of an autistic person, a concept commonly known as the double empathy problem.

In this thesis, I explore an encouraging approach to stimming, which addresses the double empathy problem by bringing neurotypical people to stim together with autistic people, by mutual sense-making in a sound collaboration activity. More concretely, the concept and system I propose in this thesis, called Stim4Sound, uses modern participatory sense-making theories within the so-called Diversity Computing (DivComp) framework. DivComp is a conceptual roadmap that envisions a type of interactive technology that connects people of diverse backgrounds (for example, neurodiverse backgrounds). This is accommodated by facilitating a new ground of mutual understanding of stimming.

Specifically, the Stim4Sound system aims to provide a safe environment and tools for people to make sounds with objects and music together intuitively, freely, and unobstructedly. The goal of the system is to provide a platform where autistic people can stim together with neurotypical people that eventually helps them establish a new meaning of stimming between them. In this way, autistic people do not feel prejudiced, and that neurotypical people gain a better understanding of stimming. As it turns out, these two goals work very well in tandem because making sounds with objects often involves repetitive and rapid movements, which is a characteristic shared with stimming. Furthermore, music and sound creation is an activity that can be appreciated by multiple groups of people with various diverse backgrounds.

To facilitate the research and design of the system, this work applies the so-called Research through design (RTD) methodology to the concept design. Here, the concept and system are iteratively designed and refined in the span of multiple design cycles. Each cycle, except the first, starts and builds on top of the previous by incorporating newly learned ideas and observations. Workshops with autistic and neurotypical users are held at the end of each cycle to better understand the effectiveness of the current design - prototype and how the concept requirements need to be revised. The evaluation of the prototype is interviewing with the users, whose feedback is then used not only to support the hypotheses put forward by this thesis but also to refine the requirements and features for the next cycle.

The main outcome of this thesis is that an early prototype of Stim4Sound has been provided. Preliminary observations show that the main concept shows great potential in achieving its primary goal of providing a safe space for neurotypical and autistic people to stim together by making sounds and music. However, much more technological innovation may need to make the concept complete. Nevertheless, this prototype provides invaluable insights and future guidelines for the design of the final product. Furthermore, this early design concept has helped address most of the key research questions put forward in the thesis. Finally, the thesis contributes to the DivComp framework, by putting its important theories into practice, and extending the existing models.

List of acronyms

OSC Open Sound Control 59

MIDI Musical Instrument Digital Interface 57

API Application Programming Interface 60

RPi Raspberry Pi 59

ML Machine Learning 70

NN Neural Network 70

SVM Support Vector Machine 71

IMU Inertial Measurement Unit 71

DoF Degrees of Freedom 71

LED Light Emitting Diode 59

HCI Human-Computer Interaction 86

In this thesis, the terms 'autistic people' and 'neurotypical people' will be used to pronounce two different neurotypes. There is no intention to offend the characteristic of each group.

Chapters 1, 2, and 3 are the extended version of the research topics that were graded before the project starts.

Contents

1 Introduction

The social phenomenon known as *stimming* is a self-stimulatory behavior most prominently observed in people with autism, and it has always been a topic of interest for clinicians since Kanner and his colleagues first identified the behavior in 1943. Stimming is typically characterized by odd and repetitive movements, such as flapping, rocking, and clapping, but also other indicators, such as sounds, facial expressions, and gestures (Lilley, 2018). A recent study by Kapp et al. (2019) has shown that typical causes for stimming are sensory overload and stress, and stimming plays an important role in relieving such symptoms.

Traditionally, research and therapeutic practice have focused on the fact that stimming can create uncomfortable social situations. People who suffer from excessive stimming traditionally undergo treatments to suppress stimming behaviors. Among the traditional therapeutic approaches is Applied Behavior Analysis (ABA), introduced by clinical psychologist Ivar Lovaas and his colleagues (1987), which aims to 'cure' and 'normalize' autistic people's behaviors into the normative ones. Furthermore, such suppression approaches have even been enhanced by interactive technologies such as social robots (Porayska-Pomsta et al., 2012; Johnson and Picard, 2017). The suppression happens because such behaviors are typically considered abnormal, and can cause distraction and/or discomfort to the surrounding individuals (Kapp et al., 2019).

However, such suppressing stimming techniques are by some considered to be unethical and have even argued to be a form of torture for autistic people (Bagatell, 2010). Specifically, the consensus is that they tend to prevent people from expressing themselves and relating to the world in a manner that fits their cognitive framework. Furthermore, given the biological function of stimming as a stress reliever, such suppressing approaches also negatively affect the mental health of individuals (Hull et al., 2017). Moreover, technology must not suppress individualistic freedom of expression and personal privacy (Friedman et al., 2013), meaning that it should let people stim, as long as it does not negatively affect other people.

The problem is that stimming is best understood by autistic people themselves, but they can still be prejudiced negatively by NT people (Kapp et al., 2019; Hull et al., 2017). For example, autistic persons can sometimes chew on a rubber wristband or a necklace, because they need to stim.

However, they feel discouraged from doing it in front of other people, because of potential negative responses, such as being given an odd look or stared at. One root cause for the negative prejudice is the lack of mutual understanding between the autistic and NT people, which has been coined as the *double empathy* problem by Milton et al. (2012). Specifically, a double empathy problem occurs when there is a *disjuncture in reciprocity between two differently disposed of social actors which becomes more marked the wider the disjuncture in dispositional perceptions of the lifeworld.* These observations suggest that a way to move forward from the traditional suppressive approaches is to reframe stimming from a negative into a positive, productive element in social interaction that leads to better mutual understanding between autistic and non-autistic people.

This leads to another important view of the topic, that stimming should be accepted and encouraged. Specifically, new approaches try to encourage stimming in a way that a mutual understanding between autistic and non-autistic people can be reached (Kapp et al., 2019). There are already various existing applications of interactive technology that could be seen as potentially alleviating the *disjuncture in reciprocity.* One interesting application of Virtual Reality (VR) is to create sensory-inclusive virtual play spaces, where the sensory needs of autistic people are supported (Boyd, 2019). Video games also provide promising means to connect people from various backgrounds, and interact with each other, regardless of their social deficits and oddities. This is because players engage in an activity of common interest, that is not concerned with other aspects other than adhering to the rules of the game, and working together to accomplish a common objective. As such, they tend to ignore and/or accept their differences, as observed in competitive e-sports. Finally, technology can provide alternative media that helps to educate the public of stimming behaviors, such as videos, interactive art installations, and others.

In response to this situation, the present design-research will explore a new and emerging alternative framework, known as Diversity Computing to address the double empathy problem. The core of the Diversity Computing vision, which will be further explained in section 3.1, is to create technologies that do not assume or reinforce the dominant social norms for interacting but instead encourage an explorative and open form of sensemaking between people of diverse backgrounds (Fletcher-Watson et al., 2018). Specifically, an application of DivComp to address the double empathy problem will be designed and evaluated, and its consequences analyzed. We will take the phenomenon of stimming as a starting point, reframing stimming from what should be suppressed into something that may help in a process of shared meaning-making between autistic and non-autistic people. The expected outcome of the study is a design guideline be constructed, which can provide useful insights to the design of empathetic devices that will help autistic people stim in a more socially accepted and ethical way.

The rest of this chapter is organized as follows: Section 1 presents the problem statement and poses the research questions and the associated philosophy of the framework. Section 2 presents the research proposal and contributions of this work. The outline of the thesis is described in Section 3.

1.1. Problem statement and research question

As earlier mentioned, traditional approaches that alleviate the social effects of stimming behaviors through suppression tend to be unethical, and even harmful to the mental health of autistic people.

The challenge is to share mutual understanding for stimming behaviors in the social context between NT and autistic people, especially in situations where the participating social actors recognize the commonality of Autism Spectrum Disorders (ASDs) because of its association with stimming. At the same time, non-technological approaches that aim to address this problem through encouragement and common understanding can be very difficult to implement in practice. However, because they have different natural cognition structures, it is hard for one to understand the other's perspective on a cognitive level. A possibility is that two different neurotypes join in and create co-meaning of stimming, formulated by the act of making sense together with the support of technology.

Also, current interactive technologies lack a common and coherent framework that specifically targets the double empathy problem. Likewise, there are limited studies of interactive technology on stimming behaviors for autistic people. At present, most interactive technologies either assume a neurotypical user, disregarding autistic experiences, or, alternatively, so-called 'assistive technologies' are designed specifically for autistic users. However, in this case, the aim is often to train users to learn to suppress their autistic behaviors in favor of more socially accepted behavior, modeled on the neurotypical norm (Porayska-Pomsta et al., 2012; Johnson and Picard, 2017).

At present, there is no comparable work that shows concretely how the problems can be alleviated while DivComp has the potential to be applied to encourage stimming and address the double empathy problem. DivComp aims to bring individuals' differences together and embraces such diversities through an interactive sense-making process. Concretely, its premise is to help the diversification of computing. As such, central to the topic of this research is answering the following questions:

- **How to use interactive technology to encourage stimming in a more socially acceptable way?**
- **How can the DivComp roadmap be applied to alleviate the double empathy problem?**

To answer this question and address the aforementioned challenges, the following sub-questions have to be answered first:

- ***What is a situated activity that an NT and an autistic person can make co-meaning by interacting with each other?***
- ***How do an NT and an autistic person make sense and establish co-meaning together in a situated activity?***
- ***How can interactive technology support sense-making process regarding stimming between an NT and an autistic person ?***

1.2. Contribution

The key contributions of this thesis are:

- **A DivComp prototype**
  A design of a system, called *Stim4Sound*, that applies the principles of the DivComp model, and helps an NT and an autistic person to foster a new co-meaning of stimming. Specifically, the system provides a rule-free platform where users are free to explore, create and collaborate sound together. Moreover, this DivComp concept provides useful guidelines for the future design of practical empathetic devices that will help autistic people stim in a socially accepted and ethical way.
- **An extended DivComp model**
  An applied model of the DivComp roadmap. The model has been applied in the participatory sense-making activity between two neurotypes: NT and autistic people to alleviate the double empathy problem in the context of stimming. The applied model not only provides guidelines for designing a DivComp device but also provides a detailed investigation of the current DivComp model.
- **Workshops with NT and autistic people**
  Workshops are conducted to evaluate the DivComp device and demonstrate the benefits of this framework. The workshops are organized in cycles wherein empirical observations are used to refine the DivComp concept and introduce new features using the Research through design.

1.3. Outline of the thesis

The outline of the thesis is as follows:

Chapter 2 exposes reader background materials about stimming behaviors. Specifically, the chapter explores the classical and modern perspectives on stimming and the behavior's biological functions. Finally, the chapter concludes with the dialectic perspectives of stimming. In Chapter 3, I summarize relevant state-of-the-art works. In particular, a more detailed overview of the DivComp roadmap and its philosophical and cognitive theories are discussed. Also, the chapter includes other related works that are used to put the contribution of this thesis into the scope of state-of-the-art approaches. Chapter 4 presents the methodology that is applied to the design and evaluation of the DivComp concept. The chapter starts with a brief introduction of the Research through design approach and how it is applied to this design process. Furthermore, I discuss the organizations and planning of the workshops and their usage in the RTD refinement cycles. Chapter 5, 6, 7 and 8 presents the four iterative cycle of the concept development. Each chapter starts with the DivComp concept requirements which are developed from the earlier cycle, and continues to discuss the process of developing and prototyping the concept. Then the chapter demonstrates how the workshops were executed, and how the data were analyzed. The chapter ends with the intermediate discussion of the results from the workshops (and expert interviews). Chapter 9 comprehensively discusses and evaluate the results used to revise the DivComp concept that eventually support our research question. The chapter further discuss the envioned version of the DivComp prototype. Finally, the conclusion of the thesis can be found in chapter 7.

2

Background

The chapter provides background information and theories about autism, stimming, and its dialectic perspectives. The chapter starts with classical and modern views on stimming. Then, it discusses the characteristics and functions of stimming. Finally, the chapter concludes with a discussion of stimming's dialectic perspectives.

2.1. Autism - classical and modern views

In this section, the notion of "autism" is formally perceived according to classical and modern literature. The section first introduces the definition and prevalence of autism. Then, the current scientific consensus on the causes of autism is also discussed. Finally, the classical and modern views on autism are discussed, and the differences are outlined.

Definition and prevalence

Autism, or formally ASD, is a term that is used to describe people who display abnormalities in social interaction, communication, and behaviors concerning the normal population. ASD is a spectrum including Autistic Disorder, Asperger syndrome, Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS), and Childhood Disintegrative Disorder (CDD) (Lai et al., 2013). Even though there are classified groups in ASD, many autistic people may not fall into any of the categories because each person has a unique set of patterns in behaviors, communication, social skill, etc. Lord et al. (2000) claim that ASD is a heterogeneous condition, meaning that there is no homogeneous profile for two different individuals. Furthermore, comorbidities are often found in autistic people. Recent study has shown that 70% of autistic people have comorbid intellectual disabilities (Matson and Goldin, 2013). It has been observed that the prevalence of autism is increasing yearly. According to a report by the Centers for Disease Control and Prevention (CDC) in the United States, the prevalence of autism was 1 per 150 children in 2000 and has risen to 1 per 54 children in 2016 (CDC, 2020). The increment of autism prevalence is also observed in the Netherlands. In 2014, 1.9% (≈1 per 50) of Dutch children are formally diagnosed with ASD (NVA, 2014). This rate has increased in the following years. In 2019, including self-diagnosing, 3.9% of parents say that their child (4-12 year old) has an ASD (CBS, 2019). Increasing rates of autistic people have also been observed in many other countries. Considering that autistic people became recognized in the society that is mostly occupied by NT, sharing knowledge about autism is important for the community to bypass bias and stigmatization against one another. Specifically, understanding how autistic people make sense of the world per se initiates mutual empathy.

Causes

According to classical research, autism is believed to develop from genetic neurodevelopmental brain disorders. In particular, a combination of causes in the form of genetic faults, brain insult, brain disease, etc, have been shown lead to different developments in the brain structure/functions (Frith, 1991).

Nevertheless, there is still insufficient evidence in modern studies to understand the true cause(s) that lead to the development of autism. Fletcher-Watson et al. (2019) state that there are no clear biological markers to explain the cause of autism. They argue that autism may be caused by genetic mutation of the typical gene traits that might result in different brain regions overgrowth. However, current studies of human genes are not mature enough to give sufficient evidence to demonstrate whether genes affect autistic characteristics.

Classical view

In classical literature, autism is viewed as an impairment. Specifically, due to abnormal mental development, autistic people are considered to have deficits in social skills and communication (Kanner and others, 1943). The proponents of this view believe that the associated behaviors do not belong to the 'norm', and should be repressed with therapy practice such as ABA. For example, Frith (1991) describes autistic traits as "handicaps", and argues that these tend to make the social life of autistic people very difficult. Thus, he encourages additional learning to "normalize" their behavior.

Many such intervention practices have been developed in an attempt to "cure" behaviors associated with autism (Lilley, 2018). Tony Attwood (1997) in particular has written a book for parents and professionals to help them understand children with Asperger's syndrome, and guide them with practical strategies to reduce the abnormalities. Though there is empathy shown from the non-autistic point of view, it is still promoting the classical approach, and view, to mould autistic traits into neurotypical ones.

Modern view

In contrast, the modern view leans towards a direction of self-advocacy that embraces individuality. Autistic people reject the term 'disorder', and claim that these are cultural differences that come from the natural diversity of human kind (Fletcher-Watson and Happé, 2019). In particular, Fletcher-Watson et al. argue that current diagnosis practices are flawed because these try to adjust autistic behaviors based on a predefined set of criteria biased towards NT behavior while neglecting the individual's priorities.

For example, the preconceived notion of increasing social contacts can make the person feel more anxious and unhappy. This example emphasizes that understanding personality and individuality should be the correct way to look at autism. These problems should be identified by autistic people first. Moreover, with regards to ethical concerns, autistic people have the right to choose what is suitable for their needs, including if they want to be changed (Bagatell, 2010). When non-autistic people assume that fitting in gives happiness for autistic people without asking their consent, it is an imposing act. Furthermore, some also perceive the 'right' way to interact in a social context differently compared to how non-autistic people perceive. For example, some autistic people believe their social skill doesn't require conversation, but sharing of energy with each other (through repetitive behaviors). In the present discussion, we restrict the analysis to the population of autistic people that do not have an intellectual impairment.

2.2. Stimming in autism

This section is dedicated to *stimming*, a particular set of routines exercised by autistic people. The section starts with a short description of stimming and the associated features. Next, stimming is contrasted with fidgeting, a similar behavior experienced by non-autistic people. The public perception of stimming is also discussed. Finally, the section is concluded with a detailed discussion about stimming as a self-regulatory tool of the mind to sensory input overload.

Description and features

The term *stimming* refers to a set of self-stimulated behaviors, which are characterized by stereotypical, repetitive motor movements. Such movements typically involve different parts of the body. Typical hand movements for example can be clapping, tapping, waving, flapping, twirling, etc, while leg movements can be kicking, stamping, shaking, and others. Full-body movements such as rocking, jumping, running, and spinning have also been observed (Lilley, 2018). The repetitive behaviors have been observed in different groups in the autistic spectrum Lord et al. (2000). Moreover, research has shown that these behaviors usually remain unchanged over time. Concretely, Ballaban-Gil et al. (1996) conducted a longitudinal examination of 102 people, including autistic adults and adolescents, and found that 50% establish stereotypical behaviors and that 66% of them have their behaviors unchanged after 5 years.

Relation to fidgeting and tics

Although stimming routines are usually attributed to autistic people, similar practices can also be experienced by non-autistic people, where these are often referred to as *fidgeting*. Similar to stimming, fidgeting describes continuous repetitive movements, which are smaller in comparison. Often they are called *tics* if the movements are sudden and short in duration. Specifically, leg shaking, pen clicking/rolling/rocking, nail biting, finger tapping, etc are typical examples of fidgeting and tics that are expressed by non-autistic people.

Studies have shown that fidgeting has an important role. In particular, one such study by Farley et al. (2013) has shown that students fidget to keep focus their attention on the lecture material. Another study suggests that fidgeting is beneficial to people that sit for too long, who fidget with their legs due to the induced endothelial dysfunction found in the leg vasculature (Morishima et al., 2016).

Public perception of stimming

Stimming is considered a controversial topic (see Section 2.3): some perceive it as an abnormal behavior that should be suppressed, while others perceive it as a freedom of expression that should be allowed. Although stimming patterns for non-autistic people come in the form of fidgeting, these behaviors seem to be more acceptable than those of autistic people. One probable reason is that fidgets are more subtle and less attentive. However, intuition also suggests that such behaviors have been normalized among non-autistic people because they can be easier to empathize with and relate to.

In contrast, stimming from autistic people is still negatively prejudged by the general population that mostly comprises of NT people. A probable explanation for this is that movements stemming from stimming seem strange to non-autistic people because these movements are preconceived as more unusual and less subtle practice. Thus, it is difficult for non-autistic people to relate to that might result in a preconceived bias against autistic people. Understanding the underlying cause of stimming from the biological point of view may alleviate this bias.

Stimming as a coping mechanism for sensory overload

Vaguely speaking, stimming can be considered as an instinctive self-defense mechanism to protect the mind from getting overburdened by sensory information. More formally, recent studies by Kapp et al. (2019) suggest that stimming functions as a self-regulatory mechanism of the sensory input to cope with and to soothe certain mental states. Typical stimuli that induce such states are overwhelming environments, *noisy* thoughts, and excessive emotions, which incidentally are the underlying influence factors that can cause stress. Concretely, an environment that is unpredictable, confusing, and overwhelming can lead to the sensory overload that contributes to such states. As an example, the study showcases a scenario where a person has urges to stim after interacting with other autistic people, because of the sensory overload caused by the overwhelming surrounding people.

Besides stressful states, states of high energy and emotion are also the target of stimming as a regulator. Such high states of energy and emotion can also be induced by sensory input overload. Autistic people in particular experience hyper-or hypoactivity from the sensory input (Lilley, 2018).

The consequence is that uncontainable emotion coming from any emotion which is positive or negative prompts them to express their feelings through body sensory output (Kapp et al., 2019). Self-regulating with a repetitive motion calms these hyperarousal states. This interpretation is similar to (Lilley, 2018; Miller et al., 2007).

Also, stimming could be attributed to the compensation of sensation shortage, for which hypoactivity is responsible. Sensation seeking occurs when the person lacks a particular sensation, or needs to increase their arousal level (Miller et al., 2007). Unsatisfied sensation sends signals to the body to generate more sensory input, which could be from sound, body movement, visuals, etc. Sometimes the craving happens when the perception of sensation is reduced and thus stimming might be an explanation for such needs (Lilley, 2018).

Sensory overload is observed to derive from Sensory Over-Responsivity (SOR) and sensory sensitivity, as observed by Miller et al. (2007), who also state that *"People with SOR respond to sensation faster, with more intensity, or for a longer duration than those with typical sensory responsivity"*. Specifically, according to the authors, the sympathetic nervous system, which is responsible for putting the body on alert by sending rapid involuntary response to dangerous or stressful situations, is activated when such SORs occur. Similar observations are also shared by Attwood (1997), who asserts that autistic people may have increased sensitivities to sounds, visuals, smells, taste and texture of food, temperature and/or pain. In the case of sounds: sudden, unexpected, high-pitched and continuous noises, and/or confusing, complex or multiple sounds can be perceived as extremely intense auditory stimuli. Regarding visuals, others can be sensitive to illumination, colours, and/or distortion of visual perception. The coping mechanism is either to avoid such disturbance, or to respond with distress behaviors. Thus, these behaviors are seen as reflexes to overwhelming stimuli. Such abnormal responses are automatic, unconscious biological reactions (Miller et al., 2007). These similar effects are also found in stimming: people describe them as automatic and uncontrollable (Kapp et al., 2019).

2.3. Dialectic perspectives of stimming: Suppression or encouragement

Stimming has been a controversial topic for therapeutic clinic and self-advocate autistic people. Classical view of practitioners suggest that stimming should be suppressed as it opposes the 'norm' while modern view and advocate autistic people dispute that view and defend stimming.

Suppressing stimming

Stimming has been socially suppressed. Autistic adults have been often told not to stim. Often to avoid marginalization or being judged, people hide their stims by masking them with more socially accepted behaviors or to stim when they are alone. According to Hull et al. (2017), these 'camouflaging' behaviors are in favor of helping them to appear socially capable, to adapt and connect with the others. The main reason is *negative prejudice* against stimming. For example, people who stim are viewed as strange, aggressive, and childish. They often put a lot effort in masking to avoid discrimination and negative feedback from the others. Furthermore, behaviour therapy practices are established to suppress stimming, especially for autistic people at the young age. For example. Applied Behavior Analysis (ABA) and *'quite hands, quite mouth, quite feet'* are two common techniques (Lilley, 2018). ABA use reward and punishment system to eliminate motor stereotypies and encourage 'normal' behaviors, the *quite* technique uses cue cards to remind them about their behaviors.

However, the consequence and effect of suppressing stims is detrimental for the mental health of the autistic people. Some feel anger, anxious and belittle; other experience stressful mood and feel insecure. The effort of hiding or masking their stims is costly and exhaustive;in several cases they need time to recover. Stimming has been shown to have important role for the body, suppressing such behaviors means obstructing their biological responses.

Encouraging stimming

Autistic self-advocates and neurodiverse people argue that stimming should be encouraged. There is no case of them saying that stimming should be eliminated. On the other hand, destigmatization occurs when there are mutual understanding and acceptance.

Stimming should be encouraged regardless of social norms that mostly formed by NT people. Social norm does not permit the 'unfit' behaviors and is a bias toward the NT people. The purpose of these suppressing techniques is to fit the 'abnormal ones' into the 'normal ones'. They try to mask, to mold the natural autistic behaviors into the ones that appear acceptable in society. Nolan et al. (2015) explain that social knowledge serves not only as a heuristic model for normative practices but also a heuristic filter for responses to stimuli. It makes us define what is 'normal' and what is 'not normal'. It is apparent that the goal of social norms is to create a respectful environment that people live in, but these norms are formed by the majority of the population, the NT people.

Stimming also represents autistic identity, culture, and sensory language. Autistic self-advocates claim it is a form of embodied semiosis that comes from a different culture, but not a deficiency. Thus, interventions that attempt to correct their sensory expression are rejected. The person who engages in the act of camouflage perceiving himself not only deceives oneself's authenticity but also derecognizes the community's identity (Hull et al., 2017). Nolan et al. claims, *'many autistic self-advocates argue that autism is a uniquely embodied and sensory language that defies NT logic and comprehension' (...) and 'stimming, emerge from within, the 'development' of shared and social knowledge serves to silence rather than enhance their language and phenomenology'*. For NT people, the sensory information is socially trained and thus results in the 'norm' behaviors. Such information is the learned knowledge, but stimming is a natural unpolished raw sensory information that then results in the 'abnormal' behaviors. However, this sensory information is not and can't be trained because of its inherent nature.

Stimming is free will behavior that the person has the right to express. As long as it doesn't affect or hurt other surroundings, people should be able to do what they like. In such an opposite case, it should only be repressed because autistic people are unconscious about their stim hurting others, even though they have no intent for harm. Having their behavior be 'fixed' may deprive their choices to behave. Many autistic people perceive interventions such as ABA without consent as unethical because they should agree first on what should be changed or what not (Bagatell, 2010). Similarly, eliminating stimming is unethical to the medical principle of 'do not harm' (Kapp et al., 2019).

Allowing and accepting stim gives benefits for the mental health (Hull et al., 2017; Kapp et al., 2019). Mutual understanding of stimming boosts productivity and relieves worries for autistic people. More importantly, autistic people can be themselves and have better social experiences.

Summary

To summarize, the literature provides a detailed background of opposing viewpoints of autism and stimming. Specifically, in contrast to the classical view where autism is seen as an impairment, modern literature perceives autism as a different culture and supports self-advocacy and individuality. These opposite perspectives of autism strongly influence the perception of stimming, which is an associated behavior commonly observed in autistic people. At present, some clinicians consider the behavior as 'abnormal' because it does not fit in the 'normal' behaviors. On the other hand, other and self-advocate claim that stimming is part of autism culture, identity, and sensory language and is a free will behavior that other people can not deprive them of.

However, regardless of all these viewpoints, we need to take a step back and truly understand the purpose and function of stimming. Stimming is a natural and biological form of self-regulation of a mind and body via the body's sensation when either of them is at an unease state. The behaviors are reported to be automatic, unconscious, and uncontrollable. The purpose of it is to generate a feedback loop to balance one's emotional level, and thus serves as a soothing and comforting mechanism for the mental state. As such, stimming is a natural body reaction that can not be changed or fixed because it comes from a human instinct, and people should be able to stim as long as they do not bother others.

The observed underlying problem here is not about stimming, but it is about the social pressure and rules that define what the 'right' behavior is, and immediately stimming is the demonstration for the discussion. By providing knowledge, context, or even experience of stimming to a larger audience, I believe that people can empathize with stimming, and autism in general, better. Then, the pressure can be relieved and the negative judgment may become less.

In the next chapter, I will discuss the state-of-the-art literature. The related interactive technology that supports stimming will be shown. Furthermore, the details of DivComp philosophy and related theories will be discussed.

3

Related work

The chapter discusses relevant state-of-the-art works. Specifically, a more detailed philosophy and cognitive theories of DivComp and design principles for participatory sense-making are discussed. Then, the chapter continues to discuss relevant interactive technologies that support participatory sense-making process and stimming to put the contribution of this thesis into the scope of state-of-the-art approaches. The chapter finally demonstrates findings that support comfort, communication, and interaction for autistic and NT people that can be applied to the DivComp design concept. The discussed related work in this chapter has been gathered in an iterative process. Some of them are collected during the design cycles. However, this chapter includes and discusses all the related works in this thesis.

3.1. DivComp philosophy, design principles for participatory sense-making and for comfort

DivComp philosophy

DivComp, which was originally proposed by Fletcher et al. (2018), is a conceptual roadmap that envisions a type of interactive technology, which connects people of diverse backgrounds (for example, neuro-diverse backgrounds). This is accommodated by building a common ground of mutual understanding, within which they can be supported in making sense and creating meaning together - *co-meaning*, while at the same time not establishing each other's normative frameworks and perceptual worldviews as the sole basis for the exchange of ideas (which each person must then adhere to).

As outlined by Fletcher et al., DivComp aims to emphasize the role of diversity computing device by providing a platform for people to perform a *participatory sense-making process*, where individuals from different backgrounds can create shared meaning and new perspective. Concretely, the idea is to help people embrace their own differences, while still being engaged and connected with others. The aim is not to seek a common goal, but rather to make sense of the present situation.

Thus, DivComp technology should be a mediator to help people constantly reflect upon themselves to create co-meaning together. Philosophical and cognitive theory, participatory methodology, and digital innovation are the main ingredients to create a DivComp application.

DivComp consists of philosophical and cognitive theories, participatory methodology, and digital innovation. The knowledge of these theories helps us to understand the underlying characteristics of how two people create shared meaning. The participatory method provides them space for self-reflection and reflecting, which are parts of making co-meaning. The digital medium supplies scaffolds to mediate the interaction. The DivComp framework is currently a rather preliminary vision that combines theories and empirical insights from multiple ranges of disciplines such as psychology, philosophy, design, and computer science. At this moment, there is a need to have a concrete application built on the concepts of DivComp to show that it is possible to create a DivComp device and to further explore what is concretely needed to design Diversity Computing artifacts.

Design principles for participatory sense-making

Participatory sense-making is the fundamental concept of the DivComp philosophy. Hence, to design an interactive DivComp device that supports sense-making between NT and autistic people, it is important to understand both perspectives of how autistic and NT people make sense of the world. The reason for this is that autistic people have trouble understanding the actions and intentions of others, and the same goes for the way NT people misunderstand the actions and intentions of autistic people. As such, the principles of how autistic and NT people make sense are served as the ground theories, and fundamental features to help us design a DivComp device that can tailor both ways of how these two neurotypes make sense and understand each other.

The *Participatory sense-making* notion, was introduced by De Jaegher and Di Paolo (2007), describes how people in close embodied interaction with one another *'make sense together'*. This collaborative, active process of sense-making is not a matter of exchanging pre-existing understandings from one mind to the other mind - as more traditional cognitive science would have it - but rather, sense-making is seen as a thoroughly situated and embodied activity: people make sense of the situation at hand through the process of interacting with one another. This interaction comprises much more than verbal language: the entire body, its appearance, and actions are at all times in an ongoing expression of the process of sense-making, and people make sense by continuously responding to such expressive activity of the other's body. Sense-making is therefore at the core a self-organizing, situated process that emerges out of the interactions between two people, not something that goes on in any individual mind.

The general principles of how people make sense together in an embodied way are described by Hummels et al. (2015). They propose that social-situatedness, scaffold, trace, interactive imaginary, dialogical system, first-person perspective, and catalyzing engagement are the important factors to create a shared experience that helps people understand, and making sense easier in the process. The social-situatedness factor helps to create the ground, the rules that the people can take part in. *Scaffolds* are seen as guidances and instructions to direct them through the desired actions and tasks. Furthermore, *traces* are physical markers from their actions that can help them establish the cause-effect connection. These factors can be implemented in the design concept to make sense-making between two individuals become easier. In addition, an intuitive way to invite people to participate in an activity is to provide them affordance, which can be achieved by supporting *perceived couplings* between digital representation and physical object (Hornecker and Buur, 2006).

Besides, autistic people are sense makers like any other, yet their ways of interacting with the world, and the specific behaviors and bodily expressions that go along with this process of sense-making, may be very different from that of neurotypical people. The differences in how autistic people make sense in a participatory interaction are that they have an asynchronous style of interaction, pay little attention to social cues (e.g eyes contact, human face, sound, movement), have rigid posture and movements (De Jaegher, 2013).Furthermore, they also have poor skills for mutual coordination, attention shifting, and imitate less of the others. Because they have different patterns of making sense of the world compared to NT people, there might be a clash of misunderstandings between autistic and NT people.

Design principles for comfort

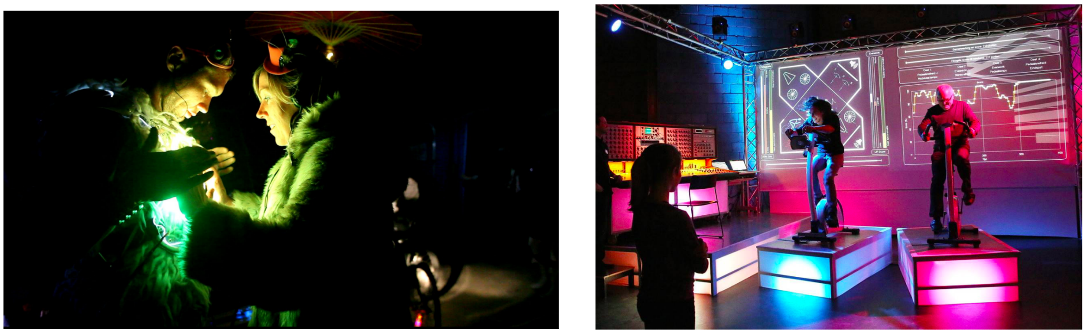
Being comfortable allows people to behave their own ways in the social context, which is one of the key goals of the DivComp philosophy. Hence, a DivComp device needs to facilitate the comfort between users to keep them ongoing to participate in a new activity, which is the *inviting* factor of DivComp.

Technology can offer comfort by providing users' personalized settings and preferences. An example to demonstrate this point is the virtual inclusive-play spaces application proposed by Boy et al. (2019). The application offers controlled parameters and customized virtual environments. These options give them the freedom to manipulate their sensory settings and thus create a comfortable interaction for them.

For autistic people, the comfort also arises from an *asynchronous* or a *non-verbal* form of communication that the technologies can equip (Zolyomi et al., 2017). For example, Zolyomi et al. find that collaborative content sharing and affecting technologies can provide social comfort for autistic users. The reason is that these platforms allow autistic people to read *fewer non-verbal social cues to interpret.*

Regarding the comfortable sound, a pleasant sound is determined by its frequency, volume, and pitch. A study that compares unpleasant and pleasant sounds shows that lower frequency sounds are pleasant while higher frequency sounds are unpleasant sound Nisha and Soruparani (2018). Moreover, loud, sharp, high-pitch, and noisy sounds lead to unpleasant feelings (Özcan and Schifferstein, 2014). On the other hand, sounds with positive associations are rated with higher pleasantness (Carles et al., 1999).

3.2. Interactive technologies



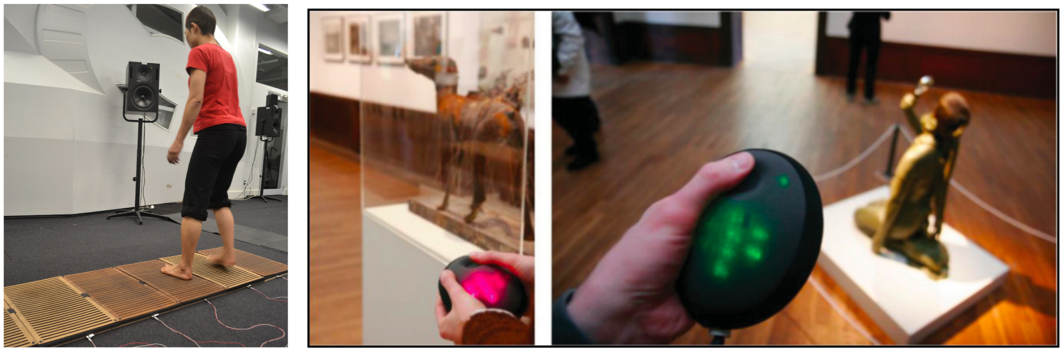
(a) (b)

Figure 3.1: a: Interaction between Performer and person (Hobye and Löwgren, 2011); b: Impression of SoundBikes (Maes et al., 2018)

Support sense-making process

Sense-making can be formulated by the way interactive technologies map certain sensed signals into a predefined set of meanings, or *rules*. For example, Hobye et al. (2011) developed a Mediated Bodysuit (Figure 3.1a) to let users engage with each other through touching-listening experience. Based on the number and duration of touch, the suit converts touch input into complex sound and light output. The coding scheme of the system leaves lots of room for the users to explore and interpret the relationships between touch, sound, and light. The authors indicate that technology changes the nature of touch-sound interaction, especially that the suit is seen as an instrument. With *exploration* and *collaboration*, the users can together engage and make sense of how the technology works in this social play context. Tune-Me-In is a sound-induced antenna that allows a person to move his hand up and down to navigate through a maze game with pitch (Stienstra, 2003). The high pitch corresponds to the 'up' and the low one corresponds to the 'down' direction. The system requires to have two sound-induced antennas operated by two users, as one is for the 'left' and another is for the 'right' direction, to control the character's navigating direction. The users need to coordinate their hand movements together to navigate in the game. The game also transforms the meaning of hand movement to the direction - rules of movement. While Tune-Me-In's sensed signal acts as a controller, Mediated Bodysuit's sensed signal - touch acts as a musical instrument. Both applications provide a platform where users can participate, explore and try to figure out how the system works. However, they also provide the meanings of the users' actions based on their interactive behaviors. In contrast to this goal, Stim4Sound accentuates the sensed signal of (bodily) movements and brings them to the foreground of consciousness. Users listen to the soundscape where the sounds of their repetitive behaviors are accentuated. The role of technology here is not to impose a certain meaning of sensed signal but to present it explicitly for users.

Self-reflection is a deeper level of consciousness that aids one to construct his perception and meaning. Technology can promote self-reflection by presenting one's sensorimotor movements and behaviors explicitly. Similar goal to Stim4Sound, the system called Slow Flow designed by Feltham et al. (2013) help people aware and reflect on their act of walking due to the responses from



(a) (b)

Figure 3.2: a: A person walking on the Slow Floor (Feltham et al., 2013); b: Scenario of visitors engaging with the Lega in the art hall (Mentis et al., 2014)

pressure-sensitive sound-generating surface (Figure 3.2a). When a system has *interactive coupling* integrated, one becomes more aware of their own body due to the consistent paired feedback. The example shows self-reflection of dancing patterns achieved by foot movements and sound mapping. On the other hand, Stim4Sound emphasizes the act of stimming and helps people reflect on their repetitive movements by the transparent produced sound. Also, reflection can be acquired explicitly by recorded self-expression. Another sharing experience device called Lega (Figure 3.2b) to record one's self-expression through tactile interaction. Lega tags his expressions on the current location and re-expresses them with another person in nonconcurrent moments (Mentis et al., 2014). Even though people experience asynchronously, it has shown that one's self-expression can be perceived and added by others. With the similar goal of sharing moments, Stim4Sound allows people to share their expression due to being in the same soundscape. The difference is that Stim4Sound delivers a physical form of memory. The whole recorded audio and taken photographs are given as take-home messages to people, which eventually foster their sense-making process.

Technology also provides a platform for an unconventional form of expression and communication such as *non-verbal communication* and *body language communication*. For example, Maes et al. (2018) developed an embodied, social, and collaborative sound application called SoundBikes (Figure 3.1b), which people cycle to control playback parameters of pre-composed songs. According to the authors, the technology allows a new form of musical expression to emerge: a combination of physical exertion and collaborative interaction. They show that participatory sense-making can be established with dynamical interaction and *co-creation* processes. Specifically, *joint synchronizations* in tempo, phase, and balance from people' collaboration are evidence for establishing the sense-making process. Stim4Sound also provides a platform where users can communicate and understand one another in an atypical form: through sounds and body language.

Manage stimming behaviors

An approach is proposed by Boyd et al. (2019), which also focuses on encouragement of stimming. Concretely, the approach is applied in the context of encouraging neurodiverse children to behave with their sensory difference, and thereby encouraging stimming. The specific design proposed by Boyd et al. is an interactive, sensory-inclusive application for NT, autistic and ADHD children, and it aims to facilitate comfortable and engaging virtual spaces for its users. It provides a customized environment with a preference for each type of sensory profile, where children can freely express their behaviors, and regulate their sensory input. The application has been shown to provide comfort by embracing neurodiverse differences, including stimming.

There are some applications and devices made that encourage people to perform self-regulating actions, such as stimming. In particular, Cottrell et al. (2018) develop a soft-bodied fidget toy to promotes self-regulation. BioFidget toy which encourages people to fidget has been shown to reduce stress level (Liang et al., 2018). Similar results of reducing stress have been seen from the study of (Alonso et al., 2008) when the users fidget with the pen with rock, roll, and squeeze motions. These self-regulation has been shown to help people relax. Specifically, Karlesky and Isbister (2016) show that the purpose of self-regulation is to help people calm, focus, and be creative. However, these approaches tend to support more 'accepted' self-regulatory behaviors. For example, the soft-bodied fidget toy only considers more socially accepted behaviors such as touching, squeezing, and padding. On the other hand, more atypical behaviors, such as flapping and clapping, are not included. Furthermore, fidget toys have been shown to negatively affect the user's attention. Specifically, Graziano et al. (2020) investigated how fidget spinners influence sixty children with Attention Deficit Hyperactivity Disorder (ADHD) in. He found that these toys tend to reduce the children's attention to activities presented in a classroom. Similar findings under similar conditions have been reported in the study of Byrne (2019), who includes six autistic children in the study, in order to understand the influence of fidgeting toys. There he shows that such toys discourage children's engagement, and academic comprehension in class. To summarize, both studies conclude that fidgeting toys significantly reduce focus and create more than necessary distractions for the users.

Other more impulsive systems that do not yet encourage atypical stimming behaviors are ECHOES (Porayska-Pomsta et al., 2012) and SPRING (Johnson and Picard, 2017), which intend to suppress stimming behaviors by applying ABA intervention in their design. Specifically, Porayska-Pomsta et al. apply the SCERTS framework in their study to guide a child to imitate NT behaviors. Thus, these approaches contribute to the idea of encouraging socially accepted self-regulatory behaviors but disregarding stimming behaviors.

3.3. Summary

The DivComp philosophy and design principles for participatory sense-making provide important guidelines for the design of the Stim4Sound system. As such, the theories provided by the DivComp framework are directly applied in the design of the physical DivComp device, and are tested with real users in the span of multiple iterations. Similarly, the literature findings for comfortable technology and sound are used as the ground up concepts to derive comfort design choices for the Stim4Sound system, especially for some autistic users that have sensitivity to sound.

Furthermore, the discussed related work about interactive technologies that support the sense-making process is used not only as a source of inspiration for building the DivComp device, but also to make direct comparisons and identify the contributions of the Stim4Sound system as a participatory sense-making platform. There a several key takeaways from this related work that helped to facilitate this process: *rules*, *exploration*,*collaboration*, *self-reflection*, *interactive coupling*, *non-verbal communication*, *body language communication*, *co-creation*, and *joint synchronization*.

However, what I have learned from the state-of-the-art literature, is that current applications that provide participatory sense-making processes have not yet targeted the interaction between an NT and an autistic. Specifically, most of the applications assume that the potential users are only NT. Though current interactive technologies can accommodate both autistic and NT users' preferred styles of communication, interaction, and comfort, they have yet to provide a more explicit way to facilitate the interaction between these two neurotypes. Furthermore, findings on how autistic people perform participatory sense-making have been found, but have not been applied in a practical interactive application.

Nevertheless, such interactive technologies provide comfort, communication, and interaction preferences of NT and autistic people, and can be applied to mediate the interaction between these two neurotypes. These are important factors that should be considered in the research process to design a DivComp device that alleviates misunderstandings in communication and interaction between a NT and an autistic user. Especially, when encountering strangers, users need to feel comfortable with one another in order to reduce their fear of being judged. Furthermore, Stim4Sound, which is a sound-making platform that users can collaborate with, needs to provide pleasant and comfortable sounds for them, resulting in a positive sensory experience.

With respect to managing behaviors stimming, the state-of-the-art technologies discussed so far in general lack an explicit method that encourages and supports atypical stimming. Besides the approach from Boy et al. which focuses on encouraging stimming, other approaches partially fall in the category of encouraging atypical stimming behaviors. Specifically, they encourage people to perform self-regulating behaviors such as fidgeting, squeezing, padding with hand motions that are more socially accepted. Furthermore, some toys have negative effects on users' attention to their activities. However, there are more impulsive systems that do not yet encourage atypical stimming by other means.

Developing concept goal

Based on the previous discussion of the literature study, an initial system goal is put forward: to encourage an autistic person to stim while informing them of how their stim affects other surrounding people. Also, the system should assist an NT person to understand the normality of stimming as an expression from the autistic person's point of view. This will be a complex challenge, especially given that people of different neurotypes have different ways of making sense of the world. It is a complicated and open challenge, because of how two such different forms of sense-making may ever converge on a shared understanding of each other in a situation.

The promise that is embedded in our design goal, is that the system would let both of the user neuro-types experience stimming, and help them make sense together so that a new shared perspective can be established. As such, it is important that the system does not impose certain meanings and interpretations of stimming behaviors, but allow them to introduce their own meanings of stiming themselves. The system shall only provide a shared platform where people can experience stimming. For example, the role of the system can be a mediator to help the users share the stimming sensation. Thus, in order to allow them to experience stimming together, there should be a certain context where they can immerse in a *shared activity.* The same activity provides a reciprocal and joint activity to help them establish sense-making (De Jaegher, 2013).

The activity can introduce a shared goal where both parties would have to solve a problem together, or let them collaborate to create something meaningful. This derives from the related work, which shows that collaborative interaction enhances coordination and engagement between peers. Collaboration is one of many participatory activities, which is a proposed approach from DivComp to foster meaning-making for peers. Thus, our goal for the DivComp concept is:

DivComp concept goal: to find a paired collaboration activity that can foster sense-making between autistic and NT people

4

Methodology

This chapter discusses the primary methodology of Research through design (RTD) that I apply in this work as proposed in (Frayling, 1993). According to the author, RTD is an approach that applies design methods to the research process. Deriving from Frayling's approach, in this thesis I applied the RTD approach by designing physical artifacts to prove and evaluate the DivComp concepts with NT and autistic people. Specifically, my approach consists of multiple prototyping-evaluating iterations in order to investigate how interactive technology can support participatory sense-making activity between an autistic and NT person in the stimming context.

4.1. The applied RTD approach

The approach consists of an iterative revision of collecting literature findings, designing prototypes, gathering users' evaluation, and reflecting. In an iteration, the literature findings provide the

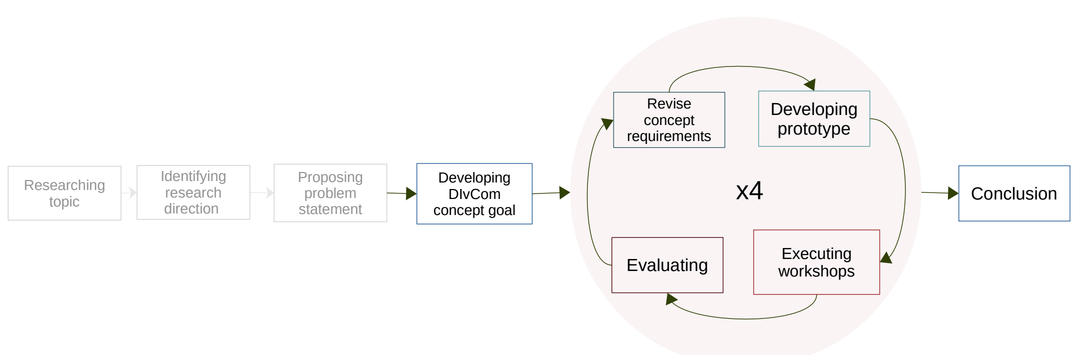


Figure 4.1: Overview method structure

ground premises to build the DivComp concept requirements. These requirements are constructed into a physical demonstration by different design techniques. Specifically, I take inspiration from existing technologies, use my creativity and design skills to successfully demonstrate the concept requirements into a concrete physical representation - a Proof of Concept (PoC). This PoC consists of a proposed activity and instruments to help users develop a participatory sense-making process. It is then evaluated in an interactive workshop, where the activity happens between two users (either between 2 NT users or an NT and an autistic user) and a post-session interview is conducted. The observation and answers from the workshop are gathered. These are the results that need to be analyzed, reflected, and conceptualized into higher abstract realizations. These realizations are compared with the intermediate concept requirements. Combined with literature findings, a new set of revised concept requirements is proposed that is used as input for the next iteration.

The iterative process of prototyping-evaluating helps me to evaluate the specific concepts, which are derived from the DivComp's abstract concepts of participatory sense-making between the different neurodiverse groups, in the real-world scenario.

The act of using design methods can refine such abstract concepts to more concrete ones. Specifically, the benefit of doing design in research is the ability to transfer abstract knowledge into a specific situation set in a situated context (Stappers and Giaccardi, 2017). Furthermore, as DivComp is a general roadmap that consists of philosophies from multidisciplinary areas, the framework still needs further exploration to provide more detailed and specific concepts of how participatory sense-making can be established. The advantage of using design methods to do research is for knowledge generation (Stappers and Giaccardi, 2017). The knowledge of how to establish and foster a participatory sense-making process between an autistic and an NT person produced in this approach can help me further explore, investigate, and extend the current DivComp framework.

The structure of the approach is as followed: To iteratively revise the DivComp concept requirements, I conducted 8 prototyping iterations, which are divided into 4 main prototyping cycles (Figure 4.1). The first aims to explore different idea prototypes to formulate the first concept requirements. The second introduces a low-fi prototype of the concept. The third cycle presents the development of the early prototype. In addition to user evaluation, three expert interviews were conducted in the third cycle. Finally, the last cycle shows the latest improvement of the prototype. In each cycle, there are four stages: concept requirements, concept development, and prototyping, workshop execution, and results, and intermediate evaluation and discussion (Figure 4.2).

The divergent-convergent design technique is applied in the concept development and prototyping stage. Divergent thinking takes open and non-criticized ideas and design features. It involves exploration and experimentation to search for the concept ideas. These ideas are converged and selected in each design cycle. They are used to make the design of the DivComp prototype concrete in order to support an informed evaluation with experts and users.

4.2. Participants

In total, five NT and one autistic people participated in the workshops. There were two NT males, three NT females, and an autistic male. The autistic person was diagnosed with Asperger Syndrome,

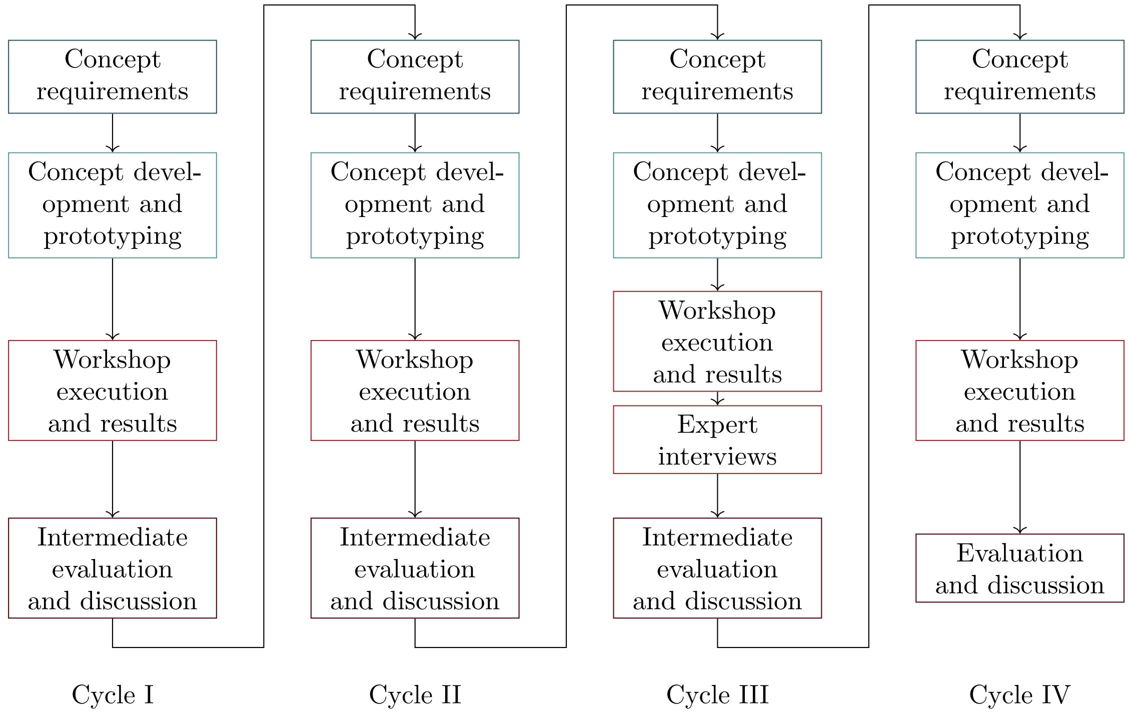


Figure 4.2: Iterative stages in four main cycles

confirmed by expert clinical opinion. All participants were young adults (22 - 30 year-olds). All participation was voluntary without reward, except for the last evaluation, an autistic and a NT participant were offered small rewards.

4.3. Experts

On a weekly basis, my main supervisor provided me insightful feedback on my progress and guided me through the process. In cycle III, 5 experts were interviewed. These experts are philosophers and psychologists in the field of autism and participatory sense-making theories. Specifically, one philosopher involves in the field of autism and the double empathy problem. Another expert who has developed the theory of participatory sense-making is a philosopher of mind and of cognitive science. A professor of developmental psychology that specializes in autism and another professor who works closely in participatory methods for designing technology for people with disabilities. The last expert is a researcher in the field of cognitive neuroscience of autism. Notably, the last four experts were the authors behind the DivComp framework.

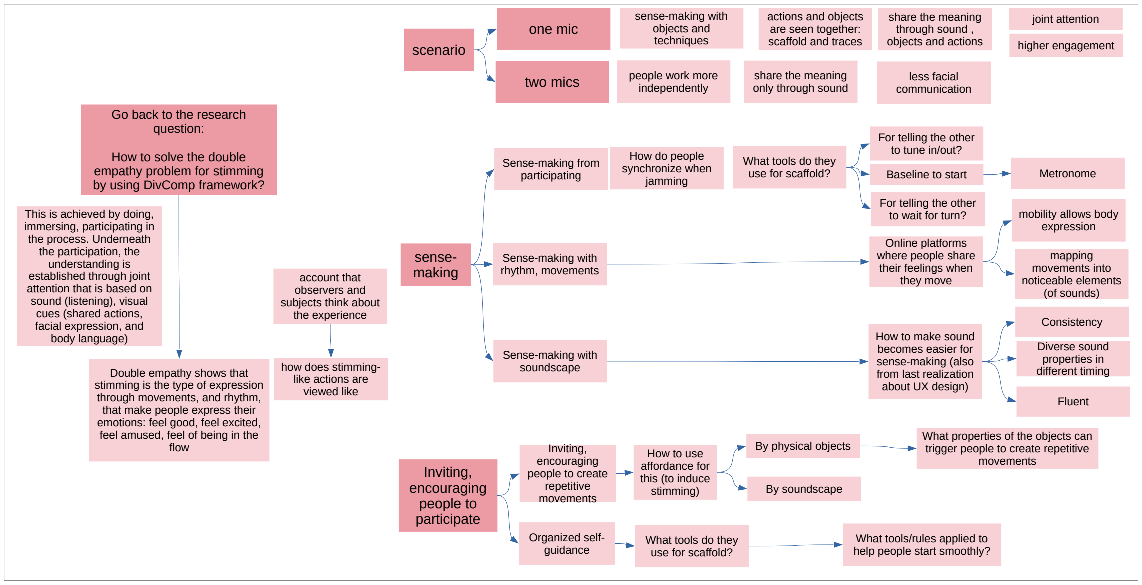


Figure 4.3: A snapshot of the third reflection session

4.4. Concept requirements

To construct the concept requirements, I hypothesized the concept requirements based on the literature review and workshop results found in the previous cycle. Literature related to concepts, design principles, and interactive technology that support participatory sense-making in autistic and NT were used as the main guidelines to design the DivComp prototype. The results from the workshops were empirical evidence to help me compare with the assumed concept requirements and provide me new knowledge of how people engage and make sense together. Based on these results, the abstract reflection points were derived. These points were presented together with literature findings in an intuitive way with a visual diagram called Miro [1] so that they could be easily compared and analyzed with the current concept requirements. For example, a snapshot of a third reflection session after the third workshop is shown in Figure 4.3.

4.5. Concept development and prototyping

In each cycle, I developed a new set of concept requirements and the DivComp prototype based on these requirements. To find a concept that satisfies the requirements, I sought inspiration and conducted an ideation process for the concept development. Before building the prototype, I researched the technical feasibility and ran multiple experiments with the software and hardware.

[1] https://miro.com/

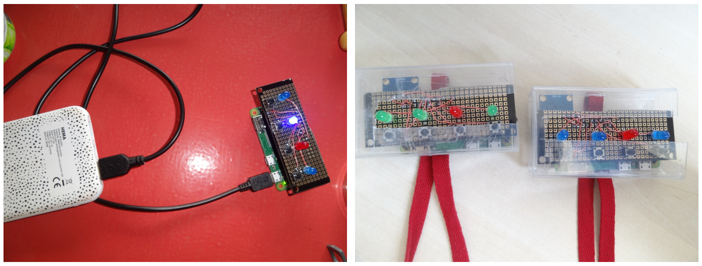


Figure 4.4: Left to right: A prototype is developed in the third and fourth cycle

Inspirational content seeking

To develop and revise the concept requirements, I used related work and digital opportunities as the sources of inspiration. Specifically, related works of concept, design principles, and technology support participatory sense-making activity. When searching for inspiration, I looked at different related content on visual platforms such as Pinterest, Youtube, Tumblr, etc. In addition, a more familiar and physical source of inspiration was ordinary objects around my home. Especially in cycle I, when searching for the shared activity that encourages stimming, I was interacting and experimenting with different tangible objects used body motors. This allows me to explore ideas that trigger repetitive behaviors.

Ideation

From the inspiration and findings input, I conducted multiple brainstorming sessions to develop the concept. The ideation process consisted of individual and group brainstorming sessions. For group sessions, only two group brainstorming sessions were conducted with an NT user in the first cycle. The goal of both group and individual brainstorming sessions was to generate as many ideas as possible and let ideas be generated unobstructedly. This method will result in a higher quality for ideas(Paulus et al., 2015). These ideas are then compared and chosen based on my skills and expertise.

Prototyping

Then, the next step was to build the prototype that demonstrated the concept requirements. Throughout four cycles, the iterative implementation results in prototyping a low-fi to a digital device. Specifically, from the first to the fourth cycle, four developed prototypes were: a low-fi prototype (with sticky notes, wizard-of-oz set up, human mediator/controller, and ordinary objects), a low-fi prototype (with primitive User interface functions), an early prototype, and a refined wearable prototype with gesture recognition (images shown in Figure 4.4).

Before designing the prototype, I conducted online research on technical tools. Especially the research focused on the feasibility and difficulty to build the physical prototype. Based on the online

Figure 4.5: Photos of two workshops during the process

survey I carried out different experiments with implementation platform interfaces such as Makey Makey tool kit, Arduino, smartwatch, and raspberry pi. Furthermore, I had to conduct multiple experiments with different software systems to find and establish the prototype's system architecture. These were experiments with different control-flow algorithms, software-hardware combinations, open-source platforms, the communication protocol between different software, etc. From here, a control-flow, state machine, and class diagram were formulated. These schemes serve as the system blueprints to guide me design the prototype.

4.6. Workshop execution

Procedure

Before carrying out the experiments, I received approval to conduct these workshops from the ethics committee (Appendix A.1). The workshops were conducted indoors, in a shared and quiet space. Before the experiment started, the participants gave their consent to be photographed, re-quoted, and have their full names used (consent form is shown in Appendix A.2). Then, they were briefed that space was rule-free for self-expression, about the collaborating sound concept and the ordinary objects provided for inspiration, and how the prototype worked. Then, the experiment between a NT and NT person or a NT and an autistic person started. After the experiment, the participants took part in a post-session paired interview. Each session usually lasted from 1.5 to 2 hours. More detailed information about the workshop procedure conducted in each cycle is discussed in the section workshop execution and results in each design cycle chapter. Photos of two workshops are shown in Figure 4.5.

User interviews

After the experiment, an interview was conducted to get more insightful information about the participants' opinions. One of the goals of the interviews was to understand what features were needed in the activity in order to help them make sense and establish mutual connections together. Regarding sense-making, they were asked about how they communicate, what components help them to communicate, when they were on the same page of understanding, and whether they understand

Questions
'What is your general opinion about the experience?'
'Without verbal communication, how did you guys communicate with each other? '
'Did you understand what he was trying to do at that moment?', 'and why did you interpret that?'
'Was there a point you don't understand or see what is going on?'
'How did you response to him at that moment?'
'Do you find this feature X helpful?'
'Do you understand what this feature X means when it Y?

Table 4.1: Examples of questions for post-session interview

one another's actions, etc. For example, I pinpointed certain actions that happened during their interaction and asked them why they performed such actions. Furthermore, related to mutual connection, I asked them questions about how much reciprocal engagement and understanding that they establish together in the activity. Moreover, another goal was to understand how and what design features can coherently facilitate their communication and interaction. Besides, the interview also provided their emotions, feelings, and expression that they were experiencing during the interaction. These were also important points to evaluate participants' level of engagement ad comfort. Some example questions from the interview are shown in Table 4.1 (the full list of 7 sessions is shown in B). Text notes were collected in all interviews. Additionally, in the last evaluation, a behavior observation note of participants' interaction was made.

4.7. Expert interviews

Two individual and a group of three expert interviews through online video calls were conducted in cycle III. All of the interviews were mostly unstructured. The first one was with an expert who is the philosopher in autism and the double empathy problem. The second one was an individual interview with the professor of developmental psychology that is specialized in autism. The goal was to ask experts about the unseen breakdown factors that could happen in an interaction between NT and autistic people. Furthermore, experts also pinpoint the missing gaps and provide suggestions for the intermediate concept requirements of DivComp device. Having these experts from different fields contribute their opinions in the developing process brings an interdisciplinary perspective for me to significantly improve the concept requirements and the prototype in the last design cycle.

Before an interview started, I showed them a pitch video of three minutes to present the information about the project more intuitively and concisely. The video [2] first describes the problem of stimming in the social context and how the Diversity Computing framework can be applied to alleviate this issue. The video continues to provide my proposal to the intermediate developed DivComp prototype. The information of how this prototype works is explained next. After the pitch video, the experts give their feedback on the intermediate DivComp prototype.

[2] https://youtu.be/bxLyNZ5ZJGo

Questions
'What is your general point of view about the concept?'
'What could be the possible interaction breakdown between two different neurotypes? '
'What is an important concern for this participatory interaction between 2 different neurotypes?'
'Are there any concerns I should take into account?'
'How can we encourage sense-making during the interaction?'

Table 4.2: Examples of questions for expert interviews

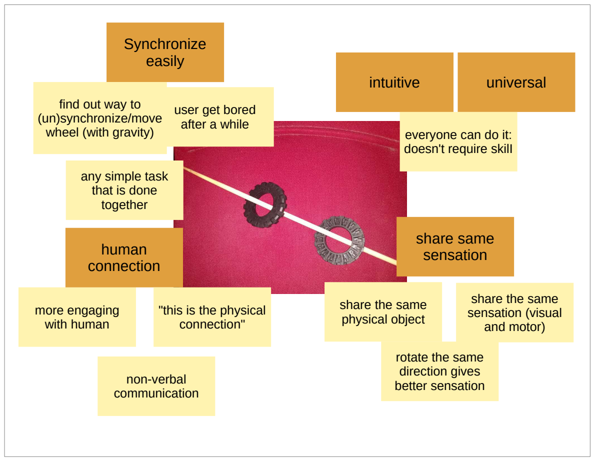


Figure 4.6: Theme map analyzed from results of a workshop

To gather their opinion, I took notes of their feedback on the first and second interviews. In the third interview, I asked and received their consent to record the online meeting. Later, I also took notes from the transcribed audio. The prepared questions that were asked during the interviews are listed in Table 4.2. Regarding the length of each interview, the first one took about forty-five minutes, the second one lasted around fifteen minutes, and the last session took about one hour.

4.8. Results analysis

To analyze the data, I used affinity diagram (Figure 4.6 to group observation notes, users' and experts' evaluation data into more similar concepts. These concepts are categorized into similar theme maps, which provide prominent topics that are emerged from the qualitative results (Courage et al., 2015). These maps help me to validate the intermediate concept requirements and design decisions and to identify the missing requirements. At the same time, additional literature research findings were used as reference sources to help me reflect and revise the concept requirements for the next iterative process.

4.9. Summary

To summarize, four main iterative cycles, which consist of 8 sub-cycles, were conducted to help me answer the research questions. Following the rest of the thesis, the next four chapters will describe four iterative cycles which are demonstrated in Figure 4.2. Each chapter will first discuss the concept requirements that are revised from the findings in the previous cycle. The chapter proceeds with the concept development and prototyping section that describes how a prototype is built based on the concept requirements. Next, it explains in detail the workshops' procedures and results. The last section in the chapter provides an in-depth intermediate discussion on the workshop results and concept requirements. The next chapter will describe the information in the first cycle.

5

Cycle I

5.1. Concept requirements

DivComp concept requirement: a paired collaboration activity can foster sense-making between autistic and NT people

The first requirement for designing a DivComp device was to find an activity that falls within the DivComp scope from the proposed DivComp concept goal, as described in the related work. The shared activity should have the potential to establish a *mutual connection* and *mutual understanding* between an autistic and an NT person. At the same time, to let both of them experience stimming, the activity should have the affordance to encourage them to stim freely. Thus, the requirement of a DivComp activity was that it should:

- establish a mutual connection between an autistic and an NT person
- establish mutual understanding between an autistic and an NT person
- encourage people to stim freely

5.2. Concept development and prototyping

The next step was to find such an activity that satisfies these requirements. To select a participatory situated activity, it was easier to first explore different activities within the same shared space. Three individual brainstorming and two group brainstorming sessions with an NT participant were conducted. Furthermore, related work of participatory sense-making in autistic people, interactive technologies that support sense-making, and digital inspirational content were used as references in the ideation process. In total, there were 13 prototype ideas proposed, and compared. They are summarized in Table 5.1. Among these ideas, 6 of them were evaluated in 2 workshops with a NT participant I would refer to as M. The results from three iterative subcycles were reflected and used to revise and formulate different activities. Six key criteria, as described below, were used to evaluate, and grade the proposed activities.

The first criterion is the amount of repetition in movements. As explained in Chapter 2, stimming is typically characterized by repetitive movements and behaviors. Based on this criterion, an activity is rated by considering the repetition rate of movements performed in the activity. Nonrepetitive movements are rated with a (- -), and highly repetitive with a (++). For example, an activity that involves shaking or tapping is considered to be highly repetitive.

The second criterion is the amount of time required to establish a connection between users. This criterion is important because when they spend more time together, they understand each other better, which will be further explained in the results section. The amount of time is quantified based on the number of actions performed in an activity, and the duration of each action. For example, an activity that requires 1-2 actions with each lasting less than a couple of seconds, e.g. lighting a candle, is rated with a (- -). On the other hand, long activity with many actions, each taking more than a minute, is rated with a (++).

The third criterion is the type of activity, which has been considered based on the related works of interactive technologies that support sense-making. The two types considered are *task-based*, and *story-building*. Task-based is an organized activity that the users are asked to perform well-organized tasks, often based on a predefined set of rules. In contrast, a story-building activity is one such that the users are left to explore, find their own creative flow and collaborate together. The observations as described later indicate that story-building activity induces higher mutual engagement compared to task-based activity.

The amount of sensation induced by the activity is the fourth criterion used in the idea selection process. This is because stimming is related to the function of self-regulating sensory input. Thus, the amount of sensation induced by activity can help an NT user experience similar sensation experiences, as an autistic user does, which is a crucial component to the final DivComp prototype.

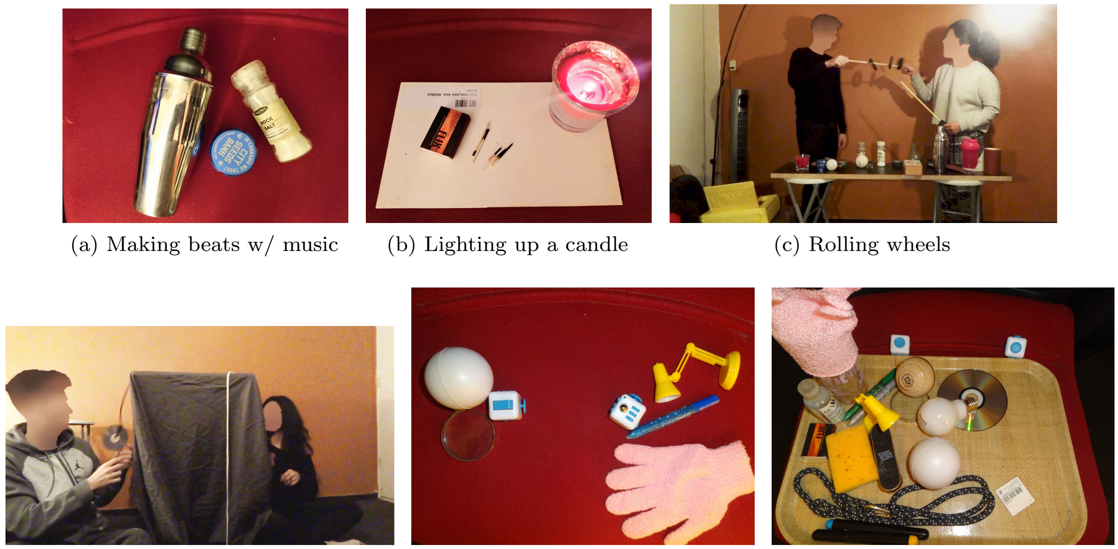
The last two criteria used are the originality of the activity and its difficulty to implement. The implementation difficulty is determined by my skills and expertise to develop within a given time frame.

The described criteria for all the activities were rated based on observations from the workshops, and self-rated.

	Idea #	Activity	amount of repetitive behavior	amount of time to establish connection	task-based, story-building	inducing sensation	originality	difficulty for implementing
Sub cycle I-a	1	Each person holds one side of a rope. They need to move in circle while rotating the rope to generate digital visual pattern come from the rope	-+	++	Task-based	-+	-+	very hard
	2	Each person has a spring-like accordion to paint a drawing together. The contracted-extracted signal from the accordion maps to the brush position	-+	--	Story-building	++	--	hard
	3	Each person has a spring-like accordion to play music. The contracted-extracted signal from the accordion maps to midi notes	++	++	Task-based	-+	--	hard
	4	Two people need to complete a map together by adding the provided objects which some of them has sensation	--	++	Story-building	-+	-+	medium
	5	Each person holds one side of a rotating rope with wheels attached. They need to keep rotating to get the wheels connected to make some fire sparks	++	--	Task-based	++	++	medium
	6	A ‘go go goat’ similar game where Each person stands in a set jumping position lets Each person controls a jumping direction	++	++	Task-based	--	--	medium
Sub cycle I-b	7	Each person has an interactive fidget where the one starting first is a controller and another is an actuator (and vice versa)	++	--	Task-based	++	--	hard
	8	Two people communicate with Each other in 2 different rooms through provided objects. Different objects have different communicative style	--	++	Task-based	--	++	very hard
	9	Each person holds one side of a elastic rope that replicates noodle. Both people need to find the right twist and shake motion to create noodle in a game	++	++	Task-based	++	++	very hard
	10	Two people need to complete a map together by searching the activated provided ‘stimming-like’ objects and adding them to the map. There are Two similar sets of item provide for 2 people. When one person interacts with a item, the other can feel the sensation from his similar item	++	++	Story-building	++	++	very hard
	11	Two people need to build a digital house together by with a stimming controller. The controller can detect different repetitive behaviors and map them with specific actions	++	++	Story-building	-+	--	medium
	12	A game where Two people are placed in a dark room. Both of them have to stim to activate the light of the room. They need to find coding scheme of the standing position to stim	++	++	Task-based	-+	++	very hard
Sub cycle I-c	13	A game where both people make music together by making sounds (beats, rhythm, melody) with repetitive behaviors and movements	++	++	Story-building	-+	++	medium

--: low -+: medium ++ : high

Table 5.1: Criteria compared table of 13 concept ideas



(a) Making beats w/ music (b) Lighting up a candle (c) Rolling wheels

(d) Interacting through a blind curtain (e) Exchanging sensory experiences (f) constructing a map from objects

Figure 5.1: Six activities from two workshops

5.3. Workshop execution and results

Execution

I conducted two workshops with M spanning two sub-cycles. The first one was to experiment with different ideas on the fly (bodystorming technique). The second one was to evaluate three ideas that were proposed in the first brainstorming session. The main goal of these workshops was to find a specific activity that can establish a mutual connection between participants. Before the workshop started, M was told to freely explore, and improvise with surrounding objects so that we could freely make activities that potentially create mutual connection. Each session lasted from 2 to 2 and a half hours because M often enjoyed the workshop. In two sessions, we evaluated six activities, of which three were created on the spot (Figure 6.3).

After each workshop, I conducted a post-session semi-structured interview to get M's opinions on the activities. The interview lasted about half an hour. Specifically, M was asked how and why he enjoyed the activities, how he felt connected and understood our actions when interacting with me, and how he engaged with the activity itself. As a participant myself, I also included my opinions after interviewing M to reduce bias.

The answers from the two interviews were categorized into themes. They associate with:

- each activity's features that are related to connectivity (in the sense of DivComp)
- the type of activity

- the activity's prerequisites
- the factors that influence the participants' interest
- the feelings that participants experience

The categories are shown in Figure 5.2 and Figure 5.3.

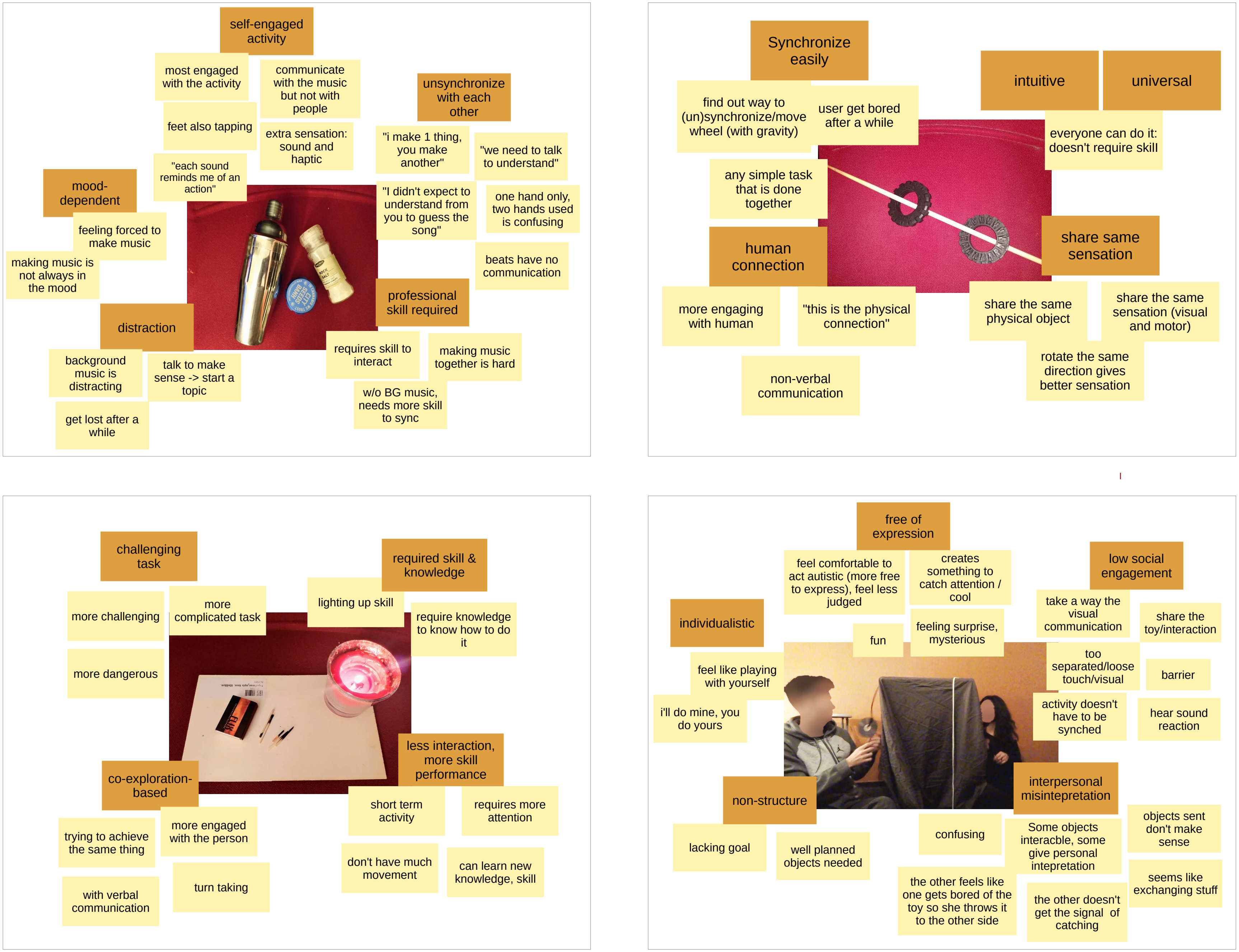


Figure 5.2: Results of categorized themes from four activities. Top left: Making beats with music, top right: Rolling wheels, bottom left: Lighting up a candle, bottom right: Interacting through a blind curtain

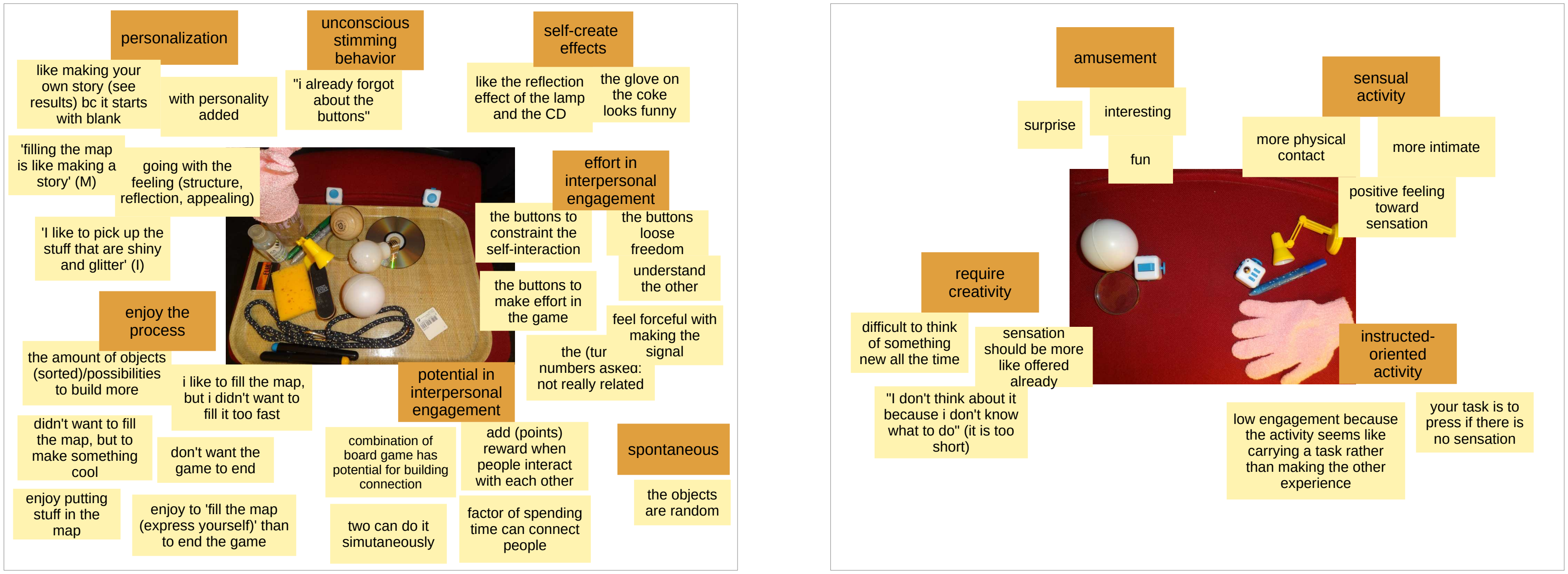


Figure 5.3: Results of categorized themes from two activities. Left: Constructing a map from objects, right: Exchanging sensory experiences

Results

The outcomes from each workshop are described as follows: For each session, the setups, environments, and activities are discussed first, followed by a summary of the key observations resulting from these activities.

In the first session, we conducted three simple activities: Making beats with music, lighting up a candle, and rolling wheels. Through these, it was observed that we could engage with each other when solving a mutual task, by sharing the same object, exploring, experiencing the same sensations, and synchronizing to the rhythm together. Mutually performing a task allows us to put effort to achieve the same goal, and sharing the same object enables physical connections with movements and sensation.

In another body-storming session, we spun wheels attached on a stick together. The main observation is that both of us felt the spinning momentum and the friction from each other's hands. Additionally, when the stick was spun the wheels were also spinning along in the same direction. This effect created a certain visual sensation where we seemed to enjoy. Thus, such repetitive physical movements help to produce a sensational experience. On the other hand, certain factors such as mood, skill, and synchronized coordination strongly determine engagement. For example, making rhythm with objects can be difficult as it requires some musical knowledge to synchronize (jamming), and sometimes people are not in the mood to make music.

It is also important that the concept takes the perspective of autistic people's interaction patterns into account, and how they make sense in an activity. Thus, findings from how autistic people make sense of the world (De Jaegher, 2013) were applied in the next workshop. In the next session, I explored different types of interactions that favor autistic people's sense-making traits, such as human interactions that require less facial communication, attention, timing, and coordination. Furthermore, I also explored asynchronous interactions in space and time that are more appreciated by autistic people (Zolyomi et al., 2017).

In the following session, I evaluated three activity ideas created from the previous sub-cyle's brainstorming session. These activities are: exchanging sensory experiences, constructing a map from objects, and interacting through a blind curtain.

I have learned that such creative, exploring collaborative activities result in higher engagement because usually people enjoy being in their creative flow, than doing tasks and solving problems. For example, the 'constructing a map from random objects' activity allowed M to put his ideas on the table. The process of thinking and putting effort into creating something on his own encouraged him to be in a *flow*, where there was a committed and uninterrupted engagement. Furthermore, time is also another factor to help us connect: the longer the activity, the more time we spend with each other.

An activity that obstructs visual contact, for example in the blind curtain activity, results in an asynchronous interaction. This type of interaction provided us more space to freely express ourselves and still ensured interpersonal engagement. However, the drawback was that both of us experience difficulty in understanding each other due to mutual assumptions. This happens because visual cues were needed for facial and body communication but they were not provided. In particular, during the ‘blind curtain’ interaction, we did not understand each other’s intended action. One perceived a thrown-over object as a used toy, while the other perceived it as an attentive communication tool.

Summarized results from workshops:

- Mutual engagement can be established by mutually solving a task, co-exploration, co-building, and sharing the same experience
- Repetitive motions can induce the sensation
- Engagement influencing factors are skill, knowledge, effort, mood, time, and coordination.
- Being in a creative process is more engaging than solving a goal process
- Turn-taking interactions discounts ‘in-synch’ moments
- A non-face-to-face interaction causes low mutual interpretation but offers more room for self-expression

5.4. Intermediate evaluation and discussion

A co-building type of activity is more engaging than a problem-solving task activity because being in the making process can be more enjoyable for people. Specifically, an activity involving creativity can provide a sense of flow to let them freely explore and improvise. Additionally, such co-building activity should be intuitive and easy for people to get engaged. Thus, it should require minimum skill and knowledge to make the activity inviting and easy for anyone to participate. However, the given tasks should not be too easy, otherwise, they will get bored and lose interest in the activity.

Furthermore, the activity should focus on stimming, such as repetitive behaviors and sensations, as it is the central theme of the design. However, the evaluated activities did not present stimming behaviors as the foreground tasks yet. As such, further explorations to find an activity that aims to bring stimming behaviors into focus should be done.

There are trade-offs among the level of engagement, freedom of expression, and synchronization between an asynchronous and asynchronous interaction.

From observation, an asynchronous-view sharing activity such as blind curtain interaction providing fewer social cues results in a lower level of mutual engagement and understanding. This happens because social cues are necessary for visual communication. Specifically, when an NT and an autistic person share the same view they have access to visual cues from eye contact, facial expression, and body language. Thus, such sharing view activity allows one to read other’s expressions and observe his actions and understand him better. However, the trade-off is that an autistic person may not prefer reading social cues. Also, he may not express himself freely because the other is observing him.

Moreover, asynchronous-time interaction for making movements such as exchanging sensory experiences does not pressure the autistic person to synchronize with the other’s bodily coordination. While autistic people may appreciate this type of interaction more, it has shown a lower level of interpersonal engagement compared to a more synchronous interaction. As such, to adapt both ways autistic and NT people make sense whilst keeping them engage and understand each other, it is important to find suitable asynchronous and synchronous features for the interaction. Thus, further exploration to find a suitable activity was needed.

Furthermore, the activity that involves more rules helps an NT and autistic user synchronize better. However, the amount of rules needs to be considered. For example, when there were too many rules, the mutual engagement between participants became less natural. Furthermore, it may lean toward goal-oriented or task-oriented activities, rather than exploration-based activities. On the other hand, lesser rules activity give them space for self-expression.

Moreover, mutual sense-making has not been observed yet. This indicates that more research on fostering sense-making is needed. Also, the activity can provide features that keep them interested in participating. Thus, In the next cycle, the concept needs to focus on feature requirements that enhance meaning-making and connection.

6

Cycle II

6.1. Concept requirements

The concept requirements for the activity have been revised based on the results and evaluation from workshops conducted in Cycle I. The new requirements entail that the activities should:

- *focus on story-building collaboration*
- *offer sensation*
- *encourage and prioritize repetitive movements*
- *require minimal knowledge and skill to perform*

6.2. Concept development and prototyping

Sound collaboration activity

Among all the proposed ideas of shared activities in the previous cycle, ***Sound collaboration*** was selected as the most suitable activity for the DivComp concept, because it best satisfied the old, and new concept requirements.

Regarding the first requirement, music, and sounds often capture the mood and events that transpire during a recording session. Hence, one may think of the constructed sound piece (a track) as a story being built through collaboration. The traces of story building, reflection, and collaboration are the sounds that have been recorded and layered on top of the next recording turn. Users can use these traces to get inspired and create their stories through sound.

Also, listening to and making music offers users sensational experiences, which is exactly what the second requirement entails. This is rather an auditory sensation that both people are creating and experiencing. Concretely, while sharing the same soundscape, they are in the flow of the creative process of producing sounds and music.

Regarding the third requirement, it is observed that people often record repetitive sound samples when they participate in the activity. This is because music is usually represented by repetitive rhythmic beats. Specifically, the beat is a fundamental element that attaches a certain structure to the final composition and makes it sound more harmonious. Also, letting people listen to the playback loop encourages them to create repetitive patterns that synchronize with the loop.

Finally, as users are let to explore freely in an open space with no specific rules, they don't need prerequisite skills or knowledge to make sound together. Still, the collaborating experience depends on the sound/music expertise of each person.

Hence, the final chosen concept of 'Sound collaboration' enables team collaboration, time spending, and repetitive (bodily) movements, and requires less facial communication and coordination. Notably, the last feature is appreciated by autistic people.

The primary focus of this sound collaboration concept is to establish an interaction between an autistic and NT person, where they come together to make sounds and music. At this point, the concept is a general conceptual idea that needs to be concretely demonstrated and evaluated in practice. Hence, this concept development cycle further explores what explicit design features are needed to physically demonstrate the requirements. Moreover, as mentioned in the intermediate discussion from the previous cycle, the sense-making process has not been observed yet. Thus, further explorations on design features that foster sense-making are explored.

In this development cycle, a more concrete design concept for the sound collaboration activity was developed. Concretely, a low-fi prototype of the concept was demonstrated with a preliminary wizard-of-oz setup. This setup allowed me to rapidly evaluate the prototype with the users. The following high-level specification of the prototype was used to guide the development towards a complete product:

Sound collaboration allows an autistic and NT person to share the same soundscape and record real-time sound feedback. A sound collaboration system shall provide a safe and comfortable space where both users are free to express themselves. They can improvise and produce different sounds however they want. There are no rules on how the sounds are created, however, there are constraints on the functionality of the system itself: When the sound is recorded, the system shall record sounds from both microphones where each person has one. There will be multiple recording iterations of the sounds that each user wants to produce. Thus, the recorded sounds will be layered on top of each other. After each recording round is over, both of them will listen to their recorded sounds that play in a loop to hear what they have produced.

In order to adhere to this specification, the system needs to fulfill a set of primary functions listed in Table 6.1.

Functions
Record sound from live microphone
Play recorded sounds
Provide multiple recording turns
Receive record signal from user
Give signal to users when the system is about to record
Give signal to users when the system is done recording

Table 6.1: Table of required basic functions

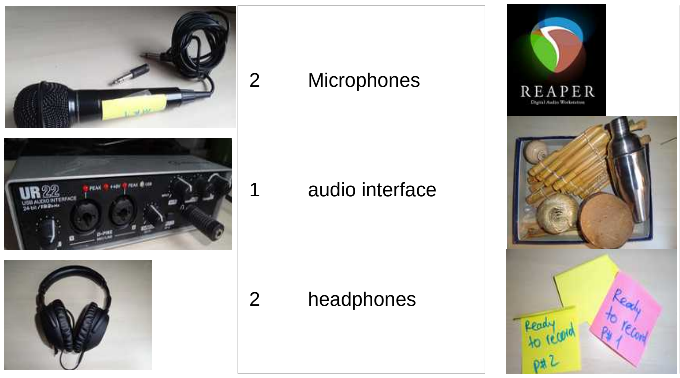
Low-fi set up

The setup of the low-fi prototype consists of a mediator - a person who performs the mentioned required functions for the users, an audio software (which is commonly referred to as a Digital Analogue Workstation by professional practitioners), an audio interface, two headphones, two microphones, physical objects, and sticky notes (Figure 6.2). The materials, hardware, and software are shown in Figure 6.1. In order to allow two users to produce sound and listen to the same soundscape, the system has to let both users hear each other's microphone live via the audio software. Also, to let them record and play sound, the system has to provide recording and playback commands. Ordinary physical objects are provided as tools and inspirational sources for users to create different sounds.

To conduct the activity, the mediator needs to perform all the required functions mentioned in Table 6.1 to the users. Specifically, by interacting with the audio software the mediator records the sound when they are ready to record, plays the recorded sounds to them when record time is over. Furthermore, she gives visual feedback (by hand- tapping or counting) and switches tracks when the software is done recording. It is important that the mediator does not interfere with users' interaction but only helps them communicate with each other through the audio software.

Different design features related to time, visual sensation, haptic sensation, physical object, and scaffold have been explored to find factors that influence mutual sense-making. These design features were iteratively revised in 3 workshops spanning 3 sub-cycles. Specifically, design principles for participatory sense-making and inspirational content were used as guidelines to develop a prototype. These design features that are integrated into the prototype were evaluated in each workshop. The goal was to find the design features that support positive sensation, meaning-making, stimming encouragement, and connectivity.

The proposed design features are shown in Table 6.2. The observations from the workshops, which will be discussed later, show that time and additional scaffold can foster a mutual sense-making process between participants.



2	Microphones	1	Digital Audio Workstation software
1	audio interface	Multiple	ordinary objects (for sound inspiration)
2	headphones	Multiple	Sticky notes as physical commands

Figure 6.1: Design equipment for the Wizard-of-oz set up

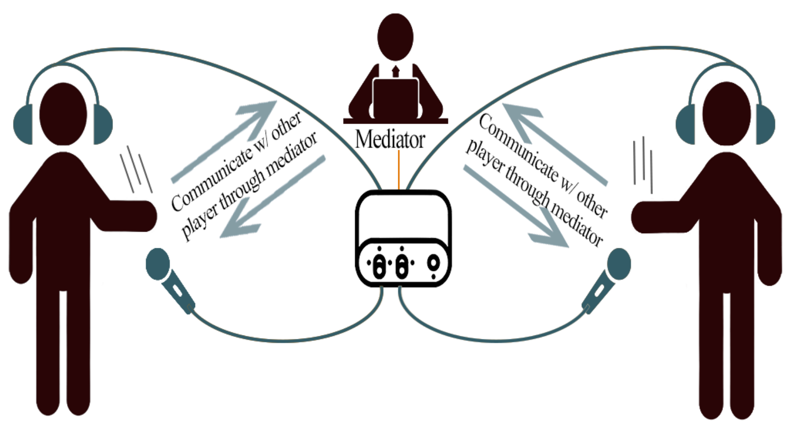


Figure 6.2: Demonstration of the low-fi prototype

I decided to integrate the feature of medium turn length with fewer turns for recording time and a metronome in the prototype. Although a shared microphone can provide more engagement, joint attention, and synchronization, it was not chosen. This is because it requires a high level of synchronization from body coordination, body language, and facial expression, which are the communicative features not preferred by autistic people. Nevertheless, the takeaway point is the shared input, but it shouldn't be the physical object. Alternatively, a shared digital input that may not involve these communicative features could be explored in the next design cycles.

6.3. Workshop execution and results

Design choices	mutual interpretation	joint attention	mutual synchronization
Short turn length(~4 secs) with more turns	--	--	-+
Medium turn length(~20 secs) with less turns	-+	-+	-+
Additional visual sensation (satisfying video)	--	--	--
Additional embodied visual cues (flash lights attached to hand)	--	-+	--
Shared physical object microphone	++	++	++
Depriving visual input (blind-fold interaction) with Haptic communication command (different tap code on mic)	--	-+	--
Additional structured scaffold (metronome added)	++	++	++

--: low -+: medium ++: high

Table 6.2: 7 evaluated design features in the workshops

(a) Workshop between M and P (b) Workshop between V and me (c) Workshop between T and me

Figure 6.3: Six activities from the first two workshops

Execution

I conducted the next three workshops with 4 NT participants, as I refer to as M, P, V, and T, to evaluate the low-fi prototype. In this cycle, the first workshop was between 2 NT: M and P (Figure 6.3a). The second one was between V and me (Figure 6.3b), and the third one was between T and me (Figure 6.3c). The goal of these sessions was to evaluate the designed features that influence the way people make sense and connect. Before the experiment started, the participants were briefed about the stimming definition and how the prototype could encourage them to perform such behavior with sound. When briefing them about the goal of the workshop, I explained that they have the freedom to create any sounds they like, without any rules restricted. Furthermore, they also got instruction on how the prototype works, and what they could do with the materials and equipment provided. After the experiment, they would participate in an unstructured interview (group interview for the first workshop).

The answers from the interviews from the workshops were categorized into themes that relate to:

- how participants communicate with each other
- information they use to communicate
- features promote participatory sense-making process
- factors hinder sense-making process
- experienced feelings by participants

The classified themes are shown in Figure 6.4, Figure 6.5, and Figure 6.6.

Results

According to the participants, they described the experience as a creative, fun, engaging activity of making sound together. The participants enjoyed the process of sound exploration and experiment: *'the creating process is engaging, it gives me flow'* (said by P), *'I want to put more effort to make the sound sounds better'* (said by T). They usually concentrate and focus on exploring different sounds, through body and hand movements. Thus, it leaves room for them to freely express their behaviors. Also, it has been observed that they do not need to mutually coordinate their body movements together, which makes it an asynchronous bodily interaction.

Furthermore, sound takes most of their attention from mutual interaction because being in the soundscape zones them out from the real world. In the fourth workshop, V said *'I was doing my own thing, I was not looking at you'* when was asked about how he was paying attention to what his peer doing.

Furthermore, the focus on facial and verbal communication is shifted. Participants communicate with each other through sounds. It is often seen that one responds with his sound when he likes a sound that his peer is making. Specifically, the pattern of how the sound is made also creates a certain meaning for both of them. For example, P said *'When he made the rhythm, I know he wants to record'*. They were making meaning through sound. In addition, sharing the same space enhances their meaning-making process because through visual cues they could observe and understand one another's ways of making sounds. In this case, visual cues of body language foster mutual understanding. However, there is still confusion in understanding the situation between them. Sometimes one participant doesn't understand the intention of the other's behaviors, other times both of them are puzzled about how and when to collaborate.

Another important factor that enhances mutual connection is time. A longer recording time allows participants to spend more time and get to know each other better through their communication style. Specifically, their sound story is more synchronized and coherent when they spend more time together. With time collaborative tasks also increase their motivation and inspiration.

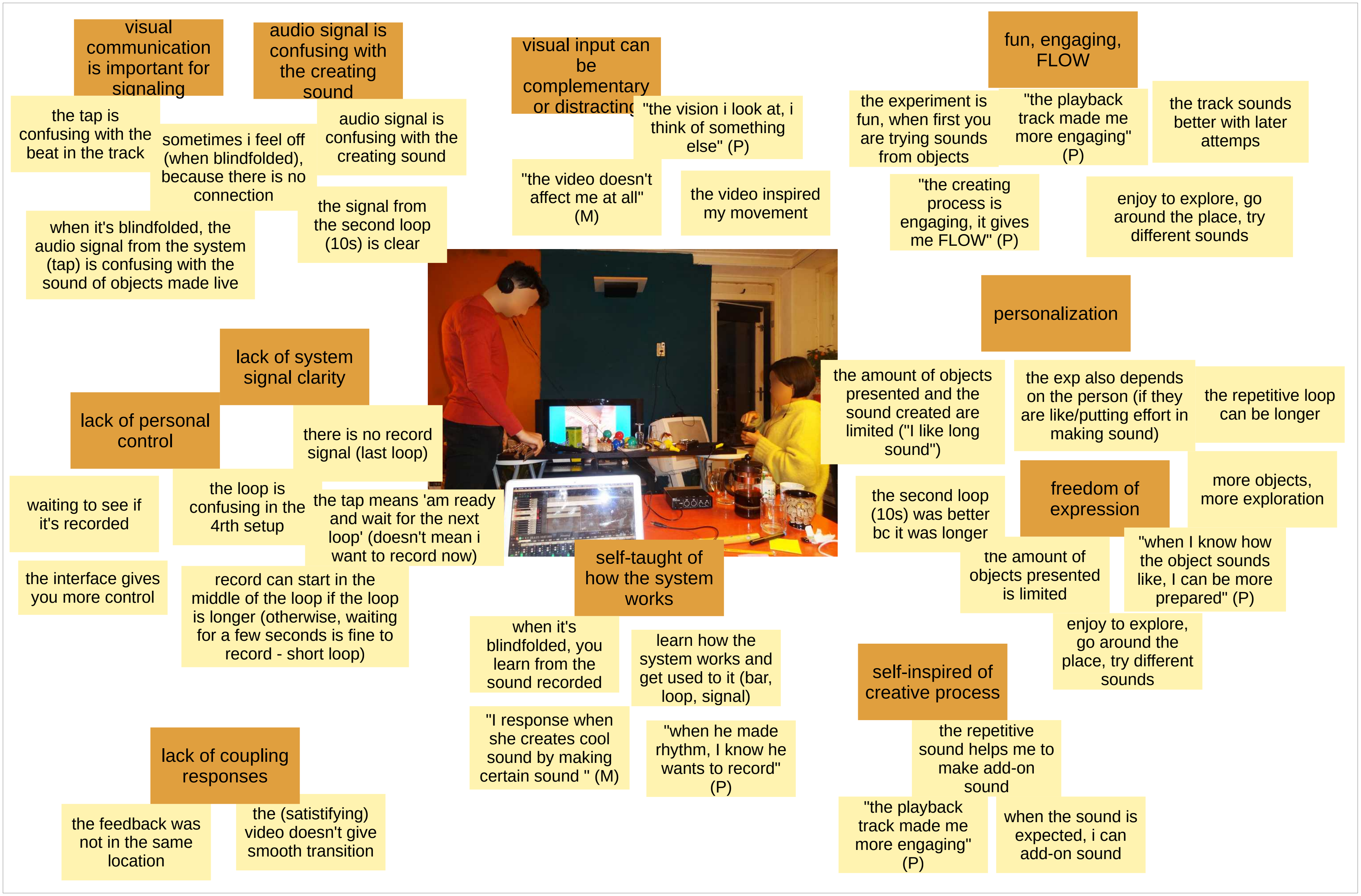


Figure 6.4: Results of categorized themes from the third workshop between 2 NT participants

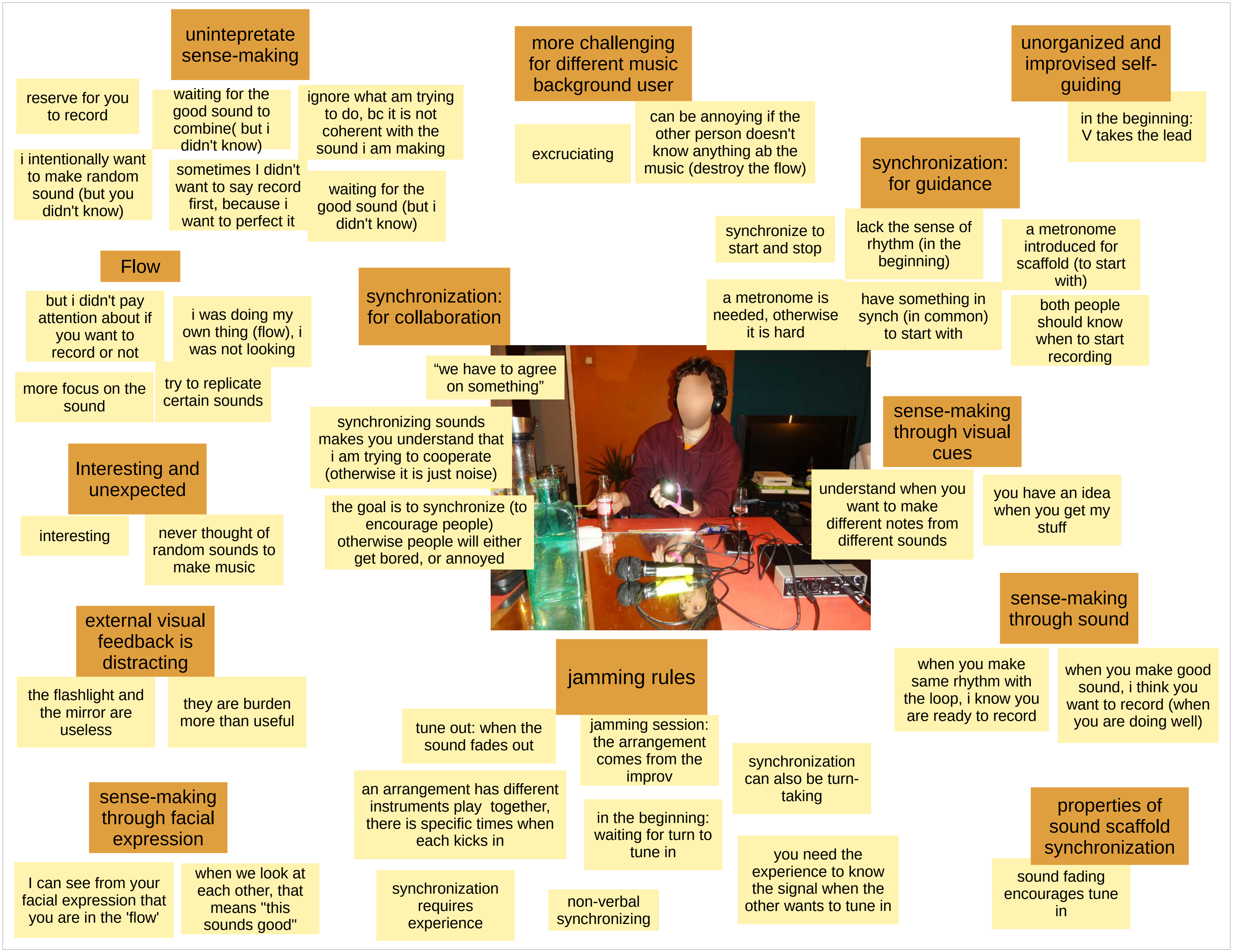


Figure 6.5: Results of categorized themes from the fourth workshop between a NT participant and me

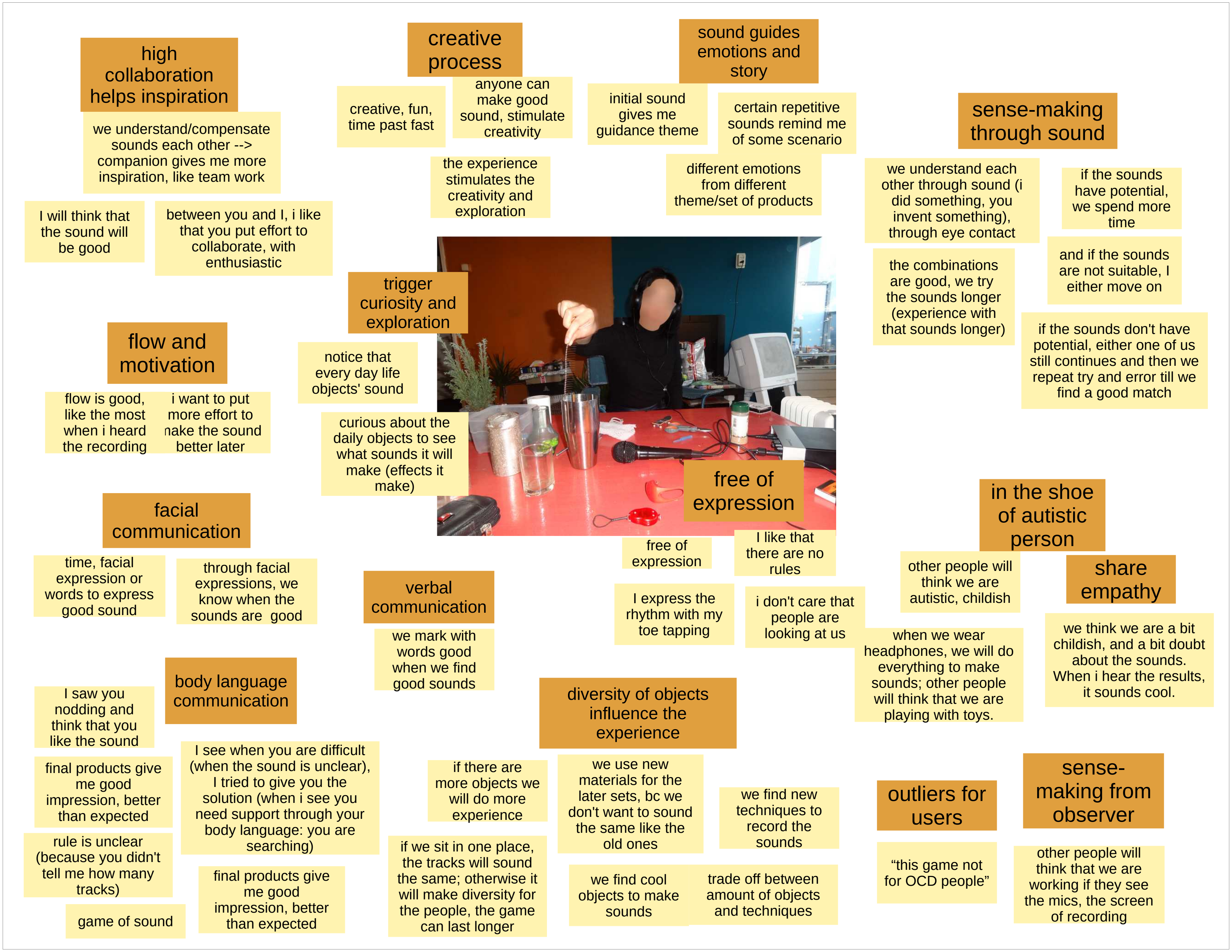


Figure 6.6: Results of categorized themes from the fifth workshop between a NT participant and me

Besides findings that are related to meaning-making, sounds also had an impact on participants' emotional feelings. Certain collaborated sound stories bring them memory, mood, and emotion. For example, T mentioned that the repetitive sounds from the first sound story remind her of being inside the peaceful temple.

To explore platforms that can help users connect and communicate, I evaluated sensational visual and light modalities. The results have shown that these non-sound modalities are distracting and adding fewer values to their creative process. This implies that sound is a very powerful platform because it takes lots of attention from participants' mental processes. Thus, I should focus on exploring different communicative approaches in the auditory modality in order to deliver a better way people connect.

Regarding the technical aspects, the incoherent and uncoupling feedback from the system makes them confused. Also, the current prototype has not yet provided independent control for participants. Thus, the next implementing prototype should take intuitive user interface design features into account to provide them a clear, coherent, and autonomous form of communication.

Summarized results from workshops:

- Sense-making is established through sound and visual cues through shared space
- Sense-making through sounds is not consistent
- Longer time experience increases mutual connection
- Different interpretation and purpose of sound occurs due to individuals' differences in sound expertise
- Freedom of expression through body languages, behaviors, and non-verbal form is observed
- Non-sound modalities input are distracting and add insignificant values in the experience
- Collaboration enhances motivation, inspiration, and meaning-making
- Sound elicits emotional feelings
- Asynchronous mutual body coordination is observed

6.4. Intermediate evaluation and discussion

From observations, I have learned that sense-making through sound is not yet consistent. One of the reasons could be the individual differences in skills, expertise, and knowledge about music and sound. Such imbalance can create different mutual expectations between peers. Another reason could be that too much freedom is introduced in the experience, which can cause a non-structure flow process. By having a norm-free space without specific rules and instructions, participants encountered confusing moments during their interaction. This indicates that a certain amount of scaffold is necessary to give them common tools to guide them through. Additional scaffolds can help them to establish better mutual ground, provide clearer directions to interact, and synchronize better. For example, a metronome can encourage them to create more repetitive and coherent sounds, which eventually contributes to enhancing their sense-making.

A possible way to foster sense-making through sound is to make stimming movements more noticeable in the soundscape. For example, the sounds made by the movements can be used to amplify the act of stimming. These sounds, which are seen as traces by (Hummels and van Dijk, 2015), can provide some sort of connection between stimming and sounds that eventually help people interpret the meaning of stimming through sound. This is rather an ambiguous idea, and needs more research done to have a concrete design feature. Hence, further exploration about the relationship between sounds and repetitive bodily movements in the next cycle needs to be done.

Besides, stimming behavior in the interaction is not yet focused. This is the direct issue stated in the main research question that has not been tackled. Stimming behavior has a clear purpose for autistic people, however, it was not clearly understood by the NT participant. In the current prototype's experience, the knowledge and purpose of stimming are not explicitly presented to them yet. For example, how do both people feel when they stim, what properties or kind of objects trigger them to make repetitive movements, etc. The problem of this prototype is that it presents the information of stimming obscurely and inexplicitly as it lets people explore the meaning themselves. Stimming relates to emotion, feelings, and sensation, such in-explicit framing is rather hard to accomplish. Thus, additional scaffolds can be used to bring stimming into focus.

Moreover, to increase the meaning-making process on a deeper level in people's perception, the prototype can give a takeaway home message to encourage them to reflect on their experience.

It is also observed that joint attention is established when the participants share the same microphone in the same physical space. They exchange objects and demonstrate different hand movements to show how they want to create the desired sounds. This scenario shows higher mutual engagement than the situation where each participant has a microphone. Thus, the next goal is to bring joint attention to the soundscape to increase mutual connection.

To tackle the technical design of incoherent feedback, I need to explore how to design a more intuitive, and flexible prototype. The next designed prototype can implement the consistency in location and time for a more natural coupling system, as proposed by (Wensveen et al., 2004).

7

Cycle III

7.1. Concept requirements

Additional requirements have been added and previous concept requirements are revised into more detailed requirements, on the results and the evaluation from the three workshop sessions from Cycle II. A list of these has been included here for completeness:

Revised:

- *focus on story-building collaboration*
- *offer sensation*
- *encourage and prioritize repetitive movements*
- *require minimal knowledge and skill to perform*

Additional:

- *explicitly present stimming as a foreground activity*
- *introduce additional structures/scaffolds to increase synchronization and clarity for the users*
- *introduce shared element for establishing joint attention*
- *consider flexibility and mobility factors to increase comfort and freedom*

7.2. Concept development and prototyping

To address the updated concept requirements, the new prototype should be able to provide the following features:

- mutual engagement: provide a longer recording time

Functions	*User Interface control*
Record sound from live microphone	receive command from user
Play recorded sounds	receive command from user
Provide multiple recording turns	
Receive record signal from user	receive command from user
Give signal to users when the system is about to record	send signal to user
Give signal to users when the system is done recording	send signal to user
Transfer commands from one user to another	send signal to user
Provide long (20-25 second) recording turn	
Add metronome for additional synchronization	
Provide shared sound effect	

Table 7.1: Table of required functions and user interface

- synchronization: a scaffold for promoting users' sound
- joint attention: a shared feature that allows both users to collaborate.
- communication: a better form of communication between peers.
- autonomy: give users more direct control with the audio software (when they want to send commands), **without** a human mediator.
- intuitiveness: a more intuitive control interface.
- flexibility: more flexibility and less restricted movement in the environment.

The functions that provide these features are summarized in Table 7.1.

Additional scaffolds are introduced in the system to provide more mutual engagement, synchronization, and joint attention for users. In particular, the system increases longer recording time to enhance mutual engagement. From the observations gathered in Cycle II, I have estimated that a maximum of 25 seconds would be required for each recording turn. This value gives users enough time to help them connect.

To scaffold users' synchronization, a digital metronome, which is provided by the DAW, is used to help the users structure their sounds making process. Because it provides constant repetitive beats that they can listen to during the playback and recording time.

The system can foster joint attention between the two users providing them a shared sound element. This element can be, for example, a sound effect controlled by the users. In this design cycle, the sound effect used was a sidechain compression effect. It works as follows: when one user's sound input volume is above a certain threshold, then his sound will be compressed by his peer's sound volume. More precisely, the level of the sound effect of one user is controlled (or compressed)

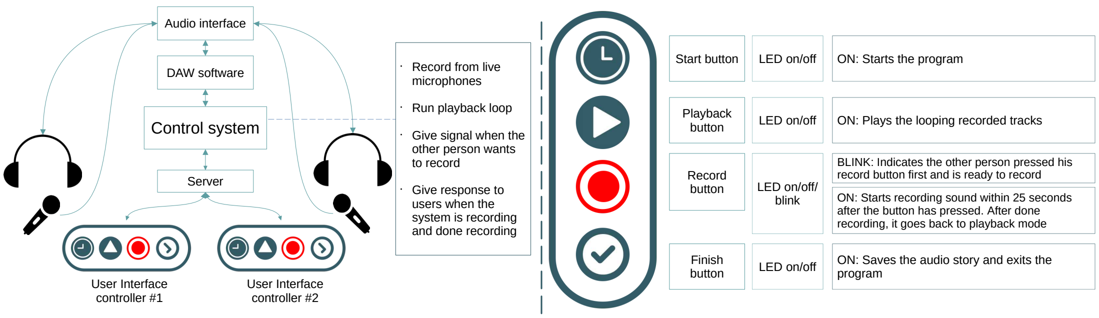


Figure 7.1: Overview of the system structure and the User Interface Controller (UIC)'s functions

by the sound volume level of the other one. This happens at occasional moments in time during playback. The idea is to get both users' attention and encourage them to be more engaged and collaborative with each other's sound activity.

A better communication form between users can be made with visual feedback from the system. As observed in the previous cycle, the additional visual sensation is more distracting than effective. Also, haptic communication is not recognizable as it blends with sound input. However, visual feedback from the mediator signal is effective because it does not interfere with users' interaction and is in the view of their space. As such, visual feedback can be implemented as a communicative tool for users.

A new set of functions has been updated to include the features of autonomy, intuitiveness, flexibility, and communication, as listed in Table 7.1. However, these new functions raise some very important issues related to the recording software and hardware. More concretely, the sound recording and music production software can be hard to use by nonexperienced users. As far as this prototype is concerned, there is too much unnecessary functionality, and too many unneeded options to choose from in existing DAWs. Furthermore, most DAWs can be controlled by external devices, such as Musical Instrument Digital Interface (MIDI) pads, keyboards, etc, which are simply too bulky and hard to use by non-experienced users in an ad-hoc way. On the other hand, it is infeasible to design a complete DAW from scratch, that is specifically tailored for this prototype. For this reason, it was easier to design a control system that is specifically tailored for this concept because it takes advantage of the essential functionalities of existing DAW software, and can still provide an autonomous and intuitive user interface. This system, shown in Figure 7.1, consists of three main components: a User Interface Controller, a *flow* controller, and a DAW control interface.

Flow controller

The first component of the system is the so-called *flow* controller. This is a dedicated piece of software that runs in parallel on the same host computer with the DAW. Its main purpose is to establish an interface between the UICs, and synchronize the sound and music creation flow of the

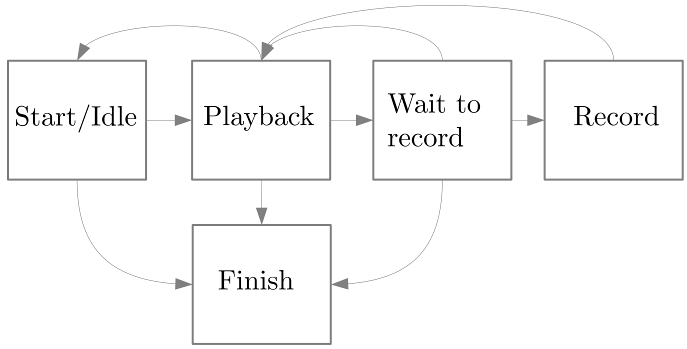


Figure 7.2: State machine

users. As such, it also takes the responsibility of exposing the functionalities and options of the DAW in an intuitive and condensed way to the UICs. For now, the flow controller has been designed to accommodate for a total of two UICs in this system prototype.

The basic idea of this component is as follows: when two UICs are connected to the flow controller, which acts as a network server, a state machine starts (Figure 7.2). There are six states: start/idle, playback, waiting-to-record, record, waiting-to-finish, and finish:

1. Start/idle state initializes the DAW software
2. The playback state plays the recorded tracks in the DAW in looped mode(if they exist).
3. The waiting-to-record state is entered when one of the users presses the record button.
   - The system remains in this state for a certain time until the other user to presses their record button.
   - When both users press their record buttons, the DAW will start a countdown, and enter record mode afterwords.
4. In record mode, the DAW software shall record sounds for a time interval of 20 to 25 seconds.
5. Finish mode is entered when both users indicate their willingness to stop the whole system.

When the recording turn is over, the system goes automatically back to the playback state to let users listen to what they have just recorded. In the next recording turn, the system will record the sounds on top of their previous recordings. Usually, four turns are available for recording. When the users recorded all the available tracks, the DAW will play an audio notification that lets the users know if they have exhausted all available recording tracks. In this case, the program automatically enters the finish state. In this state, the program will render the audio file and the system will shut down.

The UICs directly controls the switching of the state machine via network sockets. A special protocol is used that translates the user input actions into software commands and the states into LED indications.

User Interface Controller

The second component of this system is a hand-held User Interface Controller device that is to be attached to the wrist of each user. From here, the user should be able to issue commands, and receive feedback from the DAW as described in Table 7.1. This device must also connect wirelessly to the DAW and flow controller, in order to provide flexibility for the user's movement.

Hardware platforms, such as the ESP32 DEVKIT, a Raspberry Pi (RPi) Zero, a smartwatch, and an Arduino were considered for the implementation of the UIC.

However, while the ESP32 provides wireless connectivity, it is simply too difficult to program correctly. On the other hand, the Arduino is easier to program but lacks an easy-to-use wireless interface. These platforms also need to be programmed with C/C++, which are significantly harder programming languages to use, compared to e.g. Python.

Smartwatches have one important feature that makes them an attractive choice for this concept: they come with a well-designed wrist strap, and the battery is integrated inside the watch. However, smartwatches turned out to be very difficult to program, and they require the use of a smartphone as an intermediary device, which definitely reduces the flexibility factor of the UIC.

For these reasons, a Raspberry Pi Zero development board was ultimately chosen to implement the UIC, because of its ease of use (via Linux+Python), wireless network connectivity, and its small form factor, thereby being flexible enough to attach and carry on the wrist. The RPi also has considerably more processing resources, compared to the other platforms, which turned out to be a valuable asset in Cycle IV. A drawback of the RPi, however, is that it may at times consume considerably more power than the other platforms. Thus, the battery may need to be recharged or replaced more occasionally. This does not pose a big challenge to this prototype, because the system does not require long-term usage. As far as sensors and actuators are concerned, Light Emitting Diodes (LEDs) are used to indicate the status of the system to the user. Push buttons provide simple digital input. Different revisions of the UIC prototype are shown in Figure 7.3

DAW interface

The final crucial component of the control system is the interface between the flow controller, and the DAW. Either the Musical Instrument Digital Interface (MIDI), or the Open Sound Control (OSC) (Table 7.2) can be used. Without going into too many technical details, a MIDI controller can be used to issue direct commands to the DAW, such as going into record mode, turning a knob, and playing a virtual instrument. However, this interface is not suitable for providing automated and external functionality to the user from the flow controller, such as giving visual feedback and changing recording tracks, which requires low-level internal access to the DAW's

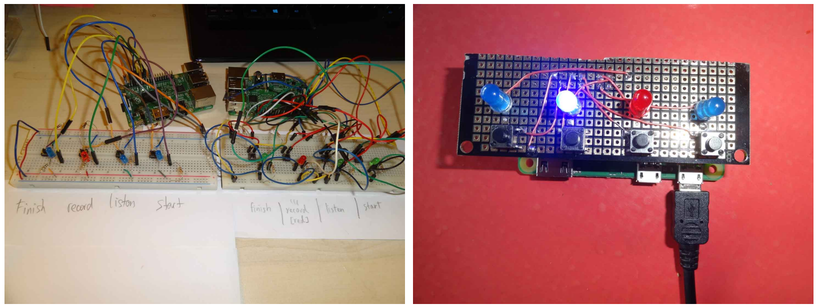


Figure 7.3: Left to right: The first and second prototypes

internal Application Programming Interface (API). On the other hand, the OSC interface can be used to access low-level commands of the DAW, without having to access the API. Thus, it is a more suitable choice in order to expose a subset of the internal DAW software functionality to the user. Once the OSC inside the DAW software opens a port and gives the IP address of the device that is running the DAW software, the flow control interface can establish a local connection with the DAW software.

Algorithm	Pros	Cons
MIDI	can control internal functions from a DAW program	can't provide external functions from DAW to send feedback to users
		focus on producing musical sounds e.g. guitar, piano and musical properties e.g. keys, notes, range.
Customized external control interface	can be built to control both internal and external functions from a DAW program	need hard coding to control internal functions
	can connect with a DAW program through Open Sound Control (OSC)	
	give participants to make sounds by exploration	

Table 7.2: Pros and cons table for comparing MIDI and customized external control interface

7.3. Workshop execution and results

(a) Workshop between P and me (b) Workshop between M and me

Figure 7.4: Two workshops were conducted

Execution

The next two workshops were conducted between me and the participants (P and M). The first workshop with P (Figure 7.4a) mostly aimed to test the functionalities and performance of the early prototype with the circuit board. The second workshop with M (Figure 7.4b) was to quickly test the revised functionalities of the second prototype that only took about an hour. Thus, in the second workshop, only a few questions were asked. Besides testing the technical performance of the prototypes, these sessions were also to evaluate the chosen designed features that affect joint attention, synchronization, and mutual connection. Specifically, how different properties of sound influence emotional feeling and joint attention. Before the session started, the participants were informed that sound effects would be used in the workshop.

The interview's answers from the first workshop were categorized into themes that relate to:

- how participants communicate with each other
- features promote participatory sense-making process
- factors hinder sense-making process
- how sound effects emotional state
- how the design features effect joint attention, synchronization, and mutual connection

However, the answers from the second workshop were not analyzed but used as feedback notes because the interview session was too short. Furthermore, technical problems were reported in both sessions. The classified themes are shown in Figure 7.5.

Results

The participants indicated that the updated prototype gives them more freedom to control and provides clearer feedback in the interaction. P had a remark: *'We don't have to think about the*

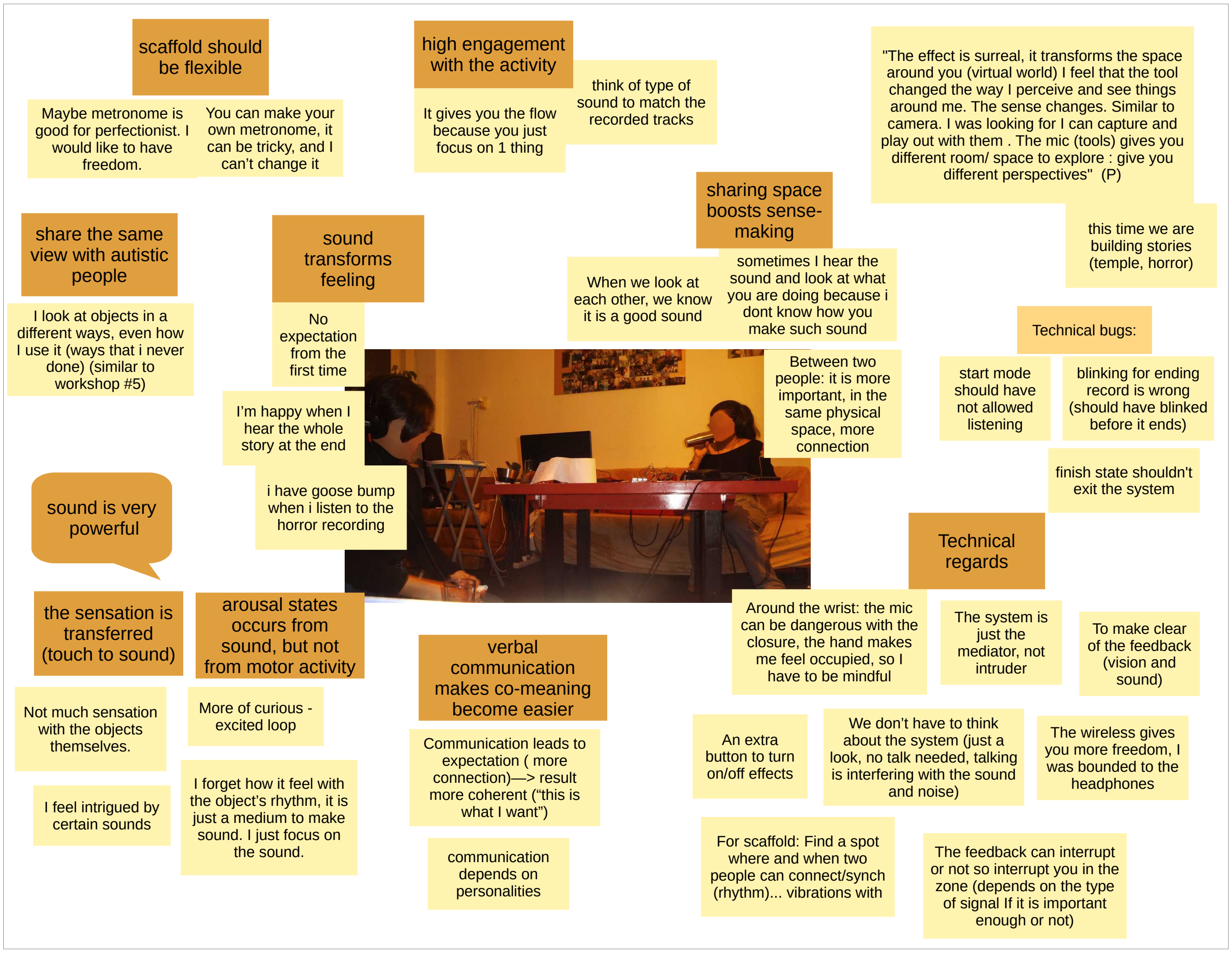


Figure 7.5: Results of categorized themes from the sixth workshop

system but just need to look at the controller (to understand what is going on). No talk is needed because talking is interfering with the sound and noise.. Furthermore, the metronome provides better synchronization and guidance. However, participant P noted that her freedom to explore was constrained: *'The metronome is good for a perfectionist, but I would like to have freedom'.*

Furthermore, the chosen joint-attention shared feature, which was a sidechain compression sound-effect, was not noticeable for participants. This happens because the effect was exposed in a very brief moment during the playback loop, hence the participants could not hear the difference. However, when the sound effect was exposed throughout the whole loop, it was recognized.

Sound effect(s) brought different feelings and perspectives to the participants. For example, the reverb effect reminded them of being in a scary place. Specifically, they made a horror sound story together when this effect was added to one of their tracks. When listening back to the story, P said

'I have goose-bump when I listen to the horror recording'. She also stated that the effect transforms her space: *'The effect is surreal, it transforms the space around you (virtual world). I feel that the tool changes the way I perceive and see things around me. The sense changes. Similar to the camera. I was looking for what I can capture and play out with them. (Here) The mic gives you a different room/space to explore, it gives you various perspectives.*

The observations show that the participants did not feel the sensation from manual haptic exploration. Instead, they described the objects used as tools for them to make sounds.

In the first workshop, P and I mostly communicated and made sense together verbally, which was different from her previous workshop with M where verbal communication happened sometimes. She remarked that this time our sound story resulted in a more coherent piece. Besides, we also understand each other's actions by visual cues provided by the same view in shared space. For example, we looked at each other when both of us liked the sound made at that particular moment.

Regarding the technical problems, the designed LED light feedback though was clear but inconsistent, due to malfunctions and bugs in the software. The participants were confused about the light signal indicating the 'ready' state. Sometimes, they could not interpret the meaning of the UIC's third light 'blinking' as 'the other one is ready to record'. The malfunction of the visual feedback also adds more confusion for them.

Furthermore, there was not enough feedback from the system to inform the participants about the status of the audio software. However, when audio notifications, which are voice notifications of countdown tics, were added in the second updated prototype, the feedback became clearer. Specifically, the countdown tics gave M more preparation for making a sound.

Summarized results from workshops:

- Metronome restraints some freedom from exploration
- Visual feedback from design features is still inconsistent
- Audio feedback from design features is clear
- The prototype gives more independence for participants' interaction
- Sound effect is not recognized when activated in a short time frame but is recognized when it spreads throughout the whole track.
- Sound (effect) transforms feeling and perspective
- Object is medium to make sound, and its interaction does not provide a haptic sensation
- Sense-making is established through verbal and facial communication and visual cues from shared space

7.4. Expert interviews

The classified topics from the experts' opinions are shown in Figure 7.6.

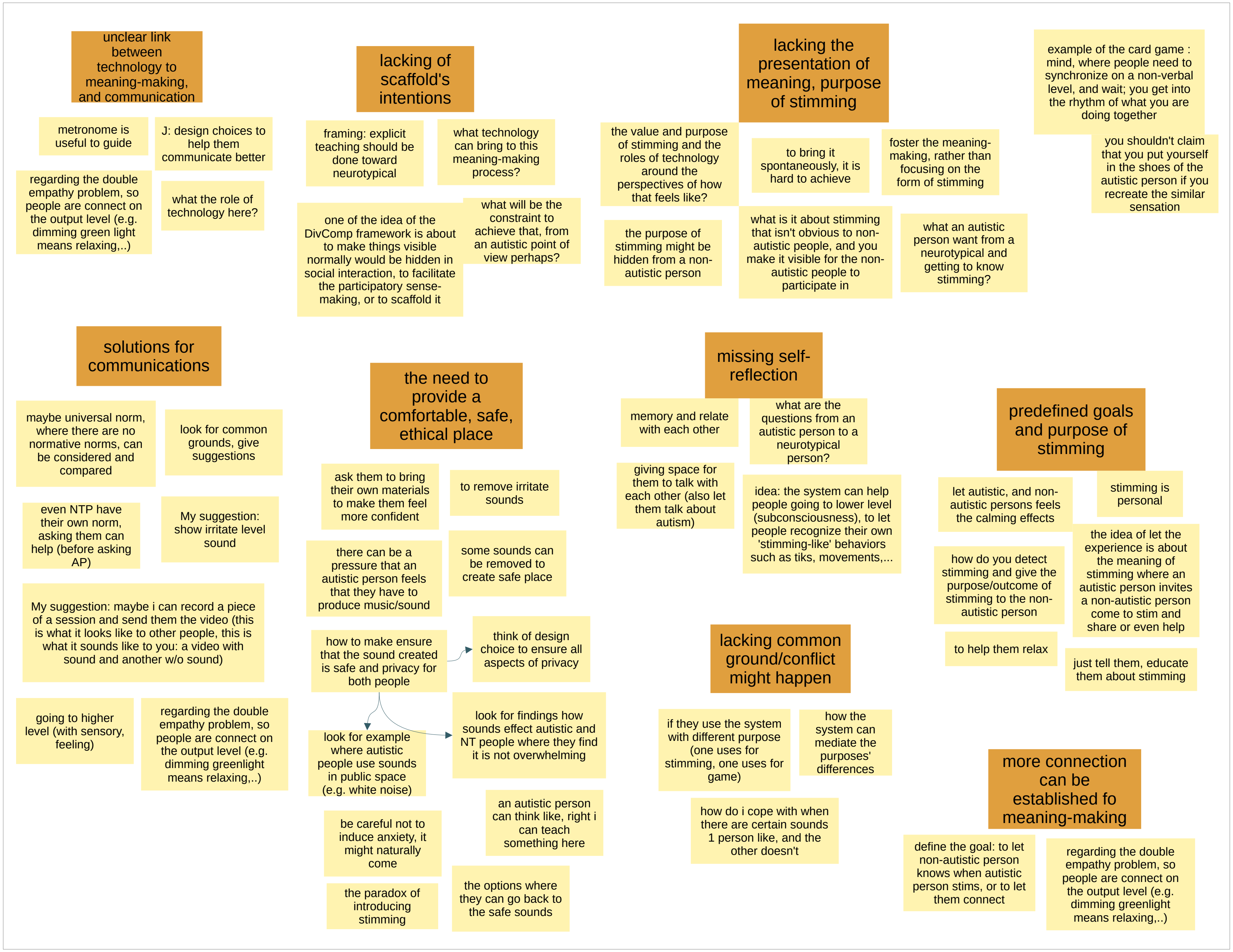


Figure 7.6: Results of categorized theme from three expert interviews

Results

According to experts, the intermediate prototype though encourages people to stim to make sounds together, still lacks presenting the role and purpose of stimming explicitly. A deeper level of understanding can be why or what an autistic wants a NT person to know about stimming. The stimming purpose can be apparent to an autistic person but often hidden from a NT person. The experts considered that the act of stimming was present but its meaning was not yet well fostered and explicit meaning was fairly hard to achieve spontaneously. It also depends on how the initial framing, a briefing will help them establish the meaning of stimming.

Moreover, self-reflection, which is part of meaning-making, has not been established in the experience. Memory which can be provided by the prototype can help them relate with each other. The activity experience can also give them an opportunity to talk about stimming and to realize their subconscious behaviors that resemble stimming (e.g. tic, fidget).

How technology can bring its benefits to the meaning-making and communication process is another concern. The raised questions were *Will people get connected on the output level (behaviors and feelings), or the input level (perception)?*, *what technology can bring to the meaning-making process?*, and *what will be the constraint to achieve that, from an autistic point of view?*. One of them further explained that *'One of the key ideas of the Diversity Computing framework is about making things visible that normally would be hidden in social interaction, and to facilitate the participatory sense-making, or to scaffold it'*. Technological features can be designed to bring its impact on the meaning-making and communication process between users.

There could be differences in using the system and personal interest between two different users. For example, one might use the system for stimming while the other one uses it for the game, or a particular sound can be liked by one and disliked by another. The latter is important because an autistic person might have a sensitivity to certain sounds and noise. This concern is linked back to the role of technology: how technological features can mediate individuals' differences in interaction.

Another important factor is a comfortable and ethical place. When they participate in such sound-making activity, certain sounds can be irritated and uncomfortable for either or both of them. Furthermore, by encouraging them to stim, the system might induce anxiety in the autistic person or make him feel calm. This is rather a paradox of introducing stimming as it can create overwhelmed or relaxation. Also, with the microphone is always on, how can privacy be assured for both people?

Summarized results from expert interviews:

The missing gaps of current prototype are:

- The purpose and meaning of stimming
- A common ground of the interactive purpose
- Self-reflection on their behaviors and relationships with the other
- The role of technological features to support meaning-making process, communication, and individuals' differences.
- Consideration about comfort, safety, and privacy

7.5. Intermediate evaluation and discussion

It is unavoidable that the people may feel constrained because the added metronome gives more structure to the activity. And, because one of the main goals is to achieve more structured and synchronized interaction from the start, lesser freedom of expression was therefore acceptable. Thus, the decision of using a metronome initially should be kept.

It makes the controller becomes a mediator but not an intruder in the activity because its visual feedback can provide system messages to people within a glance. However, the coding scheme and technical difficulties of the visual feedback have to be improved. Audio feedback is clearer because it delivers a straightforward voice message in the same soundscape. However, there could be a possibility that it is intruding with users' sound activities. This leads to the concern of when to use visual and when to use audio feedback. Hence, visual and sound feedback design features need to be revised to provide a more coherent, clear, and nonintrusive.

The goal of using sidechain compression sound-effect was to provide the ground for joint attention between people. However, this design feature can not be evaluated because people were not able to recognize the effect. A probable reason is that this sound effect is exposed in a brief moment in a track where all the tracks are layered on top of each other. Thus it is hard for people to hear the difference in sound. However, they can recognize the sound difference when the effect is exposed throughout the track. Then, the second choice for integrating the effect in the track will be used to evaluate joint attention in the next design cycle.

The observation shows that there was no haptic sensation found in this prototype. As one of the concept requirements is to provide a kind of sensation, more research about this needs to be done in the next design cycle.

Similar results compared to the previous cycle, observations in this cycle show that sound induces feeling for participants and helps them make sense in the same view space. The results also show that co-meaning is established more coherently with verbal communication.

Based on the mentioned issues from three experts' meetings, they also gave suggestions to alleviate these problems.

To foster the meaning-making of stimming better, one possible way is to use the system to help them connect at a higher level of sensory and feeling. Another way is to present the meaning more explicitly by directly introducing the meaning and purpose of stimming to them. However, in my opinion, this might change the purpose of the whole experience as the main goals are to let both of them co-create a new meaning of stimming by themselves and to behave in their newly established norm. Diversity Computing should be the mediator to provide tools and a participatory sense-making platform to help them accomplish these goals. Instead of imposing certain predefined meaning for them, DivComp lets them make meaning themselves. One possible way to present stimming more explicitly is to introduce the definition of stimming only.

Regarding the comfortable participating place, an expert suggests that letting users bring their (stimming) toys or objects, it can make them feel more relax and familiar in a new environment. Another expert suggests that the system can manipulate sounds to eliminate uncomfortable sounds or to provide them a safe soundscape to go back. These suggestions have indicated that the prototype should flexible to adapt to people's discomfort.

One possible way to both provide the meaning of stimming more explicitly and comfortable place is by creating a feeling-stimming relationship. Since the system needs to apply more scaffolds to encourage them actively perform stimming-like behaviors, it can also induce comfortable feelings when they stim. So instead of directly provide a relaxed feeling when they stim, the system can let them manually adjust their emotional state by stimming to manipulate certain comfortable properties. Then users can have control over their calming influencing variables. This way can help them connect the meaning of stimming with their act of stimming better.

Furthermore, the system can introduce more common rules and structural elements to alleviate individuals' differences. Common rules can bridge different communicative approaches between an NT and an autistic person and structural elements can help them synchronize better.

Regarding self-reflection, the system may find a way to help them bring messages home about their experience. For example, the system can deliver memory to them in multiple formats such as videos, text, audio.

8

Cycle IV

8.1. Concept requirements

Base on the evaluation from workshops and expert interviews from Cycle III, these are the *revised concept requirements* for the next prototype. It should:

Revised:

- *focus on building story collaboration*
- *offer sensation*
- *require no learnt skill*
- *present meaning of stimming more explicitly*
- *have shared element for establishing joint attention*

Additional:

- *use technology to introduce more structure and establish common rule in the system*
- *provide scaffold to strongly encourage them to stim and manipulate their own emotional state.*
- *have design features calming, relaxing variables*
- *have design feature in delivering take-away home message*
- *deliver clear and non intrusive visual and audio feedback*
- *have flexible design to adapt to and notify people' discomfort*

8.2. Concept development and prototyping

The autonomous control system prototype developed in Cycle III was a step further than the initial low-fi prototype from Cycle II, where the human mediator was needed to manually perform the tasks of the flow controller. However, in this cycle, I introduced a new shared feature in the form of a gesture recognizer, background sounds, and improved communication.

Hand-gesture recognizer

To provide a more intuitive way for the users to control their soundscape more easily, the UIC has been equipped with a hand-gesture recognizer. This recognizer uses a Machine Learning (ML) model and algorithm to identify a given gesture and produce a corresponding command that is then sent to the flow controller. As such, the hand recognizer is used to manipulate the sound effect used as the shared joint-attention element introduced in Cycle III. This new feature was introduced, because, as observed from the workshops in the previous cycle, the intermediate sidechain compression effect was not as effective as expected: it made users put too much effort into paying attention to it. The idea is to allow users to manipulate sound effects by adjusting the values of three sound properties at any time, namely pitch, frequency cut-off, and volume using three predefined repetitive gestures. These particular effects were chosen because they are important influencing factors of how people perceive the pleasantness of sounds, which was shown in the related work. There are five advantages with this design feature:

- Gestures provide more flexibility and affordance to the users to stim, as they have full control of the sound effects.
- The joint attention from one to another can be increased, the sound volume of one user is controlled by the another's gesture. Also, pitch and frequency cut-off parameters are affected on both users' soundscapes.
- By incorporating three repetitive behaviors to manipulate the sound properties, the system allows the users to directly adapt, and manipulate their own emotion and comfort.
- The feedback loop from these repetitive behaviors creates *perceived coupling* that helps users to understand the meaning of these behaviors, and specifically stimming, in their own way.
- The cause-effect relationship between the gesture and sound effect represents the exact purpose of stimming less explicitly

Classification model

Although there are libraries that provide hand gesture recognition based on a ML algorithm, such as uWave by **?**, these usually contain pre-trained gestures that are not suitable for this concept. Additionally, these libraries have been designed for more dedicated purposes, and hence are harder to use.

On the other hand, the availability of many ML training libraries in Python, such as Sci-Kit Learn, enable very easy training and deployment of an ML algorithm. Thus, it has been decided to design and train a Neural Network (NN) classifier from scratch. The choice of NN as a classification model is clear: NNs have been shown to exhibit very high classification accuracy in many popular

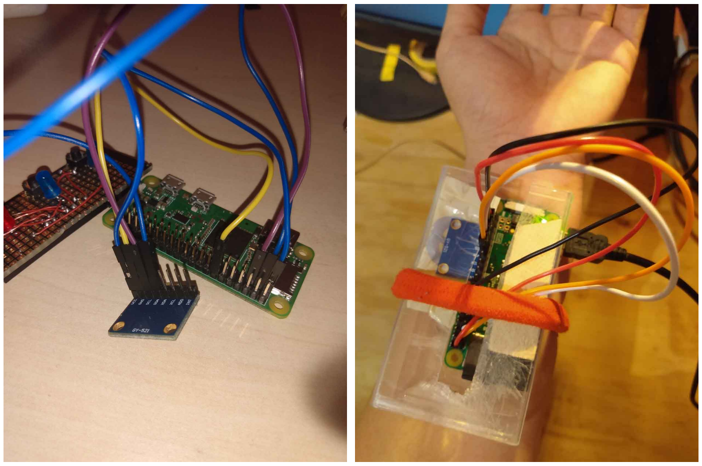


Figure 8.1: Left: MPU6050 is attached for testing, right: Attached MPU6050 for collecting data

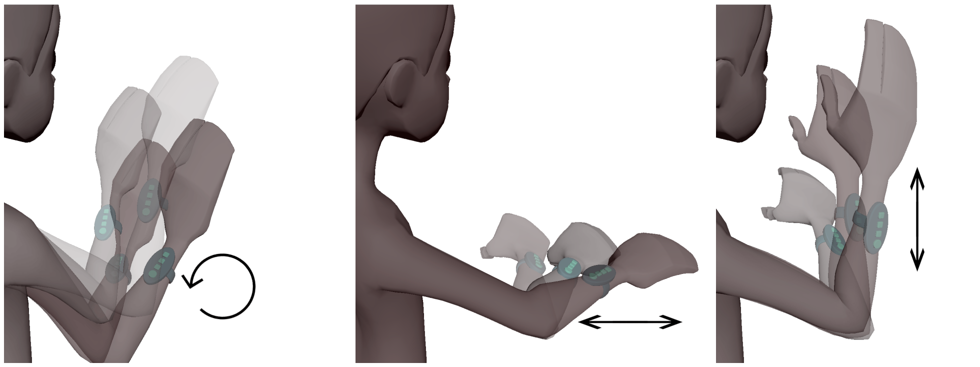
applications, such as Facebook's photo recognition feature. Also, NNs seem to handle multi-class classification better than, e.g. Support Vector Machines (SVMs). To be sure, in this cycle I have compared the NN classifier with a SVM classifier. In the end, the NN classifier beats the SVM with a $\approx 95\%$ prediction accuracy on $50-50$ training-test data split, whereas the SVM only managed to achieve $\approx 68.9\%$.

However, in practice even the NN model performed rather poorly, with the network occasionally getting fooled by unwanted gestures. This issue was addressed by introducing a post-classification algorithm on the output that tracks the occurrence count of each detected gesture. In the end, the gesture with the highest occurrence count is selected as the ground truth.

Data input and collection

The input to the classifier is a vector of samples that are periodically read in batches from a 6 Degrees of Freedom (DoF) Inertial Measurement Unit (IMU) device with a sampling rate of 10 samples per second (100 ms per sample). Specifically, the IMU consists of an accelerometer and a gyroscope, which provide measurements along each axis. The sampling rate was chosen so, to give enough time for the NN classifier to perform its inference in real-time and to reduce the data size of its input. The length of each vector depends on the time window used to capture each gesture. The specific IMU device chosen for this prototype is the MPU6000, due to its cost, and a large collection of code examples available on the internet for the RPi.

Regarding data collection and labeling, an easy-to-use setup is proposed: a Python script was developed to continuously capture 2-second recordings from the IMU, and store each recording in a file. This timing has been chosen based on the average amount of time required to perform a gesture. Thus, based on the selected sampling rate, each timing window produces a 20 sample vector that would be fed into the NN.



(a) Wiping-window gesture (b) Swipe gesture (c) Swing gesture

Figure 8.2: Three recognized gestures

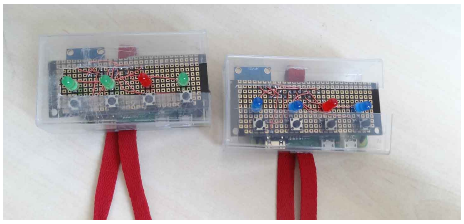
Then, within each time window, the desired gesture is performed, and its recorded samples are stored in a numbered file, which in turn is stored in a folder that is named according to this gesture's label. A small pause is also included in between each recording to allow the user to recover and prepare for the next recording. The script itself can be set up to stop whenever a maximum number of gesture examples have been collected. In the end, this setup was used to easily collect, label, and organize hundreds of training and test data for the classifier. In total, 300 training and test, examples were collected per gesture. For the unwanted gestures (negative training examples), the script is let to run continuously, by letting the user freely do whatever they want with their wrist.

The gestures I have selected for this prototype are hand- wiping-window, swipe, swing, and gestures (Figure 8.2). These gestures were chosen because they were easy to perform by users and delivered the highest prediction accuracy with the NN.

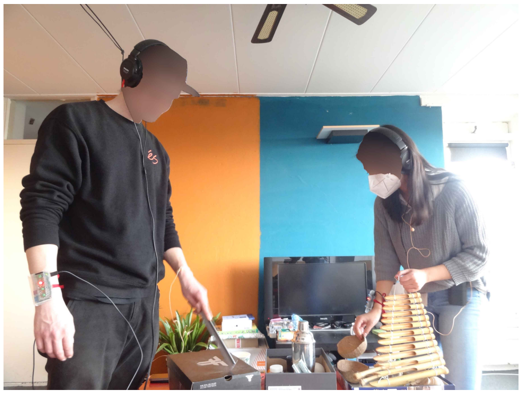
Background relaxation sounds

Another feature introduced in this cycle is relaxing sounds that are played at beginning of the interaction. This gives a good initial impression and a comfortable sound zone for users. Specifically, nature sounds such as birds chirping, ocean waves, cricket sounds were used because they are associated with positive things, as has been shown in the literature. The background sounds can also be used by the users to learn and test the gesture recognition at the beginning, which also allows them to stim without an object at first.

In addition to providing a comfortable space, users are also encouraged to bring objects that make them feel comfortable or make sounds that they prefer. This feature will make them feel more inviting and give them more affordance to participate in the activity.



(a) Final UIC prototype



(b) The workshop between the autistic and NT person.

Figure 8.3: Last UIC prototype and workshop

Improving communication and messaging

A better communication form between users can be achieved with the balance of visual and sound feedback from the system. The previous cycle's observations show that visual feedback becomes handy when it is not intervening in users' interaction. However, communicating with only visual feedback from the UIC can be hard for both users to interpret each other's UIC feedback. As such, intuitive audio feedback can be integrated to provide more straightforward messages to users. Specifically, patterned beep sounds are used to notify users when the system is about to start and finish recording. Also, voice messages are for delivering messages that visual information can hardly provide such as *no more available recording turns*.

A brief video[1], an audio story, and a photo are provided to help users take away the meaning and purpose of stimming. The textual, visual, and audio information can provide an intuitive format to users. At the beginning of the interaction, the video is shown to them. It explicitly introduces the definition of stimming at the start. It also provides instructions on how the prototype works. Specifically, the video shows how the UIC works, what and how gestures control work. At the end of the video, it informs about their participating space as a rule-free zone. Moreover, to help users reflect on their experience, they can take their collaborated audio story and video or image home as part of memory.

The final UIC prototype is shown in Figure 8.3a. It allows users to attach the device anywhere they prefer.

8.3. Workshop execution and results

[1]https://youtu.be/bxLyNZ5ZJGo

Execution

A workshop was conducted between an autistic, and an NT participant, whom I shall refer to with R and E, respectively (see Figure 8.3b). The purpose of this session was to evaluate the sense-making process, mutual connection, comfort, and the emotional feelings exhibited with real users.

Days before the workshop, the participants received workshops brochures (Appendix A.3, A.4) and were told to bring any objects that they like to stim with or make pleasant sounds. On the day of the workshop, E brought a small musical instrument toy. Also, before the workshop began, they were asked to watch the brief video. After watching the video, I further explained how the prototype works because the video was too fast. Next, they started with getting acquainted with the gesture control, using the provided background relaxing sounds to manipulate, for about fifteen minutes. Subsequently, they started with their recording session, which lasted for about an hour. Finally, a structured group interview was conducted at the end of the workshop.

The answers from the interview at the end of the workshop were categorized into themes that are related to:

- how participants communicate with each other
- how sense-making process is established
- factors that hinder the sense-making process
- how sound affects the emotional state
- the experienced feelings by participants
- how the design features affect comfort, joint attention, synchronization, and mutual connection
- how reflection is established

The classified themes are shown in Figure 8.4. In addition to the answers from the participants, observation notes were also included in the analysis.

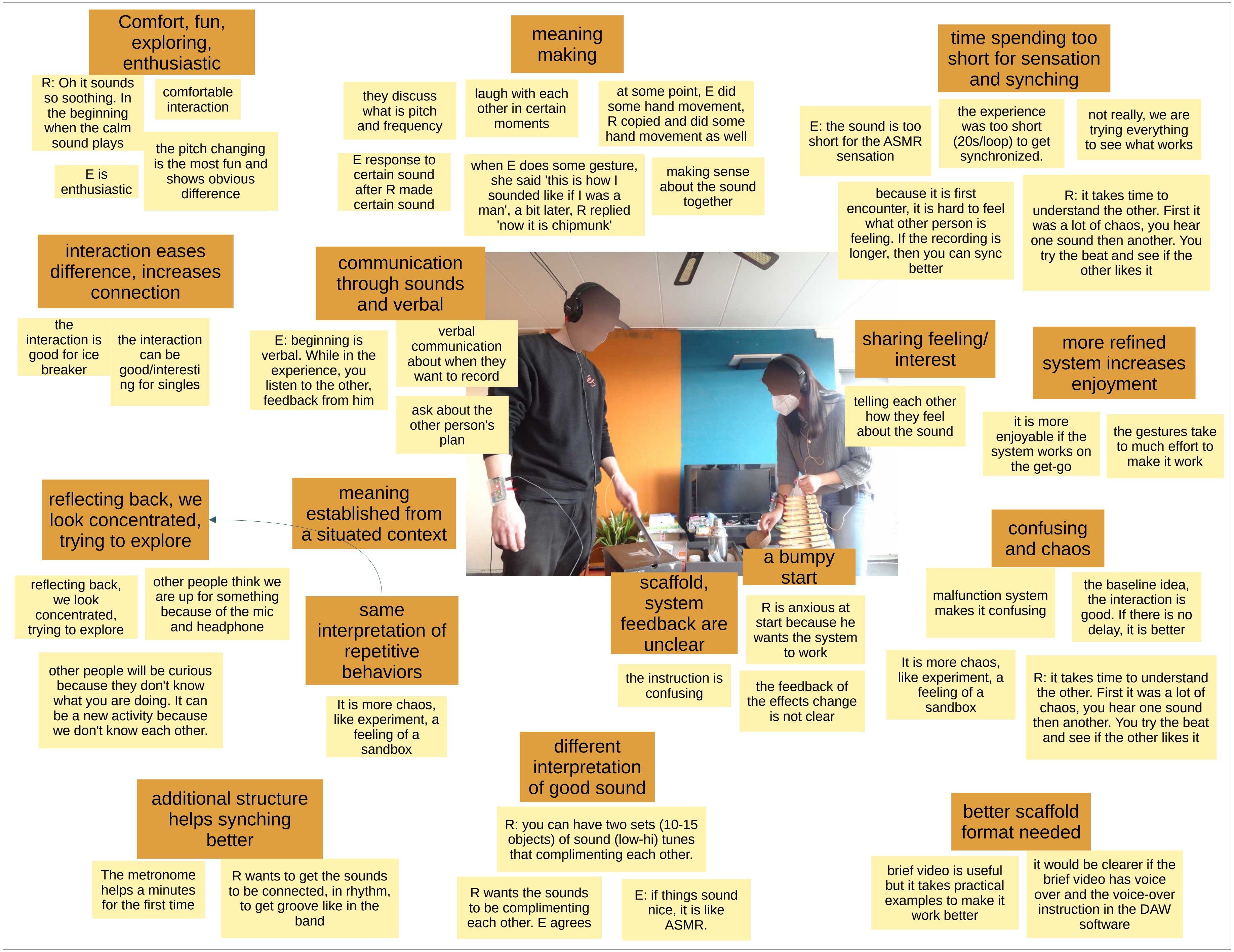


Figure 8.4: Results of categorized themes from eighth workshop between autistic and NT participants

Results

In the same space, the participants continuously conversed with each other about their tasks and shared their thoughts about the created sounds. Many times before recording, they asked each other about the planning of their sound recording process. For example, E asked: *'Which sound are you going to use?'*, then R showed her a ruler, and responded that he would use it as an object to make a sound. In another scenario, they settled on an agreement about specific objects that would produce the sounds they will record. They also exchanged their opinions about a specific sound that they were making in the same shared soundscape. One usually expressed the positive feedback of the sound if he/she liked the sound that the other was making and his/her concern whether a certain sound annoyed the other one. When the sound effects changed, they also exchanged comments on how a certain sound resembles different voice characters to them. They also discussed the differences between the two sound effects, which were the pitch and frequency.

Furthermore, they also communicated through sound in the same soundscape. For example, when one was making a sound, the other one responded right after, or almost the same time with another sound. For example, E shook the rocks-filled glass right after R made two beats with the ruler. These observed behaviors were repeated many times during their interaction.

They were also trying to make sense of two control gestures, namely swipe and swing, during the playback time. During the playback mode, there was a moment after seeing E started to perform the swipe gesture, R while experimenting with his sound simultaneously attempted to replicate her gesture. At that moment, they were collaborating and trying to see how the gesture could change the whole soundscape's effect. The observation shows that at that moment, there was joint attention between them. In particular, they were co-exploring how the recognized gestures control the sound effects. Furthermore, joint concentration was often observed when both were in 'record' mode: their actions were more assertive and less exploratory compared to those in the playback mode. Also, they had more concentrated facial expressions.

There was affordance to stim with and without objects. The initial instruction let participants learn three recognized repetitive gestures and allowed them to use these gestures during their activity. However, they both performed these gestures only once in their activity, because the low accuracy of the gesture recognizer discouraged them to try more. The provided soundscape and objects also encouraged them to make repetitive movements. For example, the autistic participant repetitively pressed a bicycle light button, made beats and slice motions using a ruler, while the NT person repetitively shook a rocks-filled glass.

Both people felt comfortable and fun interacting with each other, due to the provided nature sound, familiar space, and impression. The activity eases their differences, as they remarked: *'it is a good ice breaker'*. Especially for the autistic participant, the provided nature sound brought him comfort and pleasantness. He felt comfortable at the very start when he heard the birds chirping sound. Furthermore, the familiarity of the space also gives him comfort. The NT felt comfort because of her pleasant impression.

This norm-free space gave them room to explore and be creative, but it can create chaos and unstructured interaction. For example, E said: *'it is like a feeling of a sandbox'*. Specifically, they described their experiences as enthusiastic, fun, chaotic, anxious, and confusing. The fun can be observed during their interaction when both of them often were laughing with each other. Chaos and confusion is the mutual feeling that both of them experience in the beginning. Especially, the autistic participant felt anxious from the start because he was trying to see how the system works. Furthermore, the malfunction of the UIC and the unintuitive gestures control instructions (from the video) also caused confusing and unclear guidance. They presumed that a more refined version may alleviate such chaos and confusion.

Both participants perceived the activity as an act of exploring different sounds to find the "good" ones. E said: *We are just trying everything and see what works.* Yet, they had different viewpoints of what good sounds should be. R perceived good sounds when the sounds were complimenting, connecting each other with tunes and rhythms. E agreed to this point and also added that a good sound gives her a sense of ASMR. They both understand the meaning of repetitive sounds as beats. R mentioned that he was trying to make the beats with the bicycle light and see if E liked it. E understood that he was trying to make beats at that time.

Another issue of the prototype is that the brief recording turns hindered their mutual connection. Both of them perceived that the activity was too short due to the recording time (20seconds/turn) to establish a concrete synchronization, especially for their first encounter, even though a metronome helped them synchronize better. Their second recorded sound collaboration with the metronome was more structured and coherent than the earlier one without it.

Regarding reflection, both of them perceived the photograph of them interacting as the moment they were concentrating, trying to explore a new activity.

Summarized results from workshops:

- Sense-making has been established through verbal and sound communication.
- Joint attention was established through co-exploring how gesture control works.
- Repetitive gesture control and provided objects gave them the affordance to stim
- The interaction was comfortable and fun but chaotic and unstructured.
- The interaction eased the individuals' differences
- Same meaning of 'stimming behavior' was established as exploring, experimenting, and making rhythm.
- A more refined system with clearer and more intuitive instructions/scaffolds can boost their enjoyment and improve synchronization.
- Their perspectives on "good" sounds were different.
- Synchronization was not yet established due to the short amount of time spent.
- Reflection meaning was 'the moment of exploring'

8.4. Evaluation and discussion

Through sound production collaboration, the participants were able to make sense together through sound and verbal communication.

In this situated activity, the participants are allowed to explore and create different sounds in the same space together, which renders the meaning of stimming as an act of exploring and experimenting. Specifically, stimming behaviors are seen as tools to help them create sound. Furthermore, as observed, such repetitive-stimming behavior can have a different meaning which is beat making. This perception is similar to the one found in Cycle II, where both M and P described stimming as a beat-maker. The meaning of stimming depends on the intention of users' interaction. As such, the meaning of stimming is temporary and can be changed over time during the interaction.

They have established mutual engagement through collaboration in the shared sound world. Though the prototype was confused to use initially, it was able to guide the participants to create a story together. The situation has put both participants in a soundscape world, where all surrounding sounds are amplified, to give them a different perspective and keep their attention to the creative process.

9

General discussion

This thesis presented a conceptual DivComp-based framework and prototype that addresses the double empathy problem. Specifically, the device called *Stim4Sound* is the proof of concept of the DivComp model. Furthermore, it provides design guidelines for developing an interactive application that encourages stimming in a social, comfortable, and norm-free environment. The need to share mutual understanding in the socially situated activity between NT and autistic people was a key challenge identified in the pursuit of a suitable DivComp concept for stimming. Therefore, I addressed this challenge by designing a system that facilitates a creative and inviting environment and a playground to share mutual understanding through sound and music-making.

Coming up with a suitable design of the concept and demonstrating its effectiveness proved to be very difficult. In particular, it is hard to identify which features of the prototypical device are needed to make the experience as fruitful and comfortable as possible for both parties involved. Thus, to make the design process efficient and productive, I applied the so-called RTD methodology to our design.

The entailment of this methodology is that the design process could be split into four refinement cycles. Concretely, each cycle consists of four important phases: 1) concept requirements; 2) concept development and prototyping; 3) workshop execution and results, and 4) intermediate evaluation and discussion.

A more general and detailed discussion of the work and results is found in Section 9.2. The key contributions are outlined in Section 9.3. Finally, future improvements and research directions are discussed in Section 9.4.

9.1. Summary

In Chapter 4: I showed how the RTD is applied to the design flow of the proof of concept and prototype. The approach applied design tools to iteratively revise concept requirements participatory sense-making between an autistic and NT person.

In Chapter 5: I explored different activities that support the DivComp concepts. The activities were designed based on DivComp and participatory sense-making theories and design skills. After two iterative workshops with a NT person, I have learned that collaborative and exploring types of activity can foster mutual engagement between peers and still gives them space for self-expression.

In Chapter 6: The results from previous workshops were used to build the first concept requirements for the DivComp prototype. Sound Collaboration was proposed as the participatory sense-making activity as it satisfied the concept requirements. To develop the first proof of concept, I applied DivComp theories and used my design skills to build a concrete low-fi DivComp prototype that involves a human mediator. After it got evaluated from three workshops, the results have shown that synchronization and joint attention factors can enhance the sense-making process for the users. The synchronization factor could be achieved by structural design features and users' joint attention could be established by users sharing the same element.

In Chapter 7: The revised concept requirements and design principles for participatory sense-making have led me to develop a digital prototype without a human mediator. Then, the prototype got evaluated in two workshops and three expert interviews. The results have shown that the intermediate prototype needs to present the meaning of stimming more explicitly and design for participants' comfort.

In Chapter 8: A more elaborated prototype was built to demonstrate the revised concept requirements. Specifically, the prototype provides explicit information on stimming, provides comfort for users, includes additional rules and shared elements for users to make sense easier. It got evaluated in a workshop with real users: an autistic and an NT person. The result from this workshop has shown that participants were able to make sense together by non-verbal and verbal communication, have joint attention from the gesture controller, and feel comfortable with one another.

9.2. Discussion and conclusions

At the beginning of the thesis, I put forward the following research question:

- **How to use interactive technology to encourage stimming in a more socially acceptable way?**
- **How can the DivComp roadmap be applied to alleviate the double empathy problem?**

Regarding the first question, the developed DivComp application, Stim4Sound, contributes to a novel perspective of encouraging stimming in a social activity. Sound collaboration activity is situated in a rule-free and co-exploration environment that allows autistic user to freely express their stimming and still preserves comfort for the NT user. By allowing them to explore different movements to create sounds in their own way, the concept provides a norm-free zone that users can express any behavior. This means that any stimming behaviors are accepted and be part of their new norm formulation. Thus, the design feature of a rule-free and co-exploration environment gives autistic people more room to stim without a predefined boundary.

Regarding the second research question, the DivComp framework has been applied in the design of a mediator tool for an autistic and a NT person to establish a new meaning of stimming within a sound collaboration activity. As such, it directly addresses the double empathy problem. As discussed at the beginning of the thesis, the double empathy problem refers to the phenomenon where NT people are unable to understand autistic people's perception of the world and vice versa. Specifically, due to the difference in their cognitive structures, a mutual understanding on the cognitive level is rather hard to achieve. However, a new understanding can be created when these two neurotypes join to share an experience together. As such, the developed DivComp concept provides the participatory sense-making activity, where an autistic and NT person together share and makes sense in a specific context. The activity allows them to understand one another by expressing themselves through reciprocal body interactions. Thus, as part of the philosophy of the DivComp, Stim4Sound does not try to impose certain meaning to its users, of stimming in particular, in their interactive activity. Instead, it provides an alternative shared platform that allows them to express repetitive behaviors and communicate with each other through these behaviors and body language more explicitly. By constantly participating and exchanging their expressions, the platform allows them to form a new understanding and perception of the rising meaning, in particular stimming, in this sound collaboration activity.

9.3. Contributions

This thesis made the following important contributions:

- **A DivComp prototype**
  The prototype has fostered the a new meaning-making process of stimming between an autistic and NT users in a sound collaboration activity.
- **An extension to the DivComp model**
  The applied model provides detailed investigation and contributes more detailed concepts to the current DivComp. Furthermore, it also provides design guidelines to develop a stimming encouragement device.
- **Workshops with NT and autistic people** In particular, the results are used to evaluate the DivComp theories and design principles of participatory sense-making.

In the subsections to follow, I discuss each contribution individually in more detail.

A DivComp prototype

The proposed concept, which is the first proof of concept of its kind, has successfully implemented a subset of the goals and theories, provided by the novel Diversity Computing framework. Specifically, observations and evaluations from the workshops, as presented in Chapters 5, 6, 7 and 8, have demonstrated that the theoretical framework can be applied for the design of systems that accommodate meaning-making in the context of stimming behaviors within a situated activity, suitable for autistic and neurotypical people. Such a socially situated activity envisions interaction between autistic and NT people in spaces such as offices, schools, universities, etc. Furthermore, more public space such as a museum, or an exhibition can help to raise awareness and provide explicit knowledge about stimming to a wider audience. Moreover, the applied DivComp framework can be transferred to different neurodiverse background groups.

New meaning of stimming

Moreover, a new meaning of stimming can be achieved, provided that it is placed in situated activity. The developed prototype provides the sound collaboration activity, which is a fun and creative activity, in order to bring autistic and NT people together. As observed, the meaning of stimming was considered as an act of exploration, experimentation, and making beats. As such, stimming turns out to be closely related to this sound-making activity. As the activity brings a positive experience for both users, especially for NT user, the impression of stimming will be more enjoyable. Thus, by participating in a pleasant and comfortable experience together with the autistic person, the NT user may develop less negative judgment toward the behavior. This may contribute to the point of alleviating stigmatization of stimming by an NT person in a socially situated activity due to the new share experience of a pleasant activity. Nevertheless, this was only the first exploration that shows the opportunity of designing for stimming as a positive phenomenon. Further studies should be made to better understand the NT perception of stimming in different activities and how such perception can influence the stigmatization against autistic people.

New normative way of communication

Furthermore, the proposed concept supports a new normative way of communication formed by NT and autistic people. The experience doesn't impose an NT-like norm, where people are expected to interact with maintaining eye contact and use facial expressions. Instead, in this situated activity, the connection is made with language, body, auditory, and visual cues. This is referred to the general concept of *embodied agents, embedded in contextual settings* as discussed in the DivComp theories by (Fletcher-Watson et al., 2018).

When people can communicate in the same language, verbal communication is a semantic way to exchange communication. In other cases where verbal communication is not compulsory, people can communicate through the body. Regarding Stim4Sound, the situated activity allows people to find and define their own way to express and exchange their body movements in the shared space. They communicate with direct bodily expressions, such as body language, gestures, and behaviors. Being in the same place allows them to observe and respond to each other's actions and movements.

Though both facial and body communication is visual communication, communication with bodily expressions is less direct compared to facial communication. Since facial communication is not preferred by autistic people (De Jaegher, 2013), this new way of communication is eventually beneficial for autistic people.

As observed, a fairly new communication style (or communication ‘norm’) through sound was formed by an NT person and an autistic person in the sound collaboration activity. Specifically, they bring their sound creation into the shared soundscape and respond to each other’s samples of sounds. By exchanging their sounds in the same soundscape, they communicate through sound.

Besides sound-sharing, the visual feedback from the LEDs from the system delivers predefined commands and messages between people. This modality is necessary because: (1) the people need to have clear and structured messages to have their interaction in-synched; (2) it is a non-intrusive form of message delivery because it does not directly interfere with their sound communication.

An extension to the DivComp model

The DivComp model has been applied in a participatory sense-making activity, which is sound collaboration, between two diverse neurotypical groups: autistic and NT people. The applied model contributes additional and more detailed features into the current DivComp model and provides guidelines to design an empathetic device to encourage people to stim. Based on the findings from four iterative cycles, the applied DivComp model provides guidance to design new DivComp devices and systems in the context of encouraging stimming. In summary, the following key features and requirements have been received positively by the users that participated in the workshops and can be integrated into future designs:

- Creates a situated activity that allows collaboration activity. This feature encourages people to work together and form a mutual connection through the same experience.
- Creates a situated activity that provides a free place to express and affordance to stim. This element creates a non-normative space that encourages people to stim freely.
- Creates a situated activity that does not require expertise skill to express oneself. This requirement makes people feel invited to participate in the activity and to bring their skill set to the common ground of forming new rules.
- Creates a situated activity that provides comfort, safety, and privacy. This element assures the ethical factors of technology should provide users such rights respectively referring as calmness, human welfare, and privacy values that are proposed by Friedman et al. (2013).
- Provides a semi-synchronous interaction to adapt both of how autistic and NT people make sense. This point is a compromise between asynchronous and synchronous interactions. For example, interactive features for asynchronous mutual body coordination are appreciated by autistic people, and synchronous shared space can enhance a structural flow of interaction that boosts the sense-making process between users.
- Creates a platform where autistic and NT people can communicate in a non-verbal way such as visual or auditory contacts, and in a less normative way such as facial expression, eye contact,

and body coordination interaction. These points are communicative features of how autistic people make sense of the world that can be integrated into the design.

- Provides design features that enable users to self-reflect on the participated activity. This requirement can help them develop and sustain their newly established meaning from the activity, in particular stimming, in their perception.
- Provides design features that facilitate joint attention, mutual connection, and synchronization to increase users' mutual engagement. Joint attention can be created by providing a shared item between users' interactions. A mutual connection can be increased by longer time-spending activity. Structural scaffolding elements can enhance synchronization between them.
- Provides initial scaffolds to introduce common ground rules and meanings for users while still provide enough freedom for them to express themselves.
- Provides design features that allow users to have control over their (sensory) preference and comfort.

To demonstrate all of these concept features into a more concrete design, the final developed prototype, Stim4Sound, has shown how the situated activity of sound collaboration entails all of these factors. Figure 9.1 summarizes all of the design features of the final prototype. The required design features were revised and formulated after all findings in the workshops and expert interviews. Eventually, the prototype was able to allow people to directly control their environment, to function without a mediator person, and to provide a non-intrusive form of feedback. The compatible design features met the required criteria of providing comfort, joint attention, structured synchronization, self-reflection (from traces and memory), encouraging people to make repetitive movements and to be free to express themselves.

Workshops with NT and autistic people

The results from the workshops helped gather empirical observations and user feedback concerning the design principles of sense-making interactive systems for NT and autistic people. Specifically, Stim4Sound implemented some aspects of participatory sense-making theories proposed by (De Jaegher, 2013; Hummels and van Dijk, 2015).

In particular, Stim4Sound has taken into account the different ways autistic people make sense of the world compared to NT people. According to Jaegher et al., autistic people prefer an interaction that involves less social cues communication from facial expression and eye contact. By allowing them to explore different sounds around them, Stim4Sound lets users be in their own creative flow that does not require much facial communication with the other. Also, autistic people usually adopt asynchronous interaction that does not encourage them to coordinate synchronously with the other. Stim4Sound provides asynchronous body coordination because users do not need to synchronize their body movements and gestures together to interact with each other. Furthermore, users also do not need to record the sounds at the same time. This feature contributes to the point that users do not have to be synchronized in time.

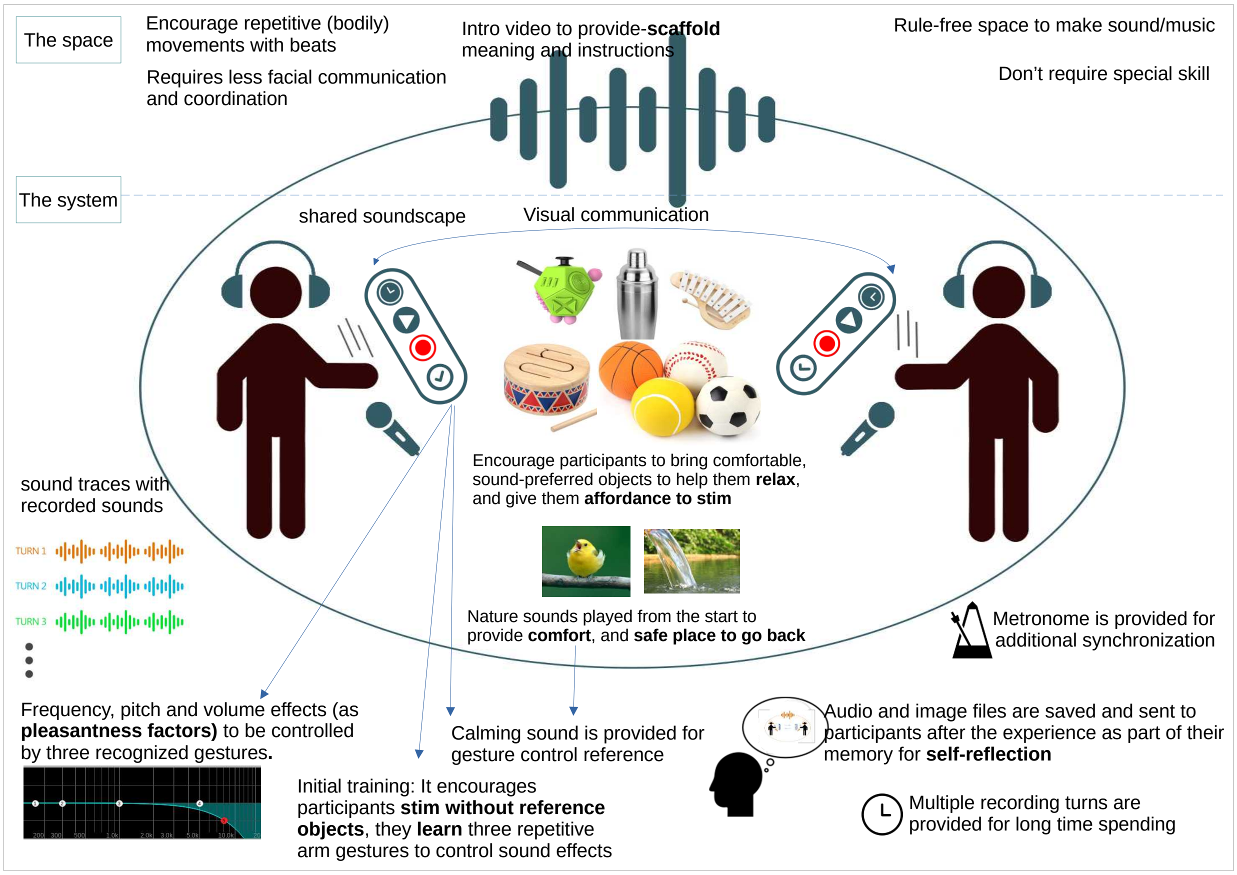


Figure 9.1: All design requirements

Also, Stim4Sound has implemented three design principles that were proposed by (Hummels and van Dijk, 2015). According to Hummels and Van Dijk, a situated activity is the social situation that gives a free space that allows people to establish their own ground of norm and communication. Deriving from this principle, Stim4Sound provides the sound collaboration activity that allows users to freely explore different sounds, made by their movements and behaviors. From observations, users were able to communicate through sound and visual cues from the shared space. Furthermore, the design principle scaffold provides a sense of direction and suggestions in users' mutual interaction. By providing a metronome, Stim4Sound helps users to structure their sounds making the process better. In addition, the initial brief video gives them instructions on how to control the effect with their gestures. Moreover, the design principle trace acts as a representation of evidence that users can use to reflect on and to react to. Stim4Sound provides users' creations of their recorded sounds and gestures feedback. These two elements serve as building blocks of their story collaboration.

9.4. Limitations and future directions

While this work makes compelling and important contributions in the field of Human-Computer Interaction (HCI), there are some limitations in the approach, the target group, and concept, which are summarized below, along with future improvements.

First, the researcher - I was present in all of the conducted workshops during the interaction between the users, which may have influenced their behaviors and communication.

Additionally, the interactive system still sometimes requires human moderation. A more polished and autonomous version of the system would fully allow its users to participate in the experience by themselves. This point might result in different communication manners and levels of engagement.

Also, although confusion is unavoidable in learning new activities, additional refinements and the introduction of intuitive scaffolds regarding the user interface can increase clarity and pleasantness for the users. Also, although the chosen target group is young adults, ranging from 23 to 30-year-old. However, the interaction between and cross middle-aged and older adults might produce different results.

Furthermore, due to extraordinary circumstances, the majority of workshops were carried out with NT people. As such, the lack of a more diverse group that includes autistic people might have biased the results more towards the NT perspective. Future research should take note of that and incorporate more workshops involving autistic people.

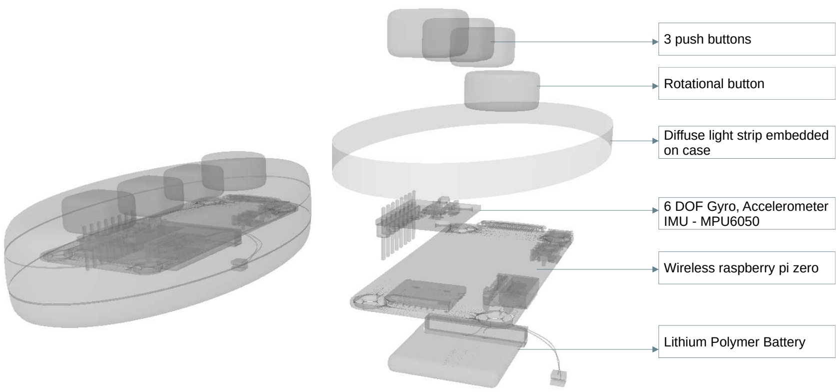
From the evaluations, users who have more musical or sound experience have a different interpretation of music and sound production. The process of sense-making in the same soundscape can be influenced by different sound expertise from users.

In this research, people participated in the activity where they were in the same physical place. However, a more asynchronous interaction in space and time may produce different results. Space and time are important variables to boost meaning-making, thus, future research needs to investigate the variations of these factors but still make sure that meaning-making is present in the interaction between autistic and NT people.

From a technical point of view, the resulting prototype devices are in a very early stage of development, and the software contains many bugs. As such, future work should try to address these issues and make the software architecture more stable. Coming up with a better hardware design is also something that should be considered. Finally, the client-server architecture can be improved further, by accommodating more UIC connections. In particular, the state machine can benefit from more data synchronization, and the communication with the DAW needs further refinement.

To elaborate further on technical specifics, additional design parameters can be explored in future designs. For example, further tuning of the hyperparameters of the NN model used in the gesture recognizer can greatly benefit its accuracy. Due to time constraints, this matter was barely given much attention, which was reflected in the overall performance during the workshop. Also, while the training data used to train the model was plentiful, the data was not diverse enough. Specifically, the gestures were collected from a limited number of people (at most 2). As such, by including more

(a) The UIC prototype



(b) Construction of the User Interface Controller (UIC)

Figure 9.2: The envision of the UIC prototype and its structure

small batches of training examples from more people with diverse physiological attributes, it would beneficial to diversify the data set in the future.

Finally, the design features of self-reflection should also be further revised to help people perceive and reflect better in a more explicit way about the meaning and purpose of stimming.

Taking all of these future considerations into account, an envisioned, fully developed prototype that meets the concept requirements is shown in Figure 9.2a.

9.5. Final words

The work presented in this thesis encapsulates new and fundamental ideas that complement the DivComp framework. Nevertheless, many interesting and fundamental issues were excluded from this thesis. As such, I hope that this work will inspire motivated researchers to continue this pursuit, and that it will stimulate further research and development into such a fresh and important framework.

Bibliography

Alonso, M. B., Keyson, D. V., and Hummels, C. C. M. (2008). Squeeze, rock, and roll; can tangible interaction with affective products support stress reduction? In *Proceedings of the 2nd international conference on Tangible and embedded interaction*, TEI '08, pages 105–108, New York, NY, USA. Association for Computing Machinery.

Attwood, T. (1997). *Asperger's syndrome: A guide for parents and professionals*. Jessica Kingsley Publishers.

Bagatell, N. (2010). From Cure to Community: Transforming Notions of Autism. *Ethos*, 38(1):33–55.

Ballaban-Gil, K., Rapin, I., Tuchman, R., and Shinnar, S. (1996). Longitudinal examination of the behavioral, language, and social changes in a population of adolescents and young adults with autistic disorder. *Pediatric Neurology*, 15(3):217–223.

Boyd, L. (2019). Designing Sensory-Inclusive Virtual Play Spaces for Children. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*, IDC '19, pages 446–451, New York, NY, USA. Association for Computing Machinery.

Byrne, M. (2019). Increasing Engagement and Academic Performance of Children with Autism SpectrumDisorder and Attention Difficulties: Do Fidget Spinners Help? *Graduate Theses and Dissertations*.

Carles, J. L., Barrio, I. L., and de Lucio, J. V. (1999). Sound influence on landscape values. *Landscape and Urban Planning*, 43(4):191–200.

CBS (2019). StatLine - Perceived health, care use and lifestyle in children up to 12 years old.

CDC (2020). Data & Statistics on Autism Spectrum Disorder | CDC.

Cottrell, P., Grow, A., and Isbister, K. (2018). Soft-bodied Fidget Toys: A Materials Exploration. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 42–48, Stockholm Sweden. ACM.

Courage, C., Baxter, K., and Caine, K. (2015). *Understanding your users: a practical guide to user research methods*. Elsevier, Morgan Kaufmann, Amsterdam ; Boston, second edition edition. OCLC: ocn918928845.

De Jaegher, H. (2013). Embodiment and sense-making in autism. *Frontiers in Integrative Neuroscience*, 7.

De Jaegher, H. and Di Paolo, E. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4):485–507.

Farley, J., Risko, E., and Kingstone, A. (2013). Everyday attention and lecture retention: the effects of time, fidgeting, and mind wandering. *Frontiers in Psychology*, 4. Publisher: Frontiers.

Feltham, F., Loke, L., van den Hoven, E., Hannam, J., and Bongers, B. (2013). The slow floor: increasing creative agency while walking on an interactive surface. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction - TEI '14*, pages 105–112, Munich, Germany. ACM Press.

Fletcher-Watson, S., De Jaegher, H., van Dijk, J., Frauenberger, C., Magnée, M., and Ye, J. (2018). Diversity computing. *Interactions*, 25(5):28–33.

Fletcher-Watson, S. and Happé, F. (2019). *Autism: A New Introduction to Psychological Theory and Current Debate*. Routledge. Google-Books-ID: ttKFDwAAQBAJ.

Frayling, C. (1993). Research in art and design.

Friedman, B., Kahn, P. H., Borning, A., and Huldtgren, A. (2013). Value Sensitive Design and Information Systems. In Doorn, N., Schuurbiers, D., van de Poel, I., and Gorman, M. E., editors, *Early engagement and new technologies: Opening up the laboratory*, Philosophy of Engineering and Technology, pages 55–95. Springer Netherlands, Dordrecht.

Frith, U. (1991). *Autism and Asperger Syndrome*. Cambridge University Press. Google-Books-ID: HoRX8s8V8WYC.

Graziano, P. A., Garcia, A. M., and Landis, T. D. (2020). To Fidget or Not to Fidget, That Is the Question: A Systematic Classroom Evaluation of Fidget Spinners Among Young Children With ADHD. *Journal of Attention Disorders*, 24(1):163–171.

Hobye, M. and Löwgren, J. (2011). Designing for Engaging Experience in Embodied Interaction. page 18.

Hornecker, E. and Buur, J. (2006). Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction. page 10.

Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M.-C., and Mandy, W. (2017). "Putting on My Best Normal": Social Camouflaging in Adults with Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, 47(8):2519–2534.

Hummels, C. and van Dijk, J. (2015). Seven Principles to Design for Embodied Sensemaking. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '14*, pages 21–28, Stanford, California, USA. ACM Press.

Johnson, K. T. and Picard, R. W. (2017). SPRING: Customizable, Motivation-Driven Technology for Children with Autism or Neurodevelopmental Differences. In *Proceedings of the 2017 Conference on Interaction Design and Children*, pages 149–158, Stanford California USA. ACM.

Kanner, L. and others (1943). Autistic disturbances of affective contact. *Nervous child*, 2(3):217–250. Publisher: New York.

Kapp, S. K., Steward, R., Crane, L., Elliott, D., Elphick, C., Pellicano, E., and Russell, G. (2019). 'People should be allowed to do what they like': Autistic adults' views and experiences of stimming. *Autism*, 23(7):1782–1792.

Karlesky, M. and Isbister, K. (2016). Understanding Fidget Widgets: Exploring the Design Space of Embodied Self-Regulation. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction - NordiCHI '16*, pages 1–10, Gothenburg, Sweden. ACM Press.

Lai, M.-C., Lombardo, M. V., Chakrabarti, B., and Baron-Cohen, S. (2013). Subgrouping the Autism "Spectrum": Reflections on DSM-5. *PLoS Biology*, 11(4):e1001544.

Liang, R.-H., Yu, B., Xue, M., Hu, J., and Feijs, L. M. G. (2018). BioFidget: Biofeedback for Respiration Training Using an Augmented Fidget Spinner. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pages 1–12, Montreal QC, Canada. ACM Press.

Lilley, R. (2018). What's in a flap? The curious history of autism and hand stereotypies. *Manuscript submitted for publication.*

Lord, C., Cook, E. H., Leventhal, B. L., and Amaral, D. G. (2000). Autism spectrum disorders. *Neuron*, 28(2):355–363. Publisher: Elsevier.

Lovaas, I., Newsom, C., and Hickman, C. (1987). Self-Stimulatory Behavior and Perceptual Reinforcement. *Journal of Applied Behavior Analysis*, 20(1):45–68. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1901/jaba.1987.20-45.

Maes, P.-J., Lorenzoni, V., Moens, B., Six, J., Bressan, F., Schepers, I., and Leman, M. (2018). Embodied, Participatory Sense-Making in Digitally-Augmented Music Practices: Theoretical Principles and the Artistic Case "SoundBikes". *Critical Arts*, 32(3):77–94.

Matson, J. L. and Goldin, R. L. (2013). Comorbidity and autism: Trends, topics and future directions. *Research in Autism Spectrum Disorders*, 7(10):1228–1233.

Mentis, H. M., Laaksolahti, J., and Höök, K. (2014). My Self and You: Tension in Bodily Sharing of Experience. *ACM Transactions on Computer-Human Interaction*, 21(4):1–26.

Miller, L. J., Anzalone, M. E., Lane, S. J., Cermak, S. A., and Osten, E. T. (2007). Concept Evolution in Sensory Integration: A Proposed Nosology for Diagnosis. *American Journal of Occupational Therapy*, 61(2):135–140.

Milton, D. E. (2012). On the ontological status of autism: the 'double empathy problem'. *Disability & Society*, 27(6):883–887.

Morishima, T., Restaino, R. M., Walsh, L. K., Kanaley, J. A., Fadel, P. J., and Padilla, J. (2016). Prolonged sitting-induced leg endothelial dysfunction is prevented by fidgeting. *American Journal of Physiology-Heart and Circulatory Physiology*, 311(1):H177–H182. Publisher: American Physiological Society.

Nisha, B. and Soruparani, D. S. M. (2018). Comparison of a Pleasant and Unpleasant Sound. 4(2):6.

Nolan, J., McBride, M., and McBride, M. (2015). Embodied Semiosis: Autistic 'Stimming' as Sensory Praxis. In *International Handbook of Semiotics*, pages 1069–1078. Springer Netherlands, Dordrecht.

NVA (2014). NVA - Prevalentiecijfers over autisme.

Paulus, P. B., Korde, R. M., Dickson, J. J., Carmeli, A., and Cohen-Meitar, R. (2015). Asynchronous Brainstorming in an Industrial Setting: Exploratory Studies. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 57(6):1076–1094.

Porayska-Pomsta, K., Frauenberger, C., Pain, H., Rajendran, G., Smith, T., Menzies, R., Foster, M. E., Alcorn, A., Wass, S., Bernadini, S., Avramides, K., Keay-Bright, W., Chen, J., Waller, A., Guldberg, K., Good, J., and Lemon, O. (2012). Developing technology for autism: an interdisciplinary approach. *Personal and Ubiquitous Computing*, 16(2):117–127.

Stappers, P. J. and Giaccardi, E. (2017). *Research through Design.*

Stienstra, M. (2003). Is every kid having fun?: A gender approach to interactive toys design. *technologies for Interactive Digital Storytelling and Entertainment *. ISBN: 9783816762768 Publisher: Stollfuß Medien.

Wensveen, S. A. G., Djajadiningrat, J. P., and Overbeeke, C. J. (2004). Interaction Frogger: A Design Framework to Couple Action and Function through Feedback and Feedforward. page 8.

Zolyomi, A., Ross, A. S., Bhattacharya, A., Milne, L., and Munson, S. (2017). Value Sensitive Design for Neurodiverse Teams in Higher Education. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 353–354, Baltimore Maryland USA. ACM.

Özcan, E. and Schifferstein, H. N. J. (2014). *The Effect of (un)pleasant Sounds on the Visual and Overall Pleasantness of Products.*

A

Appendix A: Forms

This appendix consists of ethical approval letter, consent form, Covid-19 brochure, and workshop brochure.

A.1. Letter of approval from ethical committee

UNIVERSITY OF TWENTE.

Mr./Mrs. T. Nguyen

FROM	DATE	PAGE
M.C. Kamp T 053-4892547 m.c.kamp@utwente.nl	14 December 2020 OUR REFERENCE ET/A.20.19210	1 of 1
SUBJECT LETTER OF APPROVAL	YOUR REFERENCE	

Dear Mr./Mrs. Nguyen,

The Natural Sciences and Engineering Sciences Ethics committee has reviewed your submission for "Design an interactive application for autistic and neurotypical people to stim together" and based on the submitted material has formulated a positive advice for the dean.

On the basis of this advice I approve your application and leave the responsible execution of this project in your hands trusting that you will conduct this research in a manner worthy of the University of Twente.

The request has been registered under **reference number 2020.40**

I wish you good luck with your research.

Yours sincerely,

Prof.dr.ir. H.F.J.M. Koopman
Dean faculty of Engineering Technology
University of Twente

The approval given for your research project is valid, taken into account the restrictions that you have to comply with the current RESTRICTIONS ON SOCIAL AND PHYSICAL INTERACTION set by the government regarding the COVID19 outbreak. Your study intends to make use of methods requiring social and physical interaction. This poses risks for both participants and researchers, which have to be taken into account. This may imply that you have to find alternative ways to collect data or to delay the start of your study until the restrictions have been adjusted or lifted. If adjustments lead to substantive changes in the design of your study (excluded: digital/online means to get in contact with your participants), I advise you to send the changes to the Ethics Committee stating your reference number.
Please consult the standing guidelines of the UT and national authorities on research and educational activities www.utwente.nl/corona

P.O. Box 217, 7500 AE
Enschede
The Netherlands
www.utwente.nl

University of Twente (UT) is registered at the Dutch Chamber of Commerce under nr. 501305360000

A.2. Consent form

Informed Consent for standard research

'I hereby declare that I have been informed in a manner which is clear to me about the nature and method of the research as described in the aforementioned information brochure 'Making Music Workshop'. My questions have been answered to my satisfaction. I agree with my own free will to participate in this research. I reserve the right to withdraw this consent without the need to give any reason and I am aware that I may withdraw from the experiment at any time. I am aware that my recorded data are used for the research materials.

Hereby, I would like to give you my consent:

I agree to be video recorded : Yes/No
I agree to be voice recorded : Yes/No
I agree to be photographed: Yes/No
I agree to be requoted with text: Yes/No
I would like to be mentioned/acknowledged by full name/ initials/ anonymity

In case that one of the participants in this research or a researcher contracts COVID-19, we will keep the contact data of all participants in this experiment. If a participant or researcher tests positive for COVID-19 and the research team is contacted by the GGD (Gemeentelijke Gezondheidsdienst or city health service) for contact research or otherwise made aware, you will be send a message. The message will state that someone in the experiment has contracted COVID-19 and will include a request for you to contact your general practitioner or local GGD. No names or identifying information will be shared. Your contact data will be stored for up to a month after the experiments have ended and will be destroyed after the time period is up. Your contact data is stored separately from your research data and cannot be connected to your research data.
Declaration of consent (please tick the checkbox if you consent)

I agree for my contact data to be stored as described above :Yes/No

If my research results are to be used in scientific publications or made public in any other manner, then they will be made completely anonymous. My personal data will not be disclosed to third parties without my express permission. If I request further information about the research, now or in the future, I may contact Thu Nguyen, telephone: +31 (0)6 2233 4596; email: h.t.nguyen@student.utwente.nl

If you have any complaints about this research, please direct them to the Secretary of the Natural Sciences and Engineering Sciences Ethics Committee at the University of Twente, P.O. Box 217, 7500 AE Enschede (NL), telephone: +31 (0)53 489 2547; email: m.c.kamp@utwente.nl). Signed in duplicate:

Name subject *Signature*

I have provided explanatory notes about the research. I declare myself willing to answer to the best of my ability any questions which may still arise about the research.'

Thu Nguyen
Name researcher *Signature*

A.3. Covid-19 brochure

Covid-19 Brochure Amendment

Dear participant,

Your health and safety is our main concern, in this addition we would like to explain the changes we made to the experiment set-up and procedure to protect you and the researcher during the pandemic.

- You will be asked about your health before the start of the experiment. If you have
- symptoms of a cold you will not be able to participate.
- If you have any symptoms of a cold, please contact Thu Nguyen at h.t.nguyen@student.utwente.nl to cancel your appointment.
- Doors are opened for you so you don't have to touch door handles.
- Participants wash their hands before and after the experiment.
- The researcher washes their hands before and after every experiment.
- Participants are asked to leave their contact information separately from the research data, so that they may be contacted in the case that any of the participants or researchers contract corona. The contact information will be deleted a month after the experiment has ended.
- You will be asked for permission in a separate consent form.
- If you test positive for covid-19 after the experiment, please contact Thu Nguyen at h.t.nguyen@student.utwente.nl.
- Physical contact between the researcher and participant during the experiment is limited to correctly placing the arm of the participant in the experiment set-up.
- The participants and researcher are separated by plastic screens.
- The set-up is disinfected, after each participant and at the start and end of each workshop.
- The buttons and wristband straps are covered by cling foil, which is replaced after every participant.
- Headphones and microphones placed directly on the skin will be disinfected for every participant.
- A max of 3 people are in the lab at a given time.

We would like to remind you of your right to withdraw at any time without giving a reason. If you do not find these precautions sufficient, feel uncomfortable or are in a at risk group reconsider your participation.

If you have any questions or concerns, please contact Thu Nguyen at h.t.nguyen@student.utwente.nl. If you want independent advice about participating in this research, or if you want to submit a complaint.

You can contact Maria C. Kamp, secretary of the Ethics Committee (tel. 053-489 2547, email: m.c.kamp@utwente.nl) if you have questions or complaints regarding the research.

Yours sincerely,

Thu Nguyen

A.4. Workshop brochure

UNIVERSITY OF TWENTE.

Enschede, 14th January 2021

Information brochure Department Engineering Technology

Dear reader,

In this letter, we would like to inform you about the research you have applied to participate in. The experiment will take place at Zilverling building, University of Twente. In the proposed research, entitled "Design an interactive application for autistic and neurotypical people to make movement together", where an autistic and neurotypical people make music together by interacting with physical objects. The aim of the research is to establish whether there is mutual engagement of when autistic and neurotypical people share an activity together. The research could provide knowledge of how these two neurotypes understand each other in a controlled condition. It helps us to find what factors in interaction bridge the differences between them.

In each session, there will be a workshop and an interview. The location and time of each session will be sent to your email address in advance.

First, a workshop of trying out different objects or electronic accessories to interact is performed. Therefore, if you are a neurotypical person, your partner will be an autistic person and vice versa. If you don't feel comfortable interacting with another neurotype person or vice versa, you may not participate in the research. An interview will be conducted at the end of the workshop. The goal of the interview is to get the insights from the workshop that can relate to an analysis and an evaluation of the interaction.

Furthermore, you can decide to stop at any point in the course of the experiment without this having any consequences for yourself and without giving any reasons. In addition, you can still decide at the end of the research and up to 24 hours thereafter, that your data may not be included in the research after all. Other relevant aspects are that your data will be handled in a confidential manner, the anonymity of your data is guaranteed and will never be disclosed to third parties without your permission.

You should know that the experiment lasts for a maximum of 1.5 consecutive hours and the participation is voluntary. Most of the subjects participating in similar workshops find it very fun and interesting. You are introduced to a different type of research than usual and you can even watch your own interaction in real time. At the end of the entire research, you may, if you so wish, be informed about the results obtained by means of a debriefing.

Yours sincerely,

Thu Nguyen

Dr. Jelle van Dijk
Supervisor
Research group Human Centred Design
Faculty of Engineering Technology
University of Twente
Email: jelle.vandijk@utwente.nl

Thu Nguyen
Research conductor
Department Interaction Technology
Faculty Electrical Engineering, Mathematics and Computer Science
University of Twente
Email: h.t.nguyen@student.utwente.nl

B

Appendix B: Interview Questions

This appendix includes the lists of workshop interview questions

B.1. Interviews questions in 7 workshops

These questions were the prepared questions before the semi-structured interviews were conducted. Thus, additional unstructured questions asked on the spot are not recorded here.

Ws	Interview questions
#1 and #2	What is your most favorite activity?, and why? Which activity did you feel most engaged with me / with the tasks?, and why? Did you understand what we / I was trying to do at that moment? Was there a point you don’t understand or see what is going on? Can you describe this activity with 5 words? How did you interpret X? What meaning was X to you? What element make you understand X? How can we keep it interesting? What was wrong at that time? How do we improve our mutual understanding?
#3, #4, #5	What is your general opinion about the experience? Without verbal communication, how did you guys communicate? How did you response to him (me) with sound in that moment? Do you understand what you guys (we) were doing at that moment? And why? Do you have any feeling when you look at these videos? And why? Do you have any feeling when you interact with these objects? And why? Do you have any feeling when you listen to the sound effects? And why? How do you guys (we) understand each other at this time X? What do you think about the improved version of the system? Do you find the metronome helpful? At what point do you think that you guys (we) were synchronized? How do you think you guys (we) can synchronize better?
#6	What is your general opinion about the experience? Is the feedback from the system clear for you? At that moment, do you know why the third LED light blink? How did this reverb sound effect make you feel? Did you hear the difference in sound at any point?

#8

What is your general point of view about this workshop?

In the scale of 1 to 7, how much do you rate this workshop? Why so?

In the scale of 1 to 7 how much do you feel relax and comfortable in this workshop? Why so?

In the scale of 1 to 7 how much do you feel excited and enjoyable in this workshop? Why so?

What are other feelings, moods, emotions and sensations that you feel when you interact with the system and each other?

In the scale of 1 to 7 how much do you think that you understand each other ?

What were the cues that you use to communicate? And how did you response?

Are there any moments that both of you were in the same page?

How do you communicate with each other (without words)?

How much can you read what the other person's thinking or feeling (e.g. i like this sounds, i don't like this...)?

When one person makes certain repetitive movement (both in the beginning and recording time), do you understand what that person is trying to do?

Were there any moments that you don't understand each other? Can you please explain?

(Show a clip of themselves) What do you think about this?

This is what you guys look like for observers, what do you think how they will react when they see this?