Move4Music: An Adaptive Sound System To Prompt Connection During Neurodiverse Dyadic Interactions

Lara Oral,

Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente

MSc Interaction Technology

Master's Thesis

Supervisors: dr. Jelle van Dijk, dr. ir. Dennis Reidsma

Date: 14/08/2023

Acknowledgements

This thesis process would not have been the same without the various types of support I received from the people in my life. First and foremost, I would like to thank my supervisors, dr. Jelle van Dijk and dr. ir. Dennis Reidsma for their advice and recommendations. Especially Jelle's course on 'Embodied Interaction' and his perspectives, on how technology is to approach the topic of improving neurodiverse social interactions, was what led me to my focus area in the first place. It was a great pleasure to have many interesting discussions with him throughout the last few months and to get more and more comfortable with the complex academic concepts involved in the project. To feel supported, both in terms of being given space to find my own way and voice in my efforts, and in terms of being given the necessary guidance. I felt better about whatever concerns or questions I had after each of our meetings and he, for sure, made things feel much smoother and much more comfortable for me throughout.

Yet another person that indirectly influenced my decision, to work on such a topic that had to do with the sound modality, was Rik Nieuwdorp. From him, I followed the course 'Storytelling Through Sound', after which I felt I had the adequate interest and perspectives to take upon this design challenge. His enthusiasm and external consultation, especially regarding the field and approaches of adaptive music, were very valuable for me.

Thanks to Purna Bishas and Yifan Cheng, who comprised the small research group that made this process less lonely and again, more supportive, whether in a more practical or an emotional sense. Thanks to my dear housemate, Shari Wohlenberg, who not only was there for me as a good friend through my ups and downs, but also shared with me her knowledge and insights of the music therapy field. She took the time to have conversations with me on different aspects and to give me her opinions on my sound design or procedural choices. I owe the same kind of friendship, and my first impressions of the music therapy field and practice also to Janny Amelung, so thank you. To Laure Tolsma, I owe my continuously broadened perspectives and increased desire to improve my knowledge and stance on disability – also in relation to technology and design. With Gemma Ciabattoni as well, I had several inspiring exchanges, especially at the start of my thesis when I was getting into the topic.

Big thanks, also to one of the developers of MI.MU Gloves Limited[1], Adam Stark, who was kind enough to give me free access to the Glover software for research purposes. Additionally, thanks to the MiMU Discord community, who inspired me at the beginning of my efforts.

Thanks to my dear participants, and specifically also Kinderdagcentrum De Walnoot[2]. To our enthusiastic contact person there, Jessica Copier, a wonderful teacher that helped us in better understanding and engaging with her students during the workshops.

Lastly, but most importantly, thanks to all my friends and family for being there for me, for checking in, for encouraging me, for making sure that I also made fun memories outside of thesis-writing during this process.

This would not have been possible without you all.

[1] For the company website, see https://mimugloves.com/
[2] For the school/organization website, see https://www.gemiva.nl/locaties/de-walnoot

Abstract

We believe that we must work towards social systems that are more understanding, inclusive and welcoming of neurodivergent ways of interacting. Thus, the overall aim of the current work is to create a *diversity computing (DivComp)* device, that is, a technology system that helps autistic and non-autistic (NA) people to better understand, adapt to and appreciate each other's social behaviors and expectations. Specifically, this Master's thesis investigates how neurodiverse dyadic social interactions can be supported with multi-sensory embodied interaction technologies. *Dialogical systems* can help to change and transform *social systems*. Given the autistic embodied experience as a specific case, elements, such as the *double-empathy problem*, may make neurodiverse interactions – and achieving a mutually-satisfactory *structural coupling* – more challenging. A technology-mediated dialogical system could, thus, target this challenge to help accomplish the overall goal. Embodied interaction techniques that make use of sound and movement are, here, deemed promising, based on the fact that these modalities are also used in music therapy for autistic individuals. Hence, focus is put on how real-time measurements of body movements can be mapped onto sound feedback, to create an interactive system that aids in co-located neurodiverse *participatory sense-making*. In particular, an overall design-theoretical framework based on *embodied sense-making* was applied and a *participatory research-through-design (RtD)* methodology was followed, moving through two reflective design and research iterations. These involved various methodical approaches such as literature research, co-design, contextual and semi-structured interviews, questionnaires and expert reflections and discussions. As a result, a first sketch of the system, *Move4Music*, involving a list of design guidelines and requirements was reached. Specifically, in the first iteration, an *in-service Wizard-of-Oz (WoZ) data collection* set-up was made use of, in order to obtain insights and expertise from autistic and NA dancers and music therapists, in an embodied way. This aimed to inspire the design of *Move4Music*, in terms of what aspects of the interaction and joint movements it can aim to perceive, and how these can be mapped onto different sounds and sound changes. As a result, various insights were collected, including strategies participants use to try and dance with one another (e.g. mirroring/repetition, lessened distance) and sound aspects that can help to support connection-building (e.g. volume, unexpected sounds). These were used to come to a set list of movement-to-sound mappings, which were tested through WoZ in a second iteration, in an intended use context. In particular, technology-mediated interactions between autistic children and their NA teacher were observed for further refinements in the design. For example, the set of dancing-together strategies were affirmed, while the iteration also revealed important different characteristics in the target user group and their interactions (e.g. attention). Future directions were provided, such as working towards a working autonomous technology system and exploring more use contexts and longer-term repeated use and interactions.

Table of Contents

1. Introduction

This paper and the associated Master's thesis work carried out addresses neurodiverse dyadic social interactions and how to support these with multi-sensory embodied interaction technologies. The study takes an embodied social sense-making perspective (Hummels & Van Dijk, 2015; Van Dijk, 2022; Van Dijk & Hummels, 2017), looking at social interaction as an instance of 'participatory sense-making' (De Jaegher & Di Paolo, 2007).

1.1 Participatory Sense-Making

Maturana and Varela (1987) refer to a *social system* as building on a two-way co-acting and co-regulation. Accordingly, someone is seen as a member of this social entity, only if they find a place for themselves within a reciprocal structural coupling (Maturana & Varela, 1987). Following Maturana and Varela's original enactive framework, the concept *participatory sense-making* explains how intentional activity in social interactions is coordinated, in order for the interactors to together create and appreciate the interactions' meaning (De Jaegher & Di Paolo, 2007). While at a shared location, individuals together *enact* a common co-agency in interaction (Steffensen, 2012). For example, during an improvised performance or practice, jazz musicians all listen to one another, interpret and respond to each other's musical cues, and together establish the piece they are playing real-time. Similarly, if you are at a party, dancing, you can observe the others' movements to copy or build on them as a way of indicating that you want to connect and to dance together. Perhaps, you end up taking turns being the one deciding how to move to the music at that moment. This view sees *communication* not as mere information transmission, but rather highlights how coordinated behavior comes to be through mutual *structural coupling* (Maturana & Varela, 1987).

Steffensen (2012) further distinguishes between dialogical and social systems. Participatory sense-making, a *dialogical system*, concerns situational and situated behavioral coordination that takes place across shorter time spans whereby the actors are co-present (Steffensen, 2012). Social systems, on the other hand, have to do with the ways in which these dialogical patterns get habitualized – or at further levels – routinized, conventionalized and/or institutionalized (Steffensen, 2012). For example, in daily life, we may play our respective roles together with someone else in various situations. Making the expected small talk with the waiter at a restaurant, often consisting of the same few phrases, would be an example of social systems. There, we would be on a sort of autopilot, informed by the current norms. However, we may discover that our usual ways are not received as expected, or that we are interacted with differently by someone. Say, we, for example, get told something very directly and bluntly by a Dutch colleague; we are thrown off, instead, into a dialogical system. Our own norms may dictate this sort of behavior as 'rude'. Regardless, perhaps we know our colleague and that they would not have such intentions. Thus, we are now to together figure out the interaction and its meaning for both parties, and to authentically reflect and act on the situation in improvisation. Therefore, the outcome of the interaction, in terms of its meaning, is together established within the moment. This is in contrast to social systems, where, rather, a meaning already pre-coded in norms is sustained and repeated. Dialogical systems can, thus, take up a role of a sort of 'locus', to transform and even form social systems (Steffensen, 2012): What is considered as

norms can be dictated upon through the accumulation of small interaction instances that challenge them.

1.2 The Autistic Embodied Experience

As the official diagnostic name (i.e. autism spectrum disorder (ASD)) indicates, autism can present in different ways in individuals (Lord et al., 2020). This can lead to different unique strengths and challenges, and therefore also different needs (Lord et al., 2020). A current average estimate of the prevalence of autism is about one in every 100 children around the world (Zeidan et al., 2022).

A common perspective on autism is that it consists of a combination of social, communicative, and cognitive deficits (De Jaegher, 2013). Characterized as a disorder, it is seen as the originating point of 'problems' with social communication and interaction, and also other common patterns such as repetitive behaviors or restricted interests (American Psychiatric Association [APA], 2013). A typical description of an autistic persona, from the standard clinical point of view, may, for example, look like the following: Someone that 'fails' to respond to their name, 'resists' cuddling, has 'poor' eye contact and 'lacks' facial expression, doesn't speak or can't start or keep a conversation, has 'odd, stiff or exaggerated' body language and more (Mayo Clinic, n.d.). Hence, autism is commonly framed in this way, as a brain disorder that causes problems and calls for 'treatment'. Oftentimes, autistic people's social and embodied behaviors are misjudged. They are regularly expected to be the ones to change their ways to better 'fit in'. This may, however, bring forth a negative emotional experience and self-image in themselves, as can be seen by the following personal accounts:

"(...) It was teachers yelling at me for being a smart-arse because I answered a math question that asked 'How long would it take for a frog to get to a pond a kilometer away if it jumped a certain distance each day?', by writing a full page explaining that the frog would die before it got to the pond, because amphibians can't go more than three days without water.
It was consistent reports of 'She is incredibly intelligent, but disordered and doesn't fit in with her peers at all. She needs to try harder.'" (Hayden, 2023, p. 30).

"People will see an autistic child rocking, flapping or making repetitive noises and immediately deem it as off, weird, unnecessary and uncomfortable. Society has come to fear human beings simply moving their bodies in a way that doesn't suit its current understanding of 'normal'."
(Hayden, 2023, p. 48).

"I vividly remember writing a letter to God in my journal when I was five, asking why he'd made me so different, and if he could 'magic me' to not be so different, if it wasn't too much of a hassle.The idea of being different was terrifying, and the realization that I might always be different led to tears, panic and a desperation for change." (Hayden, 2023, p. 13).

On the flip side, there is increasing awareness of the different ways of perceiving and moving involved within the autistic experience, such as hypo- and hyper-sensitivities, timing-and-coordination-related difficulties, certain stimuli causing pain and muscle tone

differences (De Jaegher, 2013). Some argue that this basic sensorimotor difference generates patterns in behavior that, in the end, lead to problems in communication and social interaction (e.g. timing and rhythmic differences, see Isenhower et al., 2012; Sheridan & McAuley, 1997 and Trevarthen & Daniel, 2005).

Autistic people are not anti-social, but they can, for example, be overstimulated easily by emotionally-loaded facial expressions and therefore, from a young age, avoid looking at faces (for a review, see Stuart et al., 2023). This, hence, constitutes an example of how different social expectations and experiences can complicate neurodiverse social interactions: The reduced amounts or lack of eye contact-making by the autistic participants in interaction may be wrongly interpreted by the NA counterpart as them not paying attention or acting rude. While meaningful for themselves, certain characteristics (e.g. preference for repetition and sameness) of autistic people may come across to NA people as odd or inappropriate (De Jaegher, 2013).

1.3 The Double-Empathy Problem

As explained by Crompton et al. (2021), there is such a mismatch regarding the social expectations and experiences autistics and non-autistics (NA) introduce to the interaction, which may lead to a breakdown in communication. In effect, it often proves very difficult to establish a satisfactory structural coupling in such neurodiverse social interactions, and this mismatch is referred to as the *double-empathy problem* (Milton, 2012). The majority of people are NA and as a result of that, social systems have evolved around NA expectations and experiences. Thus, this entails that neurodivergent groups are at risk of being excluded from being seen as constituents of the social system.

Behaviors that are a meaningful part of an individual's embodied autistic experience, in fact, do not need to be seen as something to intervene and change. The following personal account exemplifies how family members, for example, can instead frame these positively:

> *"While growing up, stimming[3] was a part of my every day. I would flap my hands when I got excited, rock when I was upset, walk on my toes when I felt anxious and repeat sounds* (i.e. 'echolalia') *when I was overwhelmed. No one ever mentioned it, no one would ever bring it up or seemed to care about it, because it was just Chloe. That was just how I was. My parents often said I looked like a little fairy walking on my toes and flapping and jumping. When I was a toddler, Mum was convinced I must have been a dolphin in my past life because my echolalia seemed so like dolphin squeaks."* (Hayden, 2023, p. 50).

The great attention to detail, preference for repetition and sameness, and restricted interests, or similar of autistic people may, in fact, be inherently purposeful for them, rather than being just 'inappropriate behaviors to be treated away or fixed'. Commonly-misunderstood

[3] *Stimming* refers to self-stimulating behaviors consisting of repetitive movements of different body parts or objects (Rajagopalan et al., 2013). Everyone showcases some sort of stimming behavior (e.g. playing with one's pen or hair). However, the way neurotypical people do so can differ from the tendencies seen in autistic people (e.g. flapping arms, clapping, or jumping). These behaviors, regardless of how they are perceived in society by others, are thought to help with sensory and emotional regulation for the individual (Kapp et al., 2019; Mazefsky et al., 2013).

behaviors such as stimming are relevant and necessary for the autistic embodied experience, with possible negative consequences when wrongly suppressed to 'fit in':

> *"Pent-up stims changed into self-destructive behaviors that I couldn't control. Instead of flapping my hands, bouncing or repeating words, I would dig my fingernails into my skin until I bled, scratch myself until my skin was raw, hit the floor and bite my lips until they bled. I developed tics that made me jerk my neck to the point of long-term damage."* (Hayden, 2023, p. 52).

The current perspectives within the neurotypical society can be said to generally misrepresent the autistic experience and see it as deeply disordered.

1.4 Autism and Participatory Sense-Making

An enactive account of autism, instead, is based on a view of cognition as sense-making: Agents make meaningful connections to the environments they inhabit, on the basis of their goals and needs, to self-organize and maintain (Thompson, 2010; Torrance, 2005; Varela et al., 1991). This perspective offers new approaches to overcoming the problems faced, when compared to traditional functionalist accounts of autism (De Jaegher, 2013).

By applying the enactivism paradigm onto social cognition, we obtain crucial insights. In particular, we get a very different understanding of everyday face-to-face dyadic interactions: Here, individuals are not viewed as agents in interaction, whose roles are to merely figure out what goes on in the other's mind, in a calculative manner (De Jaegher, 2013). They are not in the position of passively perceiving and reacting to stimuli, neither from the environment itself nor from the other agents in it; from their actions. Instead, there exists a different definition of social interaction: It is but the regulated coupling that two autonomous agents establish amongst themselves (De Jaegher, 2013). Specifically, the relational domain these interactors share has to support the upkeep of both of their own, separate, identities (De Jaegher, 2013). For instance, De Jaegher (2013) gives as an example pair dancing, during which individuals take the role of either a leader or a follower. In interaction, the follower has to and does keep their autonomy. It is not as though the leader carries them around, as if they would be an inanimate object. Rather, the dancing consists of both the leader and the follower moving, in addition to the leader making the follower move and the follower getting moved by the leader: Each party must be engaging from an autonomous standpoint during a social interaction.

The enactive account highlights, thus, the need for not only one side being expected to conform and to fit into the other's norms, but for both interactors to better make sense of why the other moves and acts in the ways they do (Crompton et al., 2021). Instead of seeing certain autistic characteristics as the core problem source and targeting this, the bi-directional nature of this social challenge asks to be acknowledged. Hence, the NA society should also be helped in better understanding and accommodating the social expectations, needs and behaviors of autistic people, respecting also their autonomous identities and positionings.

1.5 The Design Challenge and Opportunities

There exists, therefore, a new framing of challenges observed in neurodiverse dyadic social interactions, as described above. Alternatives are there, in terms of how these can be addressed. Furthermore, the intersection of these with technology hint towards a design space ready to be explored:

In a literature study done earlier, I confirmed that there are several approaches that aim to better these neurodiverse interactions (Oral, 2022). However, most of these view the situation at hand, again, as stemming from the communicative behaviors and habits of the autistics (Oral, 2022) They treat autistic ways as wrong or inappropriate, and offer interventions that can help this group to change how they act, to better fit in with the social norms currently dominated by a neurotypical perspective (Oral, 2022). Due to autism's common conceptualization as a sort of disability and a focus on mitigating the functional limitations it is perceived to introduce into individuals' lives, the current design and research space is limited to such interventions or assistive technologies (Frauenberger et al., 2016a). Such work stems from a line of thought also found in therapies historically offered to autistic people. These aim, for example, to help them to stop displaying stimming behavior (e.g. Applied Behavior Analysis (Lovaas et al., 1987)). This is the case, despite the documented positive effects stimming has on relieving sensory-overload- and stress-related symptoms (Kapp et al., 2019). Thus, it may be suggested that such therapies – as well as technology systems that embody similar perspectives – address, instead, the discomfort of NA people that stems from lack of information and understanding on why the autistics engage in such behavior in the first place[4].

In my earlier paper, I have suggested that we must not continue on replicating the common normative interventions (Oral, 2022). Instead, we must move towards the technology taking up a *bi-directional mediation* role, and our works abiding less to normative trends (Oral, 2022). In line with this, the current project adopts the following vision and understanding:

Neurodiverse dyads indeed struggle to successfully achieve a satisfactory coupling. However, the basis for this is not to be sought purely within autism as a disorder. The weight of achieving a mutually satisfactory structural coupling should not be solely put on the shoulders of the autistic participants, nor should this 'satisfaction' be assessed solely from the perspective of NA ones (Crompton et al., 2021; De Jaegher, 2013; Milton, 2012). All agents included within an interaction should be able to, ideally equally, regulate the interaction (Steffensen, 2012). This must be grounded in and towards sustaining their own unique ways of keeping their autonomy in interaction with the world (Steffensen, 2012).

In line with an embodied, enactive, neurodiversity ethics, a technology design vision has been proposed called *diversity computing (DivComp)* (Fletcher-Watson et al., 2018). DivComp highlights the possible role artifacts can play in helping individuals to together establish what the interaction means for themselves. With their physicality, artifacts can act as *scaffolds* for a constructive interaction. This can include adapting different communication aspects to one another on both sides, prompting self- and collective reflection and better communication

[4]This lack of information and understanding is not necessarily out of bad intentions nor people not wishing to know. Rather, it is due to them acting based on their own expectations of what is normal, reasonable and empathic. This is despite these possibly being unproductive and having the exact opposite effect, due to how they are actually perceived and experienced by autistic interaction counterparts.

(Fletcher-Watson et al., 2018). Technology that aims to address the enhancement of autistic and NA interactions, despite the double-empathy problem, should provide better grounds for this process to take place (Oral, 2022). It should play the role of a mediator, a prompter. As a result, it must help to establish and keep balance points across varying, possibly conflicting values and goals introduced to the interaction by the autistic and NA interactors (Oral, 2022).

1.6 Overview of the Current Work

The current work aims to investigate how technological systems can contribute to the improvement of social interactions of neurodiverse dyads (e.g. autistic and NA), in novel ways. There exists a broader goal, through also future studies: A range of technology-mediated neurodiverse dialogical systems are aimed to be created. This is in order to work towards changing social systems into ones that are more understanding, inclusive and welcoming of neurodivergent ways of interacting. Therefore, the current efforts act as an example DivComp technology application that follows this line of thought: Building further on the knowledge and insights from fields such as music therapy, as well as earlier explorative design studies (Benjamin Leonard, 2022; Nguyen, 2021; Zhang, 2023), we are exploring *how real-time measurements of body movements can be mapped onto sound feedback, to create an interactive system that aids in co-located neurodiverse participatory sense-making*.

Consequently, this Master's thesis work aims to answer, in an applied manner, the main research question of *how neurodiverse dyadic social interactions can be supported with multi-sensory embodied interaction technologies*. Furthermore, the following sub-questions are asked:

1) *What are some main characteristics of differences in autistic & NA perspectives, perceptions and expectations that have potential for design opportunities?*

2) *What kind of joint movements mapped onto what kinds of sounds and sound changes can serve a prompting role, whereby a sense of 'connection' between autistic and NA participants in a dyadic interaction could be encouraged?*

3) *In terms of being predetermined or open and adjustable, what should the movement-to-sound mappings be like in the embodied interaction technology system?*

4) *How can we define and measure this sense of derived 'connection' or the technology and mappings' success, given the existing double-empathy problem and possibly different autistic & NA perspectives, perceptions and expectations?*

5) *If translated into a real-life product, what would be a good use context, scenario (e.g. educational at schools, home-use) and goal (e.g as 'prevention', training, more general learning and reflection) for this embodied interaction technology system?*

To inform these upcoming efforts, the current paper reviews and reflects on the relationship, thus far known, between autism and sound (see ***2. Autism, Sound and Movement***), especially thanks to music therapy contexts. Furthermore, former related technology applications are reviewed, to serve as inspiration and guidance (see ***3. Related Work***). Finally, relevant methodological approaches such as RtD and participatory design, as well as the existing and derived design guidelines and requirements, are looked into, to base current efforts upon (see ***4. Methodology***). As a result, a research and design plan for the associated Master's work is drafted and applied. As aforementioned, this is in order to explore how real-time measurements of body movements can be mapped onto sound feedback, in the form of an interactive system, to support co-located neurodiverse participatory sense-making. Accordingly, the remainder of the paper goes over the two reflective design iterations (see ***5. First Iteration*** and ***6. Second Iteration***) consecutively. In each, first, conceptual designs, prototype materials (i.e. sound and interface designs), participants included and the procedures carried out are described. Then, the gathered insights are displayed and interpreted. As a result, a list of design guidelines and requirements is output for *Move4Music*. Finally, the answers to the research questions overall, limitations and future directions are discussed (see ***7. Discussion***).

---

2. Autism, Sound and Movement

2.1 A First Glance Into Music Therapy

Autistic individuals possess unique sets of traits and there are different possible target contexts for improvement (e.g. social and motor skills) for this group. Hence, it is not a surprise that several ways of intervening and therapies have been established in the past for them (Sharma et al., 2018). Autistic children generally enjoy musical activities (Heaton, 2003). Furthermore, music has been shown to benefit them in different ways. This, for example, includes dealing with their behavioral, emotional, sensory, and motor challenges (Srinivasan & Bhat, 2013). In addition, music can help to better facilitate communication and support otherwise social skills (Whipple, 2012), promote attention (Kalas, 2012) and contribute to expressing emotions (Katagiri, 2009). Therefore, it is also reasonable that there is a continuous increase in the use of music-based therapies (Cibrian et al., 2018) (for a recent scoping review of music and sound-based interventions in autism, see Shahrudin et al., 2022). Former reports, for instance, estimate around 12% of all interventions and 45% of all alternative at-school strategies for treatment being music-based (Hess et al., 2008; Simpson, 2005).

At its core, *music therapy* consists of using music and musical elements in a systematic way and within an interactive therapeutic context, in order to help the client to regain, maintain and add to their emotional, physical and mental health (Fachner, 2017). The therapist and client both use their musical experiences in connecting and interacting with one another, interacting in and about music (Fachner, 2017). It can be *active*, including singing, collaboratively composing/writing songs or improvising, or *receptive* – only focusing on the listening of the music (Fachner, 2017). As reviewed by Shahrudin et al. (2022), its specific goals, the types of sounds incorporated as a result of that, what test parameters are utilized and in what sort of timeline (i.e. the number of sessions and their durations) can all be different on an individual basis.

Likely due to its very personalized nature based on the clients' specific needs, conditions, behavior patterns, personalities and histories, research conducted within the music therapy field often involves single-case study designs (Fachner, 2017). Therapies involving music may be fruitful in the context of autism given the growing evidence hinting at this. However, there is a need for further systematic studies to verify the positive effects, as most of the work is still considered to be experimental in nature (Brbić & Tomić, 2020; Garness et al., 2016; Lyra et al., 2017; Sinha et al., 2011; Vargas & Lucker, 2016). Currently, many studies suffer from limited sample sizes that lend themselves to poor generalizability of outcomes (Garness et al., 2016), lack of control groups to ensure better reliability and validity and more (Shahrudin et al., 2022).

Also as a result of such methodological limitations in the field thus far, questions arise as to what the most suitable and appropriate sounds would be for the different therapy intentions (Singh et al., 2016). Specifically in the context of 'interventions' that involve autistic children, this selection of reinforcements have been highlighted to impact therapy outcomes (Cibrian et al., 2018). Hence, a need is put forward for music-based therapy interventions, video games or

otherwise pervasive healthcare applications to utilize engaging auditory feedback (Cibrian et al., 2018).

2.2 Common Target Contexts and Musical Cues

One of the most researched aspects of autism in clinical music therapy contexts is social communication (Shahrudin et al., 2022). This involves different focal points such as receptive communication, social awareness, joint attention and imitation (Shahrudin et al., 2022). These reflect the current perspectives, in which difficulties of neurodiverse social interactions (i.e. autistic and NA) are attributed to the autistic parties' inherent characteristics, seeing these as possible grounds for intervening (Oral, 2022).

The investigation of the types of sounds and their suitable uses, as well as methods to decide on these and to customize them for individual clients, has been highlighted as a future direction by Shahrudin et al. (2022). However, one initial look into the sound preferences of autistic children was the work of Cibrian et al. (2018).

Their findings suggested that cello melody in low pitch yielded responses closest to that of a 'pleasant' sound from the children (Cibrian et al., 2018). This was attributed to the 'melody' structure being more easy to predict and to interpret (Berger, 2002). The same held also for the 'flat' sound included (i.e. consisting of 11 repetitions of the same note) (Cibrian et al., 2018). Cello, in addition, was highlighted as an instrument that produces more harmonics. This increased level of auditory variations was assumed to be more engaging (Cibrian et al., 2018). Furthermore, such an effect matched previous studies with neurotypical adults, where the 'happier' musical instrument was thought to be a violin, another string-based instrument (Wu et al., 2013). On the flip side, natural sounds such as water were found to be too abstract for autistic individuals (Cibrian et al., 2018).

In the case of low pitch, attention was better sustained over time with more positive responses (Cibrian et al., 2018). However, a use also for high pitch sound cues was highlighted: Autistic children's attention could be drawn in this way, as a kind of negative feedback and reinforcer to be used for a short amount of time (Cibrian et al., 2018).

Also through a study of interactive *sonification* (i.e. the use of sound to interactively represent data and data relationships) of children's spontaneous movement, it was shown that sounds may affect the resulting body movements' characteristics, as well as the outcome of therapies (Frid et al., 2016). This fitted with former research in *music-induced movement*. For example, according to Leman (2007), there exist three influencers of these corporeal movement reactions to music:

- *synchronization* at the basic level of beats,
- *embodied attuning*, so going beyond beats to move to more complex musical aspects such as melody, rhythm, tonality and timbre, and,
- *empathy*, so connection to musical features in an expressing-emotions sense.

Frid et al. (2016) argued for an existing cross-modal mapping between body motion qualities such as energy, smoothness and directness to sounds. For instance, from different sound models based on filtered white noise, the movements in response to smooth and

wind-like sounds were rated to be significantly more expressive and fluid (Frid et al., 2016). On the other hand, those in reaction to choppy and clicking sounds were instead perceived as more rigid and fast (Frid et al., 2016). The sound characteristics could also play the role of providing cues, influencing social interaction, communication and connection aspects (Abedi Koupaei et al., 2013; Al-Ayadhi et al., 2013).

In conclusion, the common target contexts in music therapy with regards to autism (e.g. social communication) are fitting of the current project's focal points. Furthermore, there exist several leads in terms of musical contents and cues, such as needing to customize these per individual clients and them needing an 'engaging' quality. This, for example, can be achieved through the sound cues having more harmonics, like a cello does, in comparison to a piano. The most suitable sounds for these purposes are yet to be further researched and determined, but easy prediction and interpretation seem to be important, which can be provided through including a melodic storyline. In contrast, unexpected and attention-drawing short high pitched sounds can be utilized as negative feedback. Finally, different levels of sounds (i.e. synchronization, embodied attuning, empathy) leading to different motion qualities (e.g. energy, smoothness, directness) can help to create different social interaction, connection and connection contexts.

2.3 Improvisational Methods and Techniques

Given the trends in research and application (Shahrudin et al., 2022), a very relevant sub-category of music therapy methods and techniques are those that are improvisational. One such trending context is facilitating social engagement and verbal and non-verbal communication skills (Kim et al., 2008), also relevant to the current work.

Musical improvisation can refer to the creation of any combination of sounds, created to have a beginning and an end (Wigram, 2004). *Clinical improvisation*, on the other hand, encompasses the use of this sort of improvisation, specifically to meet a set of needs of clients, within a trusting and supportive environment (Wigram, 2004). The latter includes techniques that help to:

- build empathy (e.g. imitation),
- provide structure to the sessions and the musical interactions that take place (e.g. rhythmic grounding),
- elicit responses (e.g. repeating),
- go in a different direction (e.g. introducing change),
- build intimacy (e.g. playing the same instrument) and more (see ***Appendix A***).

Moment-by-moment musical attunement is one part of shared improvisational music-making experiences. Through these, therapists are able to better tune into their patients' behaviors, and work on building a personal medium of communication with them (Kim et al., 2008).

Furthermore, when such (improvisational) music-making takes place in a dyadic or group setting, opportunities arise for individuals to develop social connections with one another. One example of this is that the individuals can be in a state of social cooperation during

music-making and feel a sense of togetherness (Kirschner & Tomasello, 2010; Overy & Molnar-Szakacs, 2009). These settings also involve many practice moments for imitation, turn-taking, joint attention, shared affect, and empathy (Overy & Molnar-Szakacs, 2009). As a non-intimidating, but engaging shared experience, music can prompt positive emotions (Srinivasan & Bhat, 2013), and may be more helpful in supporting social connections, in comparison to only verbal and non-verbal communication (Kirschner & Tomasello, 2010).

2.4 Current Knowledge and Design Insights

Based on the music therapy field as a whole, we can get initial design ideas for an embodied interaction technology system that would involve sound and movement modalities, in order to support neurodiverse social interactions. Specifically, common target points addressed for autistic groups (e.g. social communication (Shahrudin et al., 2022)) and in what way, using what type of musical cues and methods (i.e. improvisational) can show us the following:

There exists a gap in the design space, such that most technological systems embody an interventionist perspective. These see autistic behaviors to be at the core of the communication challenges experienced by autistic and NA individuals. This is confirmed and matched by what music therapy approaches perceive as the origin of such struggles, as observed in neurodiverse embodied social interactions. Thus, the situation calls for a solution that acts as an intervention for the social interaction itself, rather than one of the individuals' own autistic ways of interacting. Such a bi-directional implementation can take inspiration from the music therapy field in different ways. For instance, this can be in the form of also adopting a dyadic or group set-up that can provide social connection opportunities, where music acts as an additional layer that provides a sense of togetherness (Kirschner & Tomasello, 2010; Overy & Molnar-Szakacs, 2009).

Say we are speaking of a system whereby real-time measurements of body movements are mapped onto sound feedback, to create an interactive system that aids in co-located neurodiverse participatory sense-making: Music therapy can provide insights on what kind of a musical baseline the system can provide, based on its improvisational methods (see ***Appendix A***). Specifically, the two structuring techniques – rhythmic grounding and tonal centering – can highlight either providing a basic beat or a tonal center, scale or harmonic basis as two options, respectively (Wigram, 2004). One of these improvisational methods, sharing a musical instrument – a technique to help build intimacy – would also support our intended technology system conceptually – given that it also targets connection-building (Wigram, 2004).

Similarly, in terms of what the technology system can aim to detect and measure, to be able to judge the technology-mediated social interaction's success, these improvisational techniques can serve as a starting point. Techniques intentionally used in music therapy to create empathy (e.g. imitation, synchronization/mirroring, incorporation), to elicit a response (e.g. repetition, making space) or to lead the session to a different direction (e.g. introducing change) (Wigram, 2004) can all be possible markers we can look out for in our observations of the neurodiverse social interactions in our own context. In this way, we can see if they perhaps make up promising acts we can aim to detect and measure our technology system's success based on.

Following this stage of inputting some meaningful aspect of the ongoing neurodiverse embodied social interaction into our technology system, music therapy can also help us in deciding what the system can output, musically speaking. There are existing gaps in the musical preferences of autistic people and the suitable uses of different sounds and sound aspects (Shahrudin et al., 2022). Hence, this can indicate a more exploratory approach as a suitable first step in the current work. The few existing pointers we have, on the other hand, can act as guides. For example, they may advise us to avoid the use of natural sounds, which may be perceived as too abstract (Cibrian et al., 2018). Similarly, aspects such as having high levels of harmonics can be adopted, to perhaps be perceived as more engaging (Cibrian et al., 2018). That and having elements that are easy to predict and to interpret may be important consideration points when selecting the musical contents of the system outputs (Cibrian et al., 2018).

The relationship between sounds and movement qualities can provide clues as to what kind of changes we can attempt within our technology system outputs (Leman, 2007). These include, for instance, synchronizing to the beats, embodied attuning to various aspects like melody or tonality or empathy and connecting to the music emotionally (Leman, 2007). There is existing literature supporting sound characteristics providing cues, influencing social interactions and connection (Abedi Koupaei et al., 2013; Al-Ayadhi et al., 2013). Therefore, we do have a basis for using music-and-movement-based embodied interaction technologies to support neurodiverse social interactions.

---

3. Related Work

3.1 Autism Interaction Technologies

Both in academia and in practice, there are several technologies that aim to improve neurodiverse social interaction contexts. These examples, even when interventionist, are possibly relevant for the current efforts. In this section, focus is put on technology examples that can add direct consideration and inspiration points to the current design process. Hence, examples that are mainly focusing on other modalities, such as visuals, are intentionally left out and not discussed in detail. Instead, technologies that feature either the movement or sound modalities are looked further into.

Upon reviewing the current state of autism interaction technologies (i.e. ones that, again, mainly focus on the movement and sound modalities), I had shown that there exist several systems that utilize different kinds of technologies (e.g. interactive surfaces, games and wearables) (Oral, 2022). These often feature contexts such as therapies or educational in-school ones and target children (Oral, 2022). For example, *OSMoSIS* is a musical motion-based game within an interactive environment, for therapy- or home-use (Ragone et al., 2020; Ragone et al., 2022). It aims to support music therapy practices that target motor-skill improvement, by providing sound feedback: The child is expected to copy the therapist's movements and when they do so, hear a certain sound that matches that specific movement, as a form of encouragement (Ragone et al., 2020; Ragone et al., 2022).

Such technology examples are useful in terms of the common design features and elements they incorporate (i.e. playfulness, *phases-structure* and *together use* (Oral, 2022)). In addition, they exemplify the addressing of the user group's, autistic people's, needs (e.g. adjustability and participatory methods); both of which serve as inspiration and guidance (Oral, 2022). However, most existing autism interaction technology systems were shown to be interventionist (Oral, 2022). These identify the origin of the challenges experienced within neurodiverse social interactions to stem from autistic characteristics. They do not acknowledge and address the interaction between these individuals and their neurotypical counterparts and overall environments, focusing on intervening to change the autistic ways instead.

The technology framework *diversity computing (DivComp)*, on the other hand, aims for the creation of artifacts that can help to mediate a bi-directional meaning-making process between individuals, abiding to an embodied, enactive, neurodiversity ethics. (Fletcher-Watson et al., 2018). It aims for artifact-mediated scenarios, in which meaning-making between people with different backgrounds (e.g. autistic and neurotypical) can be supported, in order for them to better understand one another (Fletcher-Watson et al., 2018).

There exist several projects that can fall under this technology framework's implementations. Some of them are specifically chosen to be highlighted here, rather than interventionist ones: The work of Spiel et al. (2016), for example, is highly similar to the current work, in its intentions regarding and focus on designing products that enable autistic participants to have meaningful experiences per their own definitions. There, technology plays the role of empowering its users in sharing these with their surroundings, in an open-ended manner (Spiel et al., 2016). Their *embodied companion technologies* revolve around the key element of

unlocking a new, novel way of involving others in your individual sense-making (Spiel et al., 2016). Within the four design cases that incorporate different forms of embodiments (i.e. reactive embodiment, ambient embodiment, embodied reflection with temporal separation and embodied control), two products, *BSmart* and *Adaja,* are highly relevant and together form a nice basis for the current efforts.

BSmart is a smart companion that, for example, supports its user by providing information about upcoming movies, as well as story prompts (Spiel et al., 2016). It does so, in the form of up to ten pictures that are projected on a surface, such as a wall (see ***Figure 1***). It has the quality of affording embodied interaction between people through reactive embodiment (Spiel et al., 2016). When *BSmart* is pointed towards different kinds of surfaces, there is an interplay of these location elements and the projection itself, which creates differences in the images. This can prompt the users to possibly move around *BSmart*, leading to further changes – a chain of opportunity creation for exploration (Spiel et al., 2016). This prompting aspect and the type of embodiment the artifact features can inspire the current work, even if the modalities involved are not the same.

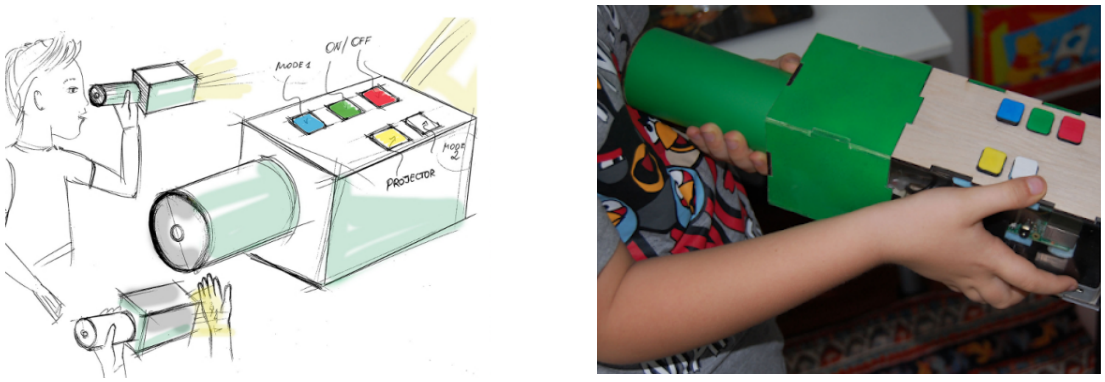


Figure 1*[5]*: The *BSmart* concept sketch *(Left)* and prototype *(Right)*. *BSmart* projects photos featuring movie information and act as story prompts onto surfaces (Spiel et al., 2016).

Adaja, on the other hand, is an ambient companion aimed at sound environment exploration (see ***Figure 2***) (Spiel et al., 2016). Thus, it has this sound modality context in common with the current efforts. It includes a microphone through which to pick up on sounds, as well as a screen to create visualizations based on their intensities. One of its use cases is that people, specifically children, can be prompted to interact and to make sounds in collaboration upon receiving these visual feedbacks, to be able to together observe how they change (Spiel et al., 2016). Therefore, *Adaja* again incorporates a prompting element, getting closer to the current modalities we are interested in: It features sound inputs and visual outputs, rather than movement inputs and sound outputs, like in the current work.

[5]Adapted from "Embodied Companion Technologies for Autistic Children.", by K. Spiel, J. Makhaeva and C. Frauenberger, 2016, *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, p. 247.

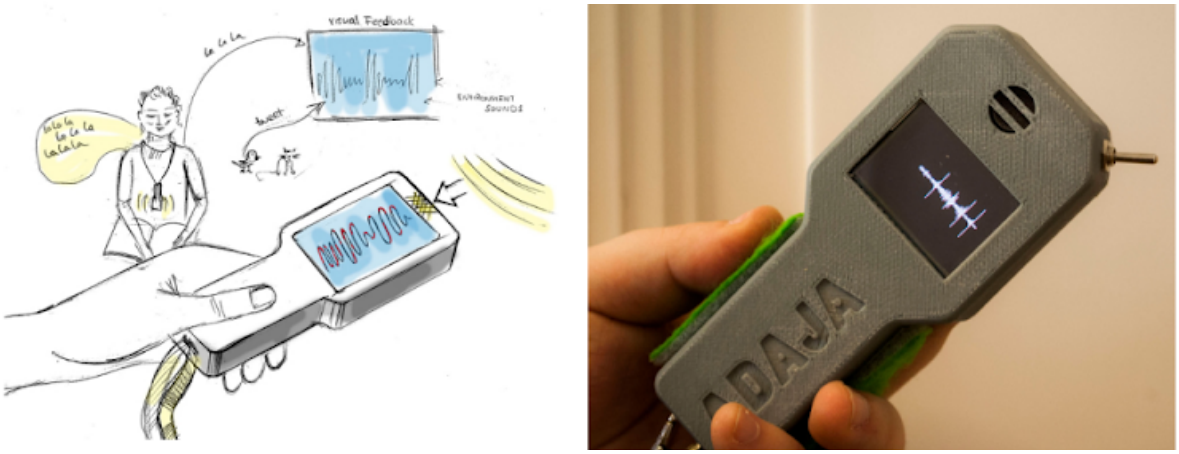


Figure 2[6]: The *Adaja* concept sketch *(Left)* and prototype *(Right)*. *Adaja* picks up on sounds in the environment to turn these into visualizations, based on their intensities (Spiel et al., 2016).

One real-life/outside-of-academia example that makes use of the intended modalities is *Visible Voice* (Monobanda, n.d.). Though not self-defined as such, it can also be categorized as a DivComp device. Rather than being positioned in a neurodiversity- and/or social-interaction-related context, *Visible Voice* aims to provide embodied interactive support tools in the context of emotion-understanding and management. As a multi-medial interactive installation, *Visible Voice* offers youth a safe and welcoming environment to better handle feelings of aggression and anxiety (Monobanda, n.d.). It does so, by mapping emotions expressed through vocalizations and movements onto visuals (see ***Figure 3***). The interaction with this installation is designed in four consecutive 'acts' (Monobanda, n.d.):

The process starts with having the youth ***express*** any built-up negative emotions, by means of moving around and making sounds. *Visible Voice*, as the name suggests, makes visible these expressive actions in different ways. For instance, this can be by following them around and drawing a line on the ground, or changing elements such as colors in response to different sound qualities. In the next act, it gets them to interact with and reflect on this visual representation, to gain ***insights***. To do so, they are offered movement-and-sound-based tools they can contribute to the visuals with. Now, in this way, they are given more control: They can act differently to end up with different paintings. For ***regulation***, in the next act, they lie down on this resulting image. They see themselves as one with it on a mirror on the ceiling. In doing so, they find calmth, also through the accompanying multimodal experience (e.g. soothing music playing). As a final act, they can have a ***conversation*** with their caretakers, using this visual representation of their emotions. This image can help them to better frame and understand their emotions and actions. What makes *Visible Voice* fit the DivComp framework lies exactly at this stage: Their individual experience and the mappings or materializations of their emotions onto these drawings can act as means for social interaction; supporting the participatory sense-making moment that comes after (De Jaegher & Di Paolo, 2007).

[6]Adapted from "Embodied Companion Technologies for Autistic Children.", by K. Spiel, J. Makhaeva and C. Frauenberger, 2016, *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, p. 249.

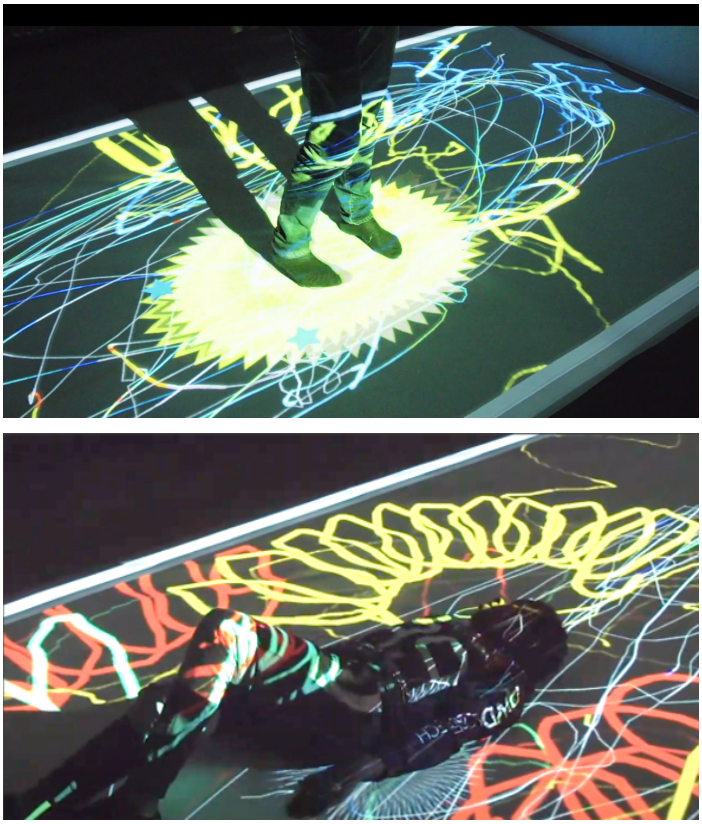


Figure 3*[7]*: *(Up:)* An example vocalization- and movement-to-visuals mapping, generated within the *Visible Voice* (Monobanda, n.d.) installation. *(Down:)* An image from the *'regulation'* act within the installation, where the youth lay down on the resulting visual, looking up at a mirror. There, they see themselves in combination with this image they created by expressing their emotions.

Another work that does explicitly subscribe to the DivComp framework, and is a precursor of the current work to be carried out, is *Stim4Sound* (see ***Figure 4***) (Nguyen, 2021; Zhang, 2023). It does not focus on the possible role music and movement – and embodied interaction technologies that incorporate these – can play in supporting neurodiverse social interactions in a general sense. *Stim4Sound* focuses, instead, on a specific behavior associated with autism, which is stimming. It aims, on the one hand, to normalize this behavior, and to help autistic individuals to unmask and to feel comfortable doing so (Nguyen, 2021; Zhang, 2023). On the other hand, it assists NA people in better understanding their peers and their stimming behaviors (Nguyen, 2021; Zhang, 2023).

Stim4Sound involves, as its main interaction, two individuals being able to make sounds with different objects surrounding them in their daily lives, through common stimulating behaviors (e.g. tapping or shaking). It involves a wearable device that is able to record this

[7] [Untitled photo of 'Act 3: Regulation' of 'Visible Voice']. Visible Voice. https://monobanda.eu/project/visiblevoice

music created. The device also senses certain actions to make changes in different musical elements, in parallel to changes in movements. Through *Stim4Sound*, the interactors can together reframe these stimming behaviors in a social context (Nguyen, 2021; Zhang, 2023). The provided tunes and rhythms, as well as the auxiliary instruments, help to inspire interaction through movement-to-sound mappings (Nguyen, 2021; Zhang, 2023).

Figure 4*[8]*: The *Stim4Sound* sketch *(Up)* and prototype *(Down). Stim4Sound* allows for recording and replaying sounds made using everyday objects . Through common stimming behaviors, these clips can get altered (Nguyen, 2021; Zhang, 2023).

Together, all these (autism) interaction technology examples serve as inspiration for the current work – whether regarding their qualities, approaches, features and/or uses of the relevant sound and movement modalities.

3.2 Adaptive Music

The term *adaptive music* refers to the generation of music, which dynamically adapts to contents and actions (Hutchings & McCormack, 2019). For example, within video games, it can help to create more immersive, memorable and emotive experiences (Hutchings & McCormack, 2019). It involves the reorganization of the existing musical material, to give rise to a new musical experience based on these (Brown & Kerr, 2009). During the process of doing so, what

[8] Adapted from “Stim4Sound: a musical interactive system promoting communication between autistic people and society.”, by Y. Zhang, 2023, p. 20.

is expected of the music in terms of mood or expressiveness is kept in mind (Brown & Kerr, 2009). This is just like other game aspects that use complex and emergent player-directed narratives: Sound changes are influenced by the players' personal decisions within the said games (Louchart et al., 2015). Furthermore, this utilization of the sound modality is a part of experience-driven procedural content generation (Shaker et al., 2016; Yannakakis & Togelius, 2015).

Within a first-person shooter game, for example, adaptive music[9] can look like the dynamic layering of instrument parts getting altered. These changes can be based on parameters such as how accurate the player is at shooting their enemies, the number of attacking opponents or the player's own health levels. There can be an addition or subtraction of such layers from the ongoing music, to create a certain affect.

A key element of adaptive music is the identification of the overall goal or context, as well as what kind of (in-game) parameters are relevant and meaningful to this goal: In a shooter game, players are aiming to achieve victory and to defeat their opponents; to stay alive despite the ongoing attacks. This makes the aforementioned parameters important ones in this case, for achieving immersion or evoking emotion (Hutchings & McCormack, 2019).

Then comes another important element, which is the mapping of these parameters onto changes in music: This should make sense in terms of the effect we aim to have in the players' experience of the game. If we want them to feel more tense as they are not successful in combat or as their life levels are dropping significantly, adaptive music may look like adding to the tempo, repeating the same bundle of chords or notes, or adding dissonance to the composition to evoke an eerie feeling, as an example.

Importantly, these mappings should be combinatory, featuring more complex – rather than one-to-one (e.g. volume going down when the player's life levels go down) – input-to-output matches. In this way, the mechanics of the adaptive music function can be made less obvious for the players. As a result, we can allow for a more nuanced impact of player actions on the sound changes.

While most common in the game industry, also within other media art works, adaptive music finds a place for itself. It has, for instance, been used as a means of opening up the use of a music jamming software also to musically inexperienced users, such as children (Brown & Kerr, 2009). Similarly, within artistic installations that require the musical material to be responsive to user actions and environment changes, adaptive music can be utilized (Brown & Kerr, 2009). Another example case, *D-Jogger,* involves the incorporation of music for prompting walkers to synchronize to a musical beat without explicit instruction (Moens et al., 2014). In this way, the possible use cases of adaptive music is extended to sports, physical rehabilitation or assistive technologies for movement performance (Moens et al., 2014).

There exist different techniques, for example in the form of algorithmic implementations that enable real-time manipulation of different musical aspects (for a review, see Brown & Kerr, 2009). These offer several opportunities when designing interactive and collaborative musical experiences. Together with the examples given and their lack thereof in other areas, such ways point towards a broad design space yet to be explored, in terms of adaptive music use.

[9] The information and examples given in this section regarding adaptive music were inspired by conversations on the topic with an expert/practitioner in the field, Rik Nieuwdorp (for his experience and work, see http://www.claynote.nl/).

3.3 Current Practices and Design Insights

Based on the existing work, both in the form of autism interaction technologies and within the field of adaptive music, we can further refine our initial design ideas for our embodied interaction technology system – as inspired by the music therapy field. In particular, we can think more specifically regarding movement-to-sound mappings' and the overall systems' qualities:

For instance, we can incorporate design features and elements commonly observed in autism interaction technologies (Oral, 2022). This, for example, can look like using gamified elements to be more engaging (i.e. playfulness) or featuring collaborative actions as a constraint and a context, so more social engagement and connection can be prompted (i.e. together use) (Oral, 2022).

We can approach the design by minding also the different sorts of embodiments we may want to incorporate. This can be regarding the way people interact with the system and/or with each other, through the system. Similarly to *BSmart*, we can aim for our system to afford embodied interaction between people, through *reactive embodiment* (Spiel et al., 2016). Rather than making use of the movement-to-visual changes – which, in the case of *BSmart*, result from people moving around the projections on different surfaces (Spiel et al., 2016) – we can imagine how this sort of embodiment could be achieved through movement-to-sound mappings. This can be attempted at, focusing on the 'interplay' aspect of the two modalities in interaction and on creating a chain of opportunity creation for together-exploration (Spiel et al., 2016). Such an approach would be consistent with the practices within the field of adaptive music: There, also, parameters relevant to the main intentions and goals of the context would be selected and combined in a meaningful way, to influence changes in sounds. That, in turn, would serve different purposes, such as immersing people (Hutchings & McCormack, 2019) and helping to create a feedback loop as intended.

This sort of a prompting role of sounds we want our technology system to possess would additionally relate to *Adaja* and its *ambient embodiment* (Spiel et al., 2016). For that object, sound-to-visual mappings are ever-present and influence people's interactions within the space and with each other (Spiel et al., 2016). In our case, we can aim to work towards designing movement-to-sound mappings that can have the same effect. Specifically, we can focus our efforts on researching different options, when it comes to different sound parameters and how helpful these are.

Furthermore, we can be similar to *Visible Voice* (Monobanda, n.d.) and also abide by the phases-structure often seen in autism interaction technologies (Oral, 2022). This quality highlights the importance of finding a balance between two things: Providing more guidance and structure versus allowing for free-form interaction (Oral, 2022). Specifically, the interaction with our technology system and the interaction of its users with each other, as a result, can be designed to consist of different 'acts', a storyline. In applying this, we can also consider the possible roles the system can play. Such a role can relate to not only the social interaction in and of itself, but also involve the technology-mediated interaction acting as a conversation-starter. At a later stage, the system can position itself between the neurodiverse participants, to support their participatory sense-making processes (De Jaegher & Di Paolo, 2007).

Last but not least, we can aim to measure our system's success based on the following line of questioning: Is everyone involved in the technology-mediated social interaction able to have a meaningful shared experience per their own definitions? (Spiel et al., 2016).

Overall, these design insights, taken together with the initial ones that come from music therapy, help to assess our research investigation and input needs. They help us in selecting appropriate approaches and planning our design and research cycles (see ***Table 1***).

Table 1: An overview of insights from literature search, to be fed into and further developed through the first iteration design and research work.

Source of Insight	Insight
Music Therapy field & Autism Interaction Technologies' examples	Given the existing gap in the design space, the system can/should be an intervention for the social interaction itself, rather than one of the individuals' own autistic ways of interacting. It can/should play the role of bi-directional mediation (Oral, 2022).
	Adopting a dyadic or group set-up for the use of the system can help to provide social connection opportunities (Kirschner & Tomasello, 2010; Overy & Molnar-Szakacs, 2009). Similarly, collaborative actions as a constraint and a context (i.e. together use (Oral, 2022)) can prompt more social engagement.
Music Therapy field	The system's use can be designed to imitate the sharing of a musical instrument; a method in music therapy, which helps to build intimacy (Wigram, 2004).
	Improvisational techniques (Wigram, 2004) can act as a starting point for what the technology can detect or measure (e.g. imitation/mirroring) as indications of social connection during interaction.
	The system's musical baseline can be based on the two structuring improvisational techniques in music therapy: Rhythmic grounding & tonal centering (Wigram, 2004).
	Different levels of sounds (i.e. synchronization, embodied attuning, empathy (Leman, 2007)) leading to different motion qualities (e.g. energy, smoothness, directness

	(Frid et al., 2016)) can help to create different social interaction, connection and connection contexts (Abedi Koupaei et al., 2013; Al-Ayadhi et al., 2013).
	Musical contents and cues need to be 'engaging', for example through having more harmonics, like a cello does, in comparison to a piano (Cibrian et al., 2018).
	Musical contents and cues that allow for easy prediction and interpretation may be important and suitable (Cibrian et al., 2018), which can be provided through including a melodic storyline (Berger, 2002).
	Unexpected and attention-drawing short high pitched sounds can be utilized as negative feedback (Cibrian et al., 2018).
	Musical contents and mappings may need to be customized.
Autism Interaction Technologies' examples	Common features and elements (e.g. playfulness, phases-structure) present in autism interaction technologies (Oral, 2022) can act as design guidelines.
	Success of the system can be measured through what is a meaningful (shared) experience per the participants' own definitions (Spiel et al., 2016).
	Different sorts of embodiment (e.g. reactive & adaptive (Spiel et al., 2016)) can be incorporated, to support/add to connection-building in different ways.
Adaptive Music field	Principles of adaptive music can help to achieve reactive embodiment within the sound modality (e.g. identifying (movement) parameters that are relevant and meaningful to the technology-mediated interaction goal, mapping these onto sound changes in a complex and combinatorial manner).

4. Methodology

4.1 Research-through-Design (RtD)

Design and *research* have historically been considered as distinct endeavors, with the former relating to practice in the industry and craftsmanship, and the latter to experimentation and reflection in academia (Stappers & Giaccardi, 2017). However, the role of research in designing products and services has come to be recognized in recent years (Stappers & Giaccardi, 2017). Designs' purpose is now not understood as solely creating solutions for certain contexts. Instead, in fields such as interaction design (IxD), design artifacts also became crucial in how knowledge is generated and further communicated (Stappers & Giaccardi, 2017). Common practices of designers such as sketching, modeling, planning, visualizing and especially prototyping (i.e. concepts and approaches' materialization and embodiment), came to serve several purposes. According to Stappers (2014), for example, they can:

- give room to experimentation through their unfinished nature,
- allow for living out a future situation,
- act as a connection point between abstract theories and experiences,
- help to mediate (interdisciplinary) discussions,
- serve as props for storytelling and other activities, and,
- be a process marker and a reference point in a project.

The act of prototyping is confrontational, in the sense that to make something is to have to handle how ideas (e.g. visions, theories) and the physical and social contexts can meet one another (Stappers, 2014).

Both design and research further have in common that they aim to build on the known, to create something new, and are interwoven: Analyzing and evaluating is necessary while designing, and research involves the development of materials (e.g. apparatus, stimuli) (Stappers & Giaccardi, 2017).

An approach given rise to as a result of such bases, *research-through-design (RtD)*, consists of the application of design methods within the research process (Frayling, 1994, 2015). In doing so, knowledge that is more concrete can be generated (Stappers & Giaccardi, 2017). Rather than staying at an abstract level, its methods help to transfer such insights onto a specific situated context (Stappers & Giaccardi, 2017): In the applied example of *MyDayLight* (van Dijk et al., 2019), for example, created design artifacts mediate autistic young adults and their care-givers in sharing, questioning and reframing their implicit understandings. In addition, this helps to expand on the role assistive technology can have in their life-worlds (van Dijk et al., 2019). In these kinds of cases, research is done through (reflecting on) a design of a product.

Furthermore, RtD can, for example, allow for prototypes themselves to determine research directions. This can be through them embodying how a certain theory approaches the given design issues, aiding research in this materialized way of hypothesizing/question-asking (Stappers & Giaccardi, 2017). Through the creation of a *group* of artifacts (for a single purpose or around a specific concept), RtD can also help with knowledge generation (i.e. framing)

(Gaver & Bowers, 2012). It can provide information based on each of these artifacts' shared qualities, similarities and differences (Gaver & Bowers, 2012). Designers can be supported in their reflections on their approaches (e.g. to weed out any dead-ends that require no further exploration). What's more, they can more easily expand on the existing design spaces, with prototypes playing both an *evaluative* and a *generative* role respectively (Stappers & Giaccardi, 2017). In this way, they fulfill purposes of both *filters* and *manifestations* (Lim et al. 2008, Wensveen & Matthews 2015). In revealing such information and patterns, RtD can further lead to *strong concepts* (i.e. knowledge that can be abstracted beyond its use in particular cases, to then play a role in new contexts and corresponding designs) (Höök & Löwgren, 2012).

A characteristic of RtD that reflects the benefits it can provide is that it involves working in iterations. Within this methodology, we alternate between and repeat steps like prototype creation, evaluation and expert reflections on these. Furthermore, coined as *reflection-in-action* (Schön, 1983), the explicit use of theory within this reflection is also a part of RtD.

Rather than prescribing a solution for an everyday situation, designing has to do with understanding and assisting with what already fits in that context (Hales, 1986). It comprises a 'reflective dialogue', rather than a search – crucially depending on the interaction between the designer and the materials (e.g. tacit knowledge, skill, socio-cultural practices) of the context (Roozenburg & Dorst, 1998). To design is to come up with an action direction through what we attend to within a context and how we organize our insights from different aspects of the said context (Roozenburg & Dorst, 1998). We are always working in indeterminate situations that involve incomplete knowledge and are having to make decisions despite this.

In the current study, various methodological approaches were woven together into an iterative reflective process. For instance, participatory/co-design (see ***4.2***) was utilized, involving prototype deployment and evaluations through Wizard-of-Oz (WoZ). Similarly, literature research (see ***4.3***) was conducted – for example, on guidelines, requirements and design principles that can later inform reflection-in-action. Additionally, contextual and semi–structured interviews, questionnaires and expert discussions (see ***4.4***)) all also found their place in the application of RtD (for a more detailed rundown, see ***4.5***).

4.2 Participatory Design (PD)

In fields such as human-computer interaction (HCI) and IxD, the involvement of the target user group in the iterative research, design, development and evaluation processes can help to gain a deeper understanding. Specifically, the problem context and the pain points, wants and needs of the different stakeholders involved can be further clarified. This can, for example, assist in the resulting products having better acceptance rates. Not only that, but these research and design processes can act as a platform of empowerment, having a transformative agenda as their backbone (Cousins & Whitmore, 1998). This relates to the positionings of marginalized groups and the existing power differences in academia or technology development contexts (Bratteteig & Wagner, 2012). It originates from researchers possessing specialized knowledge, skills, and capacities to put these into practice (Bratteteig & Wagner, 2012). As pointed out by Spiel et al. (2017), when the subject matter is design of

technological products, these practical and transformative agendas converge, especially with groups such as neurodivergent individuals, who may face increased levels of marginalization[10].

Such *participatory-design*-based methods – where users of technologies are included in shaping these artifacts that they will later be interacting with (Simonsen & Robertson, 2012) – are popular, also in contexts such as the current one. Despite this, oftentimes the partaking target user group is reduced to the data researchers obtain (Spiel et al., 2017). The researchers do so, judging the success of a design element, in the end, based on their self-defined criteria (Spiel et al., 2017). The different levels of participant involvement, as outlined by Cousins & Whitmore (1998) are not acknowledged and more effectively made use of. Instead, these input degrees are often limited and shallow. For example, in the context of neurodiversity, a technology system's success in assisting with social interaction may get assessed based on how much more eye contact autistic participants make in its use. This would be despite these individuals likely not defining eye contact to be their goal or the meaning of 'being social' themselves. The autistic individuals' participation, now, would be limited to joining in the testing stage, for the data they provide to be evaluated according to this criteria[11].

This lesser-power positioning is even more apparent in the case of technology designed and developed for *children*: They, again, are mainly involved as testers (Spiel et al., 2017). Children pose unique challenges as co-designers, as they need to be met with what they are capable of and how they are able to express their thoughts and emotions, for their inputs to be correctly understood (Spiel et al., 2017). This can, for example, look like offering methods such as giving them a camera, with which they can take a picture of what they like in a certain context. Similarly, they can be made to or assign different roles to their toys and voice different discussion aspects or opinions through these (e.g. pros and cons of a certain design element). Different sets of guidelines have been posed based on former work regarding dividing and balancing researcher and child contributions (e.g. Frauenberger et al., 2013).

Participatory design is even more important in case of autistic children (e.g. Benton et al., 2012; Frauenberger et al., 2016b; Keay-Bright, 2007; Malinverni et al., 2014; Pares et al., 2005). As pointed out earlier, many, if not most, technological projects work from an interventionist point of view (Kientz et al., 2014; Oral, 2022). In light of this, there is a need for design and evaluation processes that center on the user's authentic view and perspectives. Designing from the lived experience and needs of the user may help to create products that are less interventionist in nature.

4.2.1 In-Service WoZ Data Collection

In the application of the *Wizard-of-Oz (WoZ) methodology*, people – referred to as 'wizards' – simulate technology system actions, often without participants realizing this (Wooffitt et al., 1997). While there exist critiques on this methodology (for a review, see Eklund, 2004), it is a common way of collecting data in the field of HCI. It can be opted for for different purposes

[10] A famous quote from disability activism: *"Nothing about us without us."*

[11] It is, of course, important to also note the tyranny of participation here, where, even if steps are taken to adapt methods and processes to have participants contribute more and on more levels, the same level of interest in participation will not be there for all participants at different design and evaluation stages (Cooke & Kothari, 2001).

– for example, in cases where building of a fully-functional system is not needed, practical or cost-effective (Edlund et al., 2008).

One of the many variations of this paradigm is when the wizard is also taken as a subject (Edlund et al., 2008). There, the WoZ set-ups are used to observe, study and perhaps copy the wizard's actions (Edlund et al., 2008). In contrast to some other WoZ variations, participants being under the impression that a computer, not a human is producing whatever output and is interacting with them (Fraser & Gilbert, 1991; Wooffitt et al., 1997) is not as big of a prerequisite in this variation (Edlund et al., 2008).

Example cases of this wizard-as-subject variation include how *TeliaSonera*, a Swedish natural language call routing system, was developed (Boye & Wirén, 2007; Wirén et al., 2007). There, actual customer care operators – experts – were involved, to fulfill the wizard role and to provide insights regarding the dialogue and prompt design (Boye & Wirén, 2007; Wirén et al., 2007). They were given a prompt piano, with keys corresponding to various pre-recorded prompts that they could use to respond to callers, with these choices being logged for later design purposes, which was coined as *in-service WoZ data collection* (Boye & Wirén, 2007; Wirén et al., 2007).

Though this is not as common, in the current work, a similar flavor of WoZ fell under the PD approach applied – and is, hence, mentioned here. The in-service WoZ data collection paradigm was utilized, in order to extract situated expert knowledge and actions, as a way of having them partake in the system input and outputs' design. In particular, participants that had either dance or music therapy expertise (see ***5.3 Participants***) were made to control the sound outputs of the system in the first iteration, playing the role of the 'technology system'/DJ. This was combined with the use of another participatory method, contextual inquiry (i.e. a qualitative data-gathering and analysis method wherein users are observed and conversed with at their regular settings, as they go about their daily practices (Karen & Sandra, 2017)), as well as more traditional user research methods such as semi-structured interviews, in order to aid participants in influencing the different levels of the design process.

4.3 Guidelines, Requirements and Design Principles

The current work took as consideration points and constraints, several guidelines, requirements and design principles throughout. All of these were included at different capacities; for example, serving as theoretical bases, rough inspiration or hard requirements. Specifically, the decisions taken throughout the participatory RtD process were based on:

- theory (informing process and methods, or design),
- creative inspiration,
- empirical data (from the workshops carried out), and,
- expert reflections & discussions (involving feedback from peer researchers).

As aforementioned, additional or unique challenges are involved when applying participatory methods, especially in collaboration with autistic children (e.g. receiving concrete feedback (Frauenberger et al., 2013)). Thus, some pointers have been established in the past for this context – that is, ***theory that informs process and methods***.

One such example is the seven guidelines Spiel et al. (2017) offers for when working with autistic children and partaking in participatory evaluation. In the current work, these guidelines served as inspiration as to how to structure the sessions and the setting itself. An instance of this kind of contribution was *guideline 7**,*** regarding accommodating children, also through providing assistance during sessions. This was abided by, as it was made sure that a teacher was present at all times during second iteration workshops, where this guideline would be applicable.

By including teachers in the workshops, the more general guidelines from Spiel at al. (2017) could also be supported with contextual insights: The teachers could both share with us and could – while in interaction with children – draw from their other shared experiences and ways of working together. Namely, also fitting the *guideline 4* on organizing and sequencing, the chosen setting for the workshops, for instance, was one the children were used to doing similar activities in (i.e. moving their bodies in the school gym area) and could recognize.

Similarly, these guidelines from Spiel et al. (2017) helped in evaluating the success and the smoothness of the workshop sessions carried out. Specifically, they allowed us to spot any aspects that could have been improved (e.g. *guideline 2 - distractibility*).

Additionally, prior to the current Master's thesis work, a systematic literature review was carried out, focusing on the current state of autism interaction technologies (Oral, 2022). These insights from Oral (2022) was one other set of guidelines in accommodating the target user group, derived from practices in applied work on the topic out there. The method-related guidelines of these were consistent with the aforementioned ones (e.g. *adjustability* with *guidelines 5 (excessive prompt dependence)* and *6 (strong impulses)* (Oral, 2022; Spiel et al.,2017)). Thus, they served to further highlight the importance of these aspects. Together, the two sets of guidelines were considered in the current work, forming certain themes (see ***Table 2***).

Table 2: Theory that informed methods and process in the current work (i.e. Spiel et al., 2017 and Oral, 2022 *combined, rephrased and made into a table)*.

Theme	Guideline	Explanation
Meaningful participation & contribution during the process	1 - *Concept of Meaning* (Spiel et al., 2017)	Ensure that the child has a meaningful understanding of the research goals and the selected evaluation methods
	Participatory methods/ co-design (Oral, 2022)	To counter technology being created *for* autistic individuals, rather than *with* them, methods that allow for their meaningful contribution can be opted for, for their empowerment.
Distractibility	2 - *Distractibility* (Spiel et al., 2017)	Adapt aspects such as incorporated methods, how questions are framed, etc. to

		fit the child's hobbies and interests, to help with distractibility
Clarity & concreteness	3 - *Concrete vs. Abstract Thinking* (Spiel et al., 2017)	Keep in mind different possible interpretations of questions and instructions that may stem from concrete versus abstract thinking, and opt for clear and unambiguous phrasings, as well as contextual closed questions
Visual cues	4 - *Organizing and Sequencing* (Spiel et al., 2017)	For the different options you make available, provide visual identifiers
Flexibility & adjustability	5 - *Excessive Anxiety/Prompt Dependence* (Spiel et al., 2017)	Make sure to have alternative prompts, etc. at hand for during different engagement moments, to have the flexibility to adapt if needed or preferred
	6 - *Strong Impulses* (Spiel et al., 2017)	Build on the child's own unique abilities and preferences for ways of interaction, rather than forcing a way onto them that they do not like
	Adjustability (Oral, 2022)	Given highly individualized features, needs and preferences of autistic people, it should be possible to alter aspects of procedures (e.g. session duration) as needed, to accommodate these.
Comfort & assistance	7 (Spiel et al., 2017)	Mind the balance of freedom and structure, and accommodate child's needs by flexibly providing assistance (e.g. own therapists, teachers) during sessions

	Comfort, trust & relations (Oral, 2022)	It should be ensured that the participants feel comfortable, for example through establishing a trust relationship through a first meeting with them. Additionally, a safe space can be created through involving familiar people such as teachers, therapists and family members in the sessions.

This former systematic literature review that was carried out, which focused on the current state of autism interaction technologies (Oral, 2022), also included functional guidelines. These served in the current work as ***theory that informs design***.

The design features and elements commonly found in similar example technology systems were fed into the design of the prototypes. For example, the system heavily focused on a *'together use'* – specifically, movement and dance – context. In addition, the workshop set-ups prioritized a *phases-structure*, going from a more guided warm up onto a more open-ended interaction. In this way, the aforementioned guidelines and insights, which were derived from other fields (e.g. music therapy) – as already displayed in ***Table 1*** – were further supported.

4.4 Participatory Sense-Making as Designers

Throughout the participatory RtD process, at each phase, the work was conducted in parallel with other designers that were targeting the same overall goal and question: 'How can neurodiverse (dyadic) social interactions be supported with multi-sensory embodied interaction technologies?'

With this similarity, but also their differences, these projects and processes were a rich breeding ground for insights that could benefit each of the individual projects. Specifically, they each had their specific social interaction context and focus. Furthermore, the sensory modalities the designs featured differed (i.e. movement and sound in the current work, as opposed to haptics and visuals in the other two). As noted before, having this *group* of artifacts that were being designed and evaluated together helped to generate knowledge that otherwise would not be possible to generate (Gaver & Bowers, 2012).

Discussions were held with these three experts, also at some specific moments. This, for example, included validating workshop result interpretations and together brainstorming regarding the next steps in the process (see ***Figure 5***).

Thus, the reader is, here, informed of these processes' role in the current work and investigations. It should, however, be noted that these contributions are not discussed explicitly and in detail in the rest of the current paper.

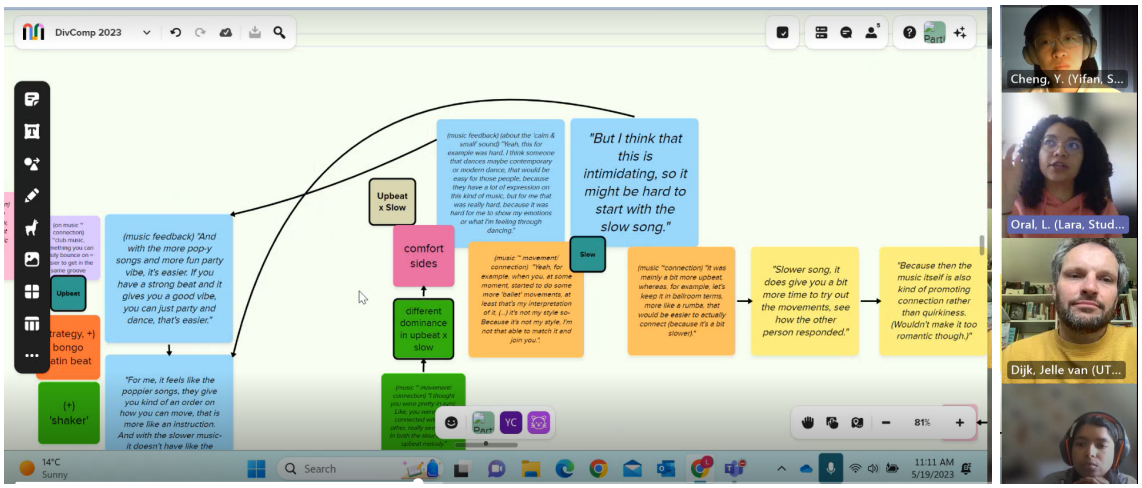


Figure 5: A screenshot taken during one of the (online) expert meetings, in which the results from a workshop were discussed all together.

4.5 Overall Methodology Structure

Overall, the participatory RtD approach utilized in the current work involved two reflective design and research iterations. These involved various different methods, as aforementioned (see ***Figure 6***). Both of the iterations were structured in such a way that they included:

1. seeking of relevant theory and creative inspiration
2. designing a prototype and a surrounding user research set-up,
3. collecting empirical data (through the said research set-up) and analyzing it, and,
4. having an expert meeting in order to discuss and reflect on the obtained insights, and to operationalize them.

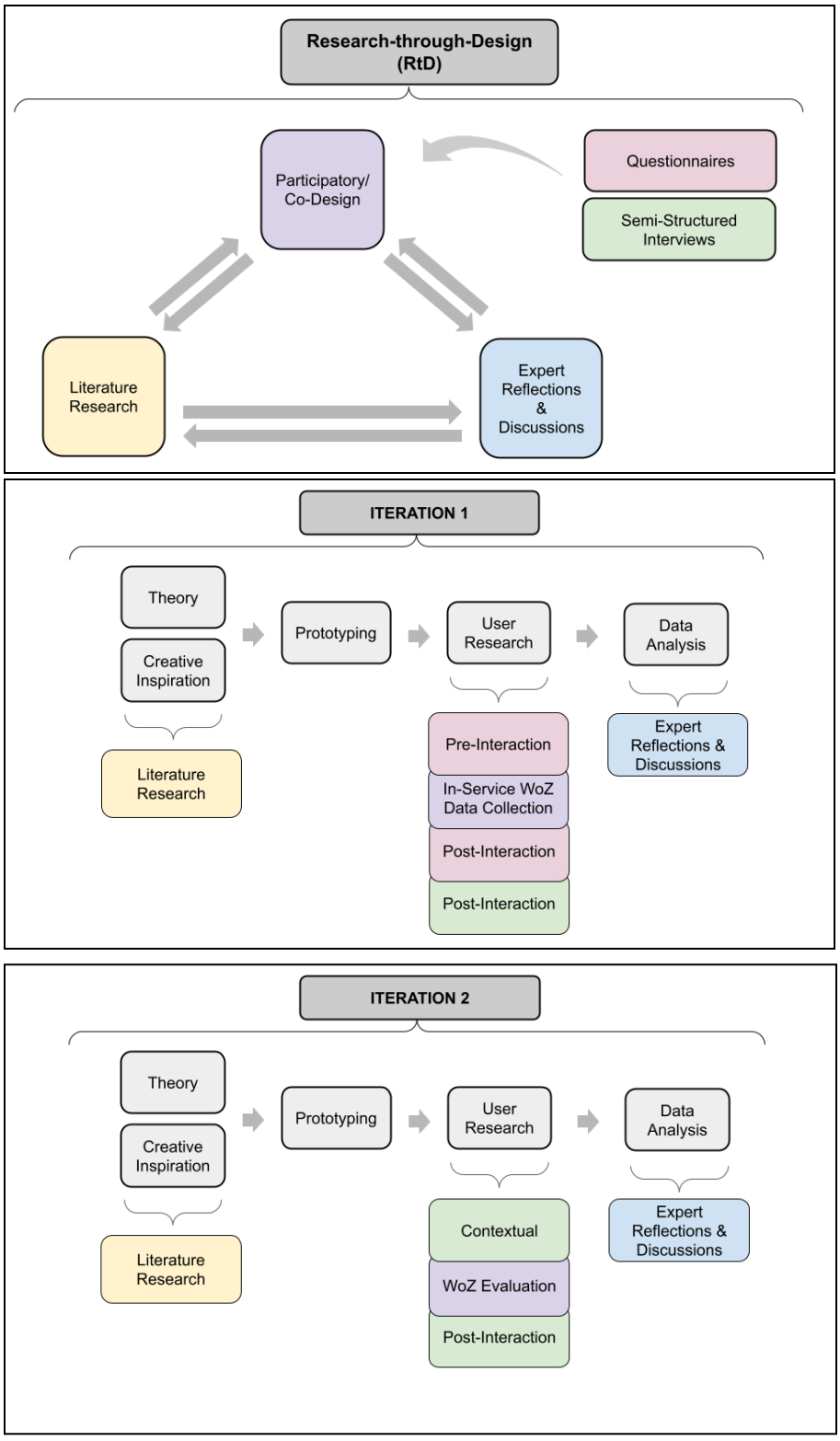


Figure 6: A visualization of the overall methodology structure, where different approaches influence one another and together form an overarching research-through-design (RtD) process *(Up)*. Matching colors between the top and the latter two images show, more specifically, how certain approaches are included within the first *(Middle)* and second iterations *(Down).*

Within this set-up, the first iteration connected to and informed the second iteration. This, for instance, was through the operationalization of the insights; by them getting turned into concrete design decisions. The connection between the iterations, however, was not only in the sense of prototype change and improvements, but also in determining the next most interesting research directions, in order to try and answer the main research questions. In both of the iterations, different methods were utilized to address the research questions (see ***Table 3***).

Table 3: The different methods used in addressing the different research questions (RQs) and in what way, in each of the two iterations.

Iteration #	RQ Addressed	Methods Used	How?
Iteration 1	RQ 1 ***[autistic versus NA]***	- RtD - PD (in-service WoZ)	- inclusion of <u>both autistic and NA</u> dancers and music therapists as experts (in-service WoZ) - dancers' main dance style (i.e. hip hop/breakdance or ballroom/salsa) as <u>a proxy 'difference'</u>
	RQ 2 ***[mappings' connection-prompting role]***		collecting: - types of (common) movements, e.g. to dance together - 'technology system'/DJ strategies - types of sounds and sound changes and whether they are successful or not
	RQ 3 ***[mappings: open/adjustable versus pre-determined]***		^ minding the similarities and differences
	RQ 4 ***[measuring connection/success]***		collecting: - types of observations (used as input or measure of success)

	RQ 5 ***[real-life product]***		-
Iteration 2	RQ 1 ***[autistic versus NA]***	- RtD - PD (WoZ evaluation)	- inclusion of autistic children and their NA teacher
	RQ 2 ***[mappings' connection-prompting role]***		testing based on mappings decided on, given Iteration 1 (WoZ evaluation)
	RQ 3 ***[mappings: open/adjustable versus pre-determined]***		^
	RQ 4 ***[measuring connection/success]***		^ & insights from teachers
	RQ 5 ***[real-life product]***		- contextual inquiry (i.e. a first school visit)

The rest of the paper, in reflection of this methodological structuring, is as follows: First, the first iteration is described, including the initial conceptual system design (see ***5.1***), the design and the implementation of the prototype utilized (see ***5.2***), the participants involved (see ***5.3***), the procedure (see ***5.4***), the data analysis approach overview (see ***5.5***), the results (see ***5.6***) and their analysis (see ***5.7***). Then, the conceptual system design in the second iteration is shared, which follows from the insights obtained (see ***6.1***). The same details as the first iteration are included for this iteration (for all, see ***6. Second Iteration***). Lastly, all of the knowledge gathered from the two iterations are discussed, in accordance with the research questions and including related limitations, as well as future directions (see ***7. Discussion***).

5. First Iteration

5.1 Conceptual Design

Prior to conducting research as a part of this first iteration, theory and creative inspiration were utilized, to come to a first conceptual system design (see ***Figure 7***). Specifically, a system that targets co-located neurodiverse dyadic interactions and their improvement was imagined. Building on the insights gathered from the literature search carried out (see ***Table 1***), this system would play a bi-directional mediator role (Oral, 2022). In particular, it would intervene with the social interaction and the participatory sense-making processes (De Jaegher & Di Paolo, 2007) themselves. This would be instead of targeting autistic social expectations and behaviors, in a way where these are aimed to better fit existing neurotypical norms. The system would achieve this, by introducing a novel technology-mediated interaction to the neurodiverse dyads. The intention would be to get the interactors into dialogical, rather than social system processes (Steffensen, 2012), given this newness element. Specifically, the autistic and NA individuals would be prompted to together make sense of this context and each other in the moment, in place of abiding to and replicating neurotypical social norms (De Jaegher & Di Paolo, 2007).

A dyadic set-up would not only be adopted by the system, but also used as a sort of a 'constraint' (Oral, 2022). This would be in terms of using as input the combined movement data of the interactors: The real-time measurements of their *joint* body movements would be mapped onto the sound feedback they both hear. In this way, the system would act as a sort of a shared musical instrument, which could increase the intimacy of the technology-mediated interaction that takes place (Wigram, 2004).

Inspired by improvisational techniques in music therapy – specifically two that serve a 'structuring' purpose (i.e. rhythmic grounding and tonal centering) (see ***Appendix A***), sounds that fit either the 'rhythm/beats' or 'melody' categories would be included. On top of this baseline, the sound output would be manipulated. These manipulations would serve one of the two purposes; either *connection-reflecting* or *exploration-prompting:*

Sound changes could allow connection or interaction aspects to be reflected onto the music, for the participants to experience these in a heightened way. This would be similar to adaptive music approaches, as utilized as a possible way *reactive embodiment* (Spiel et al., 2016) could be achieved through the audio modality.

Alternatively, these changes could prompt exploration and keep the interaction going, and, in this way, add to the participants' connection. In this case, they would be acting as another layer of input, similarly to the *ambient embodiment* quality these kinds of technologies could incorporate (Spiel et al., 2016).

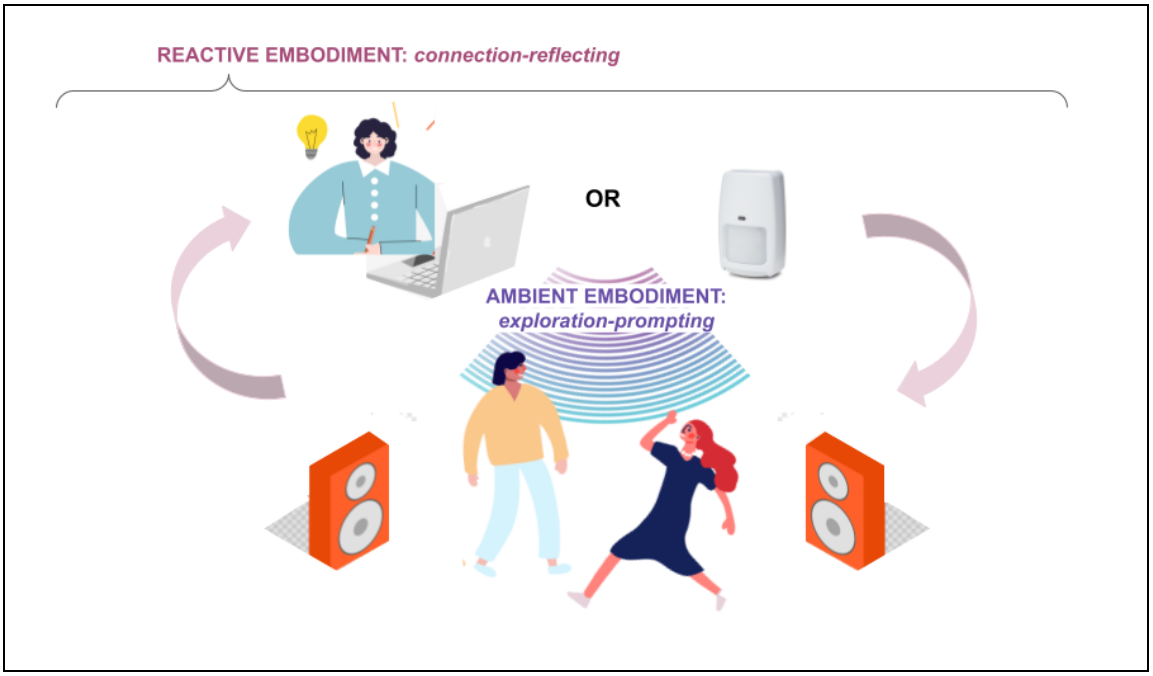


Figure 7: A visualization of the conceptual design for the first iteration. Two types of embodiment (i.e. reactive and ambient (Spiel et al., 2016) are incorporated in the system.

At this point in the current work, within the conceptual system design, the specific joint movement inputs and sound outputs to be incorporated were left open. This was so that experimental data could inform and guide these. Instead, for the design of the prototype to be used during the first iteration workshops, the following quality was highlighted: This prototype would have to make available different alternatives for in-service WoZ data collection (Wirén et al., 2007; Boye & Wirén, 2007). This included sounds that can be played and the sound changes that could be applied by the participants with the 'technology system'/DJ role. The aim of this approach was to discover – through the prototype and in an exploratory manner – which of these would be relevant for our purposes and could be included in the next iteration or the final system design.

5.2 Prototyping

5.2.1 Hardware and Software

For the workshops,

- the Glover software[12] (Version 1.1.2; MI.MU Gloves Limited, 2010),
- the (free) Gliss app (Version 1.0.0; MI.MU Gloves Limited, 2010), and,
- the Ableton Live 11 (Trial) software (Version 11.3.3; McKenzie et al., 2014)

[12] Free access to this paid software was granted to the researcher by one of its developers, Adam Stark, as it was to be used for research, rather than commercial purposes.

were used in combination (for a glimpse at what these software/app interfaces look like, see ***Appendix B***).

Glover was originally created to provide artists with expressive control over the music they make digitally. It allows for composing music using movements and gestures, and supports a range of different sensor-enabled devices' use as controllers, such as, in this case, the Gliss app. Ableton Live 11, on the other hand, was utilized for music creation within the prototype (for the reasons behind all prototype software/app decisions, see ***Table 4***).

Table 4: Prototype software/app decisions and their reasonings.

Prototype Software/App Decision	Why?
Glover (software)	- Supports different sensor-enabled devices' use as controllers and thus, makes possible the extension of the prototype to also have automatic movement detection and consequent sound alteration possibilities in the future.[13] - Is able to send the movement inputs from sensor-enabled devices (such as the Gliss app) onto third-party music software (such as the Ableton Live 11 software), in the form of *musical instrument digital interface (MIDI)* or *open sound control (OSC)* messages. - Supports Max/MSP / Jitter, a visual programming language for music and multimedia. Using Max can allow for possible prototype extensions in the future, concerning automatic movement-to-sound mappings based on sensors. Furthermore, the visual nature of it makes working with it more attainable, with less technical skill requirements.
Gliss (app)	- Is one of the methods supported by Glover for inputting movement data and controlling sound parameters. - ^Has easy WiFi connection (on a phone) to the Glover software (on a laptop). - Provides an immediate and easy-to-use interface with different input options such as

[13] This feature is, however, not relevant for this iteration of the prototype, nor for a research design that uses WoZ for movement detection and system response.

	buttons, sliders, and gestures, which can be assigned to specific ways to alter sounds (e.g. using a specific slider or raising your arm to increase the volume). - ^A suitable alternative for providing the participants with the 'technology system'/DJ role with a reduced-complexity interface.
Ableton Live 11 (software)	- Is one of the third-party music software supported by Glover. - Comes with effects, sounds and other creative features, as well as supporting musical improvisation beyond the traditional linear arrangement, making it possible to add and change musical elements real-time (see ***Appendix B***) - Supports Max/MSP / Jitter.

The Glover and Ableton Live 11 softwares were run on a laptop, whereas the Gliss app was used on a smartphone (for all prototype hardware decisions' details, see ***Table 5***). Additionally, an Internet router was used for the communication in between these. The sessions were video- and audio-recorded. The audio was played using the laptop's built-in speakers.

Table 5: Prototype hardware decisions' details.

Prototype Hardware Decision	Details
Laptop	Intel(R) Core(TM) i7-8550U CPU, 16.0 GB RAM
Smartphone	*iPhone* that has an iOS version 10.0 or later
Internet router	*for secure and stable communication between the laptop and the smartphone*
Cameras	- *Panasonic HC-V720* - *Samsung WB1100F*
Built-in laptop speakers	Realtek High Definition Audio, 48 kHz/24-bit

5.2.2 Sound Design

Due to limitations in the researcher's musical creation capacities and in the interest of time for prototyping, all sound clips utilized were pre-made ones. These were either included within the Ableton Live 11 software itself or taken from online sites. Such sites enabled finding non-copyright music based on different criteria such as the songs' genre or mood (for a full list of the used sound material in this iteration, see ***Appendix C***).

As aforementioned, within this iteration, the aim was to discover the types of sound and sound changes that are relevant for our purposes – in an exploratory manner. Thus, the focus when making sound design decisions was on making available different alternatives for the 'technology system'/DJ to be able to choose from or play around with.

As described, sounds that fit either in the 'rhythm/beats' or 'melody' categories were included, with the latter further including either upbeat or slow melodies. The sound clips included for the 'rhythm/beats' category, for example, differed from one another in terms of their speed, richness (i.e. sounds of how many different timbres they had layered together) or rhythmic patterns. All melodies were instrumental and did not include any lyrical contents. This was due to wanting to focus on other sound aspects that rely less on a 'cognitive' level. What's more, our current knowledge and insights were not mature enough to go for such a level of complexity yet, when it came to what sounds would be 'useful' for our context. Sound clips that belong to the pop genre were preferred, as most participants reported enjoying this (see ***5.3. Participants***).

In addition to giving the 'technology system'/DJ the option to play these different rhythms or melodies separately or layered with one another, an upbeat and a slow melody that were split into parts and could be 'progressed' were included. Specifically, the participants could move on to the next clip – so part – of these songs. Beyond such a 'progression' of the said two songs or layering possibilities, the available *sound changes* were reverb effect, pitch changes, delay effects, panning and volume changes (i.e. only of the melody or of the sound output overall) (for definitions and further details on these, refer to ***Appendix C***). These sound manipulation alternatives were made available, based on them fitting into one of two sound category changes that were pre-defined in the conceptual design (i.e. connection-reflecting or exploration prompting (see ***Table 6***), as per the researcher's own rationale and speculations.

Table 6: Sound changes included in the first iteration and why.

Sound Change	Selection Reason	How?
Song progression	Connection-Reflecting	The sound clip can be moved on to its next section and progressed in terms of its 'storyline', mirroring the progression in the connection being built between the interactors.
Layering		More layers can be added onto the sound playing, developing it and making it richer; mirroring the progression in the connection being built between the interactors.
Reverb effect	Connection-Reflecting or Exploration-Prompting	If less reverb effect is applied, the sound played gets clearer, as if it would be closer in space. Thus, this alteration can mirror the different levels of connection and closeness the interactors feel with one another. In addition, this effect can change the way interactors move to sounds, for example by attempting to dance in a more 'muted'/toned-down way when more of it is in place.
Volume changes (melody-only or overall)		Similarly to the reverb effect, a lower volume can give a sense of being at a distance, as well as loud sounds having more energy. In this way, volume changes can mirror the different levels of connection the interactors feel with one another. In addition, interactors can be influenced in how they move, for example, making larger movements for louder sounds

		and being more quiet in their movements to match lower volume levels.
Pitch changes		Pitch changes can influence the emotive feeling of sounds, for example when they cause atonality that raises an eerie and uncomfortable feeling. These can be made to match the different levels of connection the interactors can feel with one another. In addition, these changing frequencies can inspire different movements of the interactors. For example, more high-pitched and electronic-sounding settings could prompt them to act in a more robotic way.
Delay effects		The delay effects in the software make the sounds appear fuller thanks to the accompanying echoes. Thus, this can mirror the progression in the connection being built between the interactors. In addition, interactors can be influenced in how they move through these echoes present.
Panning	Exploration-Prompting	The distribution of the sound played between the left and right sides of the speakers create an interesting spatial element. The participants can then be prompted to make use of their environment, inspired by this materialization.

5.2.3 Interface Design

The participant or researcher taking upon the ‘technology system’/DJ role during the workshops had two interfaces in front of them: The one from the Ableton Live 11 software on the laptop and the one from the MiMU Gliss App on the phone. These allowed them to access and play with the different sound and sound change options. In the former, so the laptop interface, the different sound categories previously mentioned were put in different columns for easy access (see ***Figure 8***).

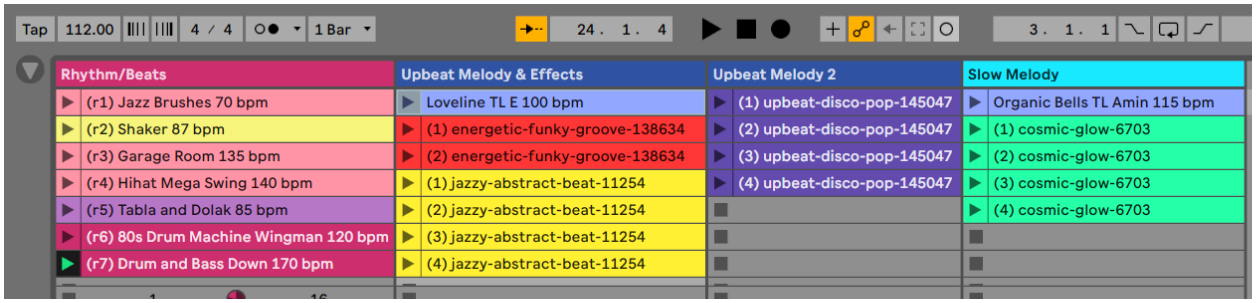


Figure 8: An overview of the Ableton Live 11 (Trial) software (Version 11.3.3; McKenzie et al., 2014) interface included in the first iteration, where different categories of sound clips (e.g. ‘rhythms/beats’) are placed into their own separate columns.

The version of Ableton Live 11 being used during this iteration was a Trial one. Hence, practical constraints that stemmed from this situation impacted certain design choices. For example, the number of individual tracks/sound clips that could be added to each of the columns (i.e. 7) determined the max number of options provided for the ‘technology system’/DJ.

Similarly, the interface on the Gliss app was intended to be kept simple enough. Thus, other practical constraints such as the number of sliders included in that app interface were abided to, when deciding on how many different sound change alternatives were to be included (i.e. 6). Furthermore, these sound changes were impacting a single column (i.e. ‘Upbeat Melody & Effects’) within the Ableton Live 11 interface. This meant that if the participant or researcher decided to use the sound changes on clips other than the ones already within the said column, they would have to move them there as well.

The following were paid attention to, in terms of other visual elements of the interface:

- The columns’ order was organized in such a way that the rhythm and the melody ones would be separate, and the upbeat melody columns would be next to one another (and in the same color), followed by the slow melody column.

- Different or same colors were used to make it easier for ‘technology system’/DJ to be able to recognize sound clips as different or parts of the same one (i.e. progressing).

- Each sound clip was given an associated number in parentheses at the start, to again help with recognition, as well as remembering the sound contents.

The motivation to make use of another interface, on the other hand, stemmed from wanting to reduce the interface and use complexity and the number of sound change options available for the 'technology system'/DJ (see ***Appendix B***). This was despite the participants taking upon this role consisting either of the researcher herself or of music therapists that have had some experience with the software, for ease of control.

5.3 Participants

As aforementioned, during the first iteration, *in-service WoZ data collection* (see ***4.2.1***) was made use of. For this purpose, participants with expertise/experience in either music therapy or dance were deemed suitable and recruited. Additionally, it was aimed to keep this group neurodiverse; a combination of autistic and NA individuals.

Convenience and snowball sampling were utilized, given the researcher's existing relevant contacts (e.g. own membership in two student dance associations, specializing in hip hop/breakdance and ballroom/salsa respectively). Participation was on a voluntary basis and the participants were not provided with any additional compensation.

In total, 10 participants (3 males, 7 females) that included both autistic and NA music therapists, as well as autistic and NA dancers took part in the workshops. Specifically, three female autistic participants (2 dancers and 1 music therapist) were involved (for related limitations, see ***7.6 Limitations***), one of whom also had ADHD[14].

Four individual workshop sessions were held, where in each one, either two or three participants took part. This was due to two participants having the 'interactor' roles and the role of the 'technology system'/DJ being fulfilled either by one other participant or the researcher herself, who is female, NA and has dance experience (see ***Table 7***).

[14] All of these labels were based on self-report, given the difficulties people face upon getting an official diagnosis (e.g. long waiting times, expenses, etc.).

Table 7: Workshop sessions' breakdown in terms of participant amounts and gender-, expertise/experience- and neurodivergence- division.

Workshop #	Participant #	Role	Gender	(Main) Expertise/ Experience	Neurodivergent ?
W1	P1	Interactor	F	Dancer (Ballroom/ Salsa)	Yes, autistic
	P2	Interactor	M	Dancer (Ballroom/ Salsa)	No
W2	P3	Interactor	M	Dancer (Hip Hop/ Breakdance)	No
	P4	Interactor	F	Dancer (Ballroom/ Salsa)	No
W3	P5	Interactor	F	Dancer (Ballroom/ Salsa)	No
	P6	Interactor	F	Dancer (Hip Hop/ Breakdance)	No
	P7	'Technology System'/DJ	F	Music Therapist	No
W4	P8	'Technology System'/DJ	F	Music Therapist	Yes, autistic & ADHD
	P9	Interactor	M	Dancer (Hip Hop/ Breakdance)	No
	P10	Interactor	F	Dancer (Ballroom/ Salsa)	Yes, autistic

The participants all fell into the 18-44 age range. In terms of musical experience, as well as musical habits and preferences, the participants had their similarities and differences. For example, the type of music they listened to varied (for more details, also regarding other demographic information, see ***Appendix D***). The top-5 genres listened to by the participants were pop, rock, classical music, hip hop and electronic music (see ***Table 8***).

Table 8: The top music genres listened to by the participants in the first iteration.

	Pop	**Rock**	**Classical**	**Hip Hop**	**Electronic**
Popularity Ranking	1	2	2	3	3
# Participants	8	6	6	5	5
Participants	P1, P2, P3, P4, P5, P6, P8, P10	P1, P3, P5, P7 P8, P10	P1, P2, P4, P5, P6, P7	P1, P3, P4, P6, P7	P1, P4, P5, P6, P7

All participants had dance experience and they were recruited and characterized by mainly focusing on and having experience either in *hip hop/breakdance* or *ballroom/salsa* dancing. However some participants had experience in both (see ***Table 9***), as well as some other dance styles such as contemporary/modern dance, ballet and physical theater.

Table 9: The division of participants' (main) dance expertise/experience areas in the first iteration.

	Only Ballroom/Salsa	**Only Hip Hop/ Breakdance**	**Both Ballroom/ Salsa & Hip Hop/ Breakdance**
# Participants	1	1	6
Participants	P2	P6	P1, P3, P4, P5, P9, P10

5.4 Procedure

The participants were recruited and divided into different workshop sessions according to the sessions' neurodivergence (i.e. autistic versus NA) and main dance style expertise/experience compositions. Following this onboarding, all of them were sent an information letter (see ***Appendix E***) electronically, to be read in advance. This was to allow for

enough time to consider their participation, as well as to save time on the day of the experiment. Furthermore, they had to already fill in the descriptive questionnaire that included questions on their musical and dance-related experience, habits and preferences and personal information. Especially their preferences in terms of music genres were important, so that the musical material provided for the 'technology system'/DJ would fit what the majority found enjoyable – which, in this case, was pop music. Along the same lines, they were also asked about any (especially sound-wise) sensory sensitivities they may have that should be taken into account. For one autistic participant, this included sounds that are very high pitched or loud, so this was minded within the workshops.

On the day of the workshops, upon arrival, participants were welcomed and asked if they needed to review the printed-out information letter prior to starting the interaction. Following this, both the participants and the researcher signed the consent form (see ***Appendix F***), of which the participants were also given a signed copy. Afterwards, the procedure of the study was also explained to them verbally and they were given the opportunity to ask any questions they may have.

If applicable, so if a third participant had the role of the 'technology system'/DJ, they were shown around the prototype. This was to familiarize them with the interface and the different sound and sound change options they could make use of during the session. A decision was made to also have the other two participants present during. This was in both cases of the researcher or another participant having the 'technology system'/DJ role. Therefore, all participants would have also heard each sound at least once. This could help to combat novelty and (especially autistic individuals) feeling discomfort due to unfamiliar sounds. They could be given a say in having any unpleasant or uncomfortable sounds removed from the options provided to the 'technology system'/DJ. The 'technology system'/DJ was also provided with template sheets that included tables and additional blank spaces for them to be able to make notes on, if needed. This, for example, could help them to more easily characterize or to remember different sound clip options for themselves. After they had been sufficiently familiarized with the prototype and had no further questions, the session moved on to the warm-up.

The activity selected for the warm-up was to have one participant, one of the interactors, take the lead in moving to the music in a way that feels comfortable, natural and enjoyable to them (for the exact wordings of the instructions given to the participants, see ***Appendix G***). The other, on the other hand, would follow their lead and do the same and mirror their moves. The 'technology system'/DJ would provide the music and instruct the interactors to switch the leading-following roles whenever they deemed suitable. Through the sounds they play and the sound changes they make, their goal was to help the interactors in building a connection through moving together.

Following this warm-up activity and a small break came the main interaction, where, now, both of the interactors were asked to move in any way that felt comfortable, natural and enjoyable to them. They could listen to the music, look at each other's movements if they wanted to, but without any specific instructions or constraints to connect with one another (in a certain way). The role of the 'technology system'/DJ, on the other hand, remained the same and was to help them in building a connection through moving together.

Participants of all roles (i.e. 'technology system'/DJ and interactor) were video- and audio-recorded in both the warm-up and the main interaction, with consent.

After this main interaction had also come to an end and the participants were sufficiently rested, the audio recording was started again. The researcher conducted a semi-structured interview with the participants. On average, this discussion part lasted about 10-15 minutes, after which the participants filled in a 5-to-10-minute online survey that matched the roles they had during the interaction (i.e. interactor (see ***Appendix H***) or 'technology system'/DJ (see ***Appendix I***)) Each of the workshop sessions lasted for at most an hour and a half in total.

5.5 Data Analysis

The first iteration involved the following data sources that could or needed to be analyzed:

- interactor and 'technology system'/DJ post-questionnaires,
- post-interaction semi-structured interviews' audio recordings, and,
- (warm-up and) main interaction video recordings.

For the post-questionnaires, the online interface of the questionnaire provider (Jotform (n.d.)) was made use of, to look into the gathered data. This was both so that certain mean and standard deviation values could be calculated, and to add any other interesting insights from these onto the results from those of the semi-structured interviews.

The interview recordings were first manually transcribed and mapped onto a Mural (n.d.) board (see ***Appendix J***). Then, thematic analysis was conducted, creating another board), using Braune & Clarke (2006)'s methodology as a guide

Specifically, based on the developed interview questions, broad themes were decided upon. These acted as codes for analyzing the transcript contents. Furthermore, the semi-structured nature of the interviews and the exploratory attitude of the first iteration workshops were taken into consideration. Consequently, the data was analyzed to also include any emerging themes not covered by these predetermined codes. The codes could have sub-themes and were, at last, also grouped under major themes for their interpretation. These were derived based on the research questions, as seen below:

- *[RQ 1] Autistic versus NA,*
- *[RQ 2] Mappings' connection-prompting role,*
- *[RQ 3] Mappings: Open/adjustable versus pre-determined,*
- *[RQ 4] Measuring connection/success,* and,
- *[RQ 5] Real-life product.*

Within this iteration, video recording data was not analyzed in-depth, but was rather used as supplementary material. This involved, for example, being able to locate any critical events the participants mentioned in discussion. In this way, the presented findings could be supported with the accompanying visuals.

5.6 Results

Participants with the interactor role (N=7) rated the interactions on a 5-point scale as comfortable (*M* = 4, *SD* = 0.58), natural (*M* = 4.43, *SD* = 0.98), as well as enjoyable (*M* = 4.14, *SD* = 0.38).

As aforementioned, the results obtained from the semi-structured interviews were also grouped into and analyzed through several themes and sub-themes (see ***Table 10***). Within the following subsections, these are presented one-by-one. In each of them, associated raw participant quotes are also included.

Table 10: Overview of themes and sub-themes the first iteration results were grouped into.

Theme	Sub-Theme(s)
Felt connection	Warm-up versus main interaction
	Dancing together versus alone
Strategies and indications to dance together	Eye contact
	Mirroring
	Lessened distance
	‘Matching’
	Physical contact
	Exploring other ways
Sounds and sound changes	Volume
	‘Change’ and unexpected sounds
	Storyline, expectation and anticipation
	Musical connection
	Upbeat versus slow music
‘Technology system’/DJ strategies	Frequency changes
	Repetition without enjoyment
	Melody versus rhythm
Sound-related suggestions	Musical layers
	Personalization, choices and dynamics

Context-and-procedure-related suggestions	Comfort zones and expression
	Interaction setting
	Target user group: children

5.6.1 Felt Connection

Following the workshop sessions, the participants talked about whether and during which part of the session they felt a connection; when they felt as though they were in interaction and dancing with the other person.

- Warm-up versus main interaction

Upon discussing how the mirroring warm-up exercise compared to the more open-ended main interaction, one point the participants brought up was how having the other in your space impacted things. This had an influence, even without an official constraint to make a connection with each other – such as, through the mirroring exercise. Specifically, **P6** stated, talking about the main interaction, that *"You could do your own thing, but at the same time, it was also difficult to ignore each other, like, the presence."*

According to a 5-point scale, all interactors could make a connection of similar strengths (M = 4.14, SD = 0.69) through this interaction. However, there was a divide between the participants, when it came to which part of the session they felt more connected in:

Half of the participants felt that the ***main*** interaction allowed for a stronger sense of connection. This seemed to relate to having the *choice* regarding and *control* over whether or not to connect. **P5**, for example, made the comment:

"I would say that we even had more connection (...), because then [in the warm-up], you are 'supposed to' (...) and here [in the main interaction], it was more like 'Do I want to make a connection?'. Yes, (...) I want to, and then we were looking for it. It felt more– I feel more connected."

Another aspect was the open-endedness allowing for more – and, specifically, more interesting – ways of connecting. This included getting to see the other person's own style and manner of moving. **P10/A**[15] highlighted:

"I really liked that maybe you, like, mirror, but when I do a turn, you do a turn in your own way. That's way more interesting than trying to do exactly the same as the other person. Cause it also tells something about the person."

[15] In the reporting of the results, for easier viewing and later interpretation (i.e. not having to reference back as often) autistic participants were marked in the text in this way. This aimed to make easier the comparisons between autistic and NA individuals' actions and opinions.

Thus, participants not only derived more enjoyment from 'also doing more of their own stuff' (such as **P5**), but also enjoyed this *revealing* aspect: Through the interaction, they could get to know each other in this different movement sense. This, for example, included their tendencies and preferences in what kind of dance moves they go for. In addition, their reactions to different sound inputs and changes were another point of disclosure with one another.

What's more, the participants talked about being able to see the *person*, rather than their bodies and movements in a sort of an isolation, thanks to the freedom in the main interaction. According to **P9**, there, his attention could be on the person and their movements as a whole. This allowed 'dancing together' to have a broader definition and meaning. This, he explained by saying,

"*You don't really have to do the same motion. You can also do motions that complement the person, which makes it so that I don't only have to be looking at her hands or her feet. I can just look at her as a person, instead of certain parts of her movements."*

P10/A agreed that *"Yeah, you can just pay attention to your energy, instead of what your body is doing."*

Half of the participants, on the other hand, highlighted the benefits the ***warm-up*** offered. For example, participants could derive comfort from the 'togetherness' constraint. **P3** said, on the topic, *"The second part felt more free, but I think sometimes you just want to match the person's energy. And yeah, if you mirror the dance move, I think it helps to be in the same… Yeah, I don't know, same vibes. I don't know how to explain it <laughs>."*

Commentary from **P4** was also in support of this. The novel interaction context was scary. However, the other's presence and especially the constraint to do the same moves allowed less attention to be given to the 'optics' or this discomfort:

"Yeah, I think, when you mirror each other, the first time it's intimidating because you are really doing the same moves. But, it also gives you a bit of comfort, because you are not alone. And I feel like if you are doing the same moves, it's not weird to do the moves. [Versus when alone you may think] 'Okay, maybe it looks weird.' "

This clear interaction structure and division of roles, therefore, helped them in not focusing on themselves and how things may look. **P2** made a relevant comment, pointing out *"She is going to copy me, so if I do strange or sudden stuff, the assignment is to have to copy."*

Furthermore, this constraint within the warm-up also meant that there existed an already established 'connection'. You had the same expectations as your fellow interactor. Specifically, you knew how you were to act and connect with one another. In the words of the same participant, **P2**,

"The warm-up had the additional benefit that you had to make a connection to your partner. Which, at this moment [in the main interaction], was a decision from both parties; where some

moments you connect and some moments you take to yourself. And that differs. (...) *Because of the constraint, you already knew which language she is going to speak. She is going to copy me."*

In comparison, within the main interaction, both interactors had to decide for themselves when to connect or not. They had to indicate these intentions, initiate the consequent actions and somehow come to an alignment, which was mentioned to be a struggle. **P1/A** pointed out that *"Yeah, I feel like sometimes I also felt very uncertain about connecting. (...) The starting was sometimes, like, 'How do you start dancing together again?' I especially like taking some moments to myself. But, at other times, it was like 'Hmm, I would have liked to connect with my partner here, but (...) I don't really know how.'"*

As **P2** put it, the participants *"would like to know now [if they are] actually dancing with someone, instead of 'I'm dancing by myself.'"*; whether they are connection-building at that moment or not.

- Dancing together versus alone

When asked regarding instances of having the sense that they are dancing with the other person versus not, the participants appeared to less often feel as though they were dancing alone. **P10/A**, for example, said *"We sometimes just tuned out, but most of the time, we at least knew where the other person was."* Similarly, **P5** commented that *"Even though at some points, I really focused on myself, those were short parts, I would say."*

This did, however, bring forth the question of what it even meant to be dancing together. As per **P4**'s reflection, this was hard to put into exact words. Instead, it was a sort of a feeling you would have: *"I mean, we did not touch each other, but, like, I had, at some point, the feeling that we would dance together, not alone. (...) Or at least that we enjoy the same vibe, energy."*

Yet another important point seemed to be that both interactors needed to be open to dance together or connect in this way to begin with. **P5** highlighted:

"It's about your own attention or intention, or actually both. At some moments, I was really just focusing on myself and doing my own stuff. Then, when I would want to make contact, I would focus more on you, trying to get your attention."

5.6.2 Strategies and Indications to Dance Together

To dance with one another, the participants seemed to apply various converging strategies. However, **P1/A** *"ironically felt like it's easier to clearly express that you don't want to dance together, rather than show with your body that you want to do so."*

- Eye contact

A major indicator of wanting to dance with the other – that was mentioned by almost all of the participants – was eye contact. **P4** thought this to be "*the first thing."* This tendency was understandable, given the current neurotypical societal norms that favor engaging with one

another in this way. Despite this, perhaps also as expected, autistic and NA participants diverged on the matter to a certain degree:

Interestingly, **P1/A** attested to copying **P2**'s movements without making eye contact, in order to make a connection and to dance together. However, following her failed attempts at achieving so, she pointed out that *"That's why it's hard. If I start copying you without the first [eye] contact, the other person doesn't notice that you are trying to make contact."*

On the other hand, **P10/A** talked about how eye contact within an everyday conversation versus in a dancing context, for herself, was different. She experienced it to be less overwhelming and more helpful now, given that it was a main way you can read someone's energy:

"When I'm talking, it [making eye contact] is difficult and I'd rather look at someone's mouth. Because it's easier to follow what [they are] saying. But in dancing there is no– The only input you get is someone's movements and someone's energy. And like, [for] energy, you can really look into the eyes. And you can also see in the eyes, what someone is going to do, actually. (...) You are focusing on energy and not on something someone wants to communicate to you."

- Mirroring

Another strategy that came up when trying to dance together was mirroring (see ***Figure 9***). This was as expected, also due to the chosen warm-up exercise. It was a way the participants already knew how to get each other's attention. For example **P3** said on the topic:

"Yeah, if someone is doing a dance move and you consciously mirror them, you are seeing that person and trying to– You are saying 'I like that move and I'm going to do it with you.'"

Figure 9[16]: Example instances of *mirroring* as a strategy and indication to dance together, from three different workshops. The first, middle and last rows correspond to **W1**, **W2** and **W3** consecutively.

The participants, such as **P5**, shared how and why this enabled connecting with the other person. It could keep the dance interaction going for them, because *"Like partying, it feels very much like a social interaction. You see people doing cool moves and you are like 'Oh, this is really nice, I can copy them.' and you try to, after copying them, see how you can make it your own."*

P4 highlighted one use of adjusting to each other's movements, which was that it *"decreases the barrier of dancing together, because if you do similar moves, (…) you feel not alone."* The meaning behind this action, what they believe they are able to convey through it, was also further explained by **P3**, who said *"Yeah, [when the other person mirrors your moves] you feel seen as well."*

- Lessened distance

What one participant, **P5**, referred to as *'anti-mirroring'* was yet another common approach to attempting to dance together with the other person (see ***Figure 10***). **P9** put it very bluntly, when asked regarding how one can indicate their intention to connect: *"This is gonna sound very simple, but just move closer."* This point was agreed upon by other participants as well. Speaking of such indicators, of showing wanting to dance together or not, **P1/A** exemplified and explained,

[16] The participants' faces are hidden in the figures included, in order to protect the privacy of their data and personal information.

"I think the decreasing [of] the distance significantly? Cause I felt like the opposite was also true, that sometimes one of us would take a step back to dance by ourselves. And by stepping closer again, you give an indication of 'Oh, I wanna dance together.'"

In this regard, participants, such as **P3**, not only highlighted moving towards or away from one another, but also other body language aspects. This included where your body was pointed towards: *"And maybe also standing in front– Yeah, just facing them. I think it's an indication that you are dancing with them, right? Or trying to."*

Another related strategy had to do with the participants' use of the space and their positioning (see ***Figure 10***). **P1/A**, in this regard, claimed *"I feel like dancing together doesn't always need to be mirroring. We were also dancing around one another. That was nice."*

Figure 10: Example instances of *lessened distance* – and related body language and spacing aspects, such as going around one another – as a strategy and indication to dance together, from two different workshops. The first and second rows correspond to **W3** and **W4** consecutively.

- 'Matching'

Whether it was their energy, or as **P6** put it, *"Matching the pace, rhythm, or something. Matching small things."*, a (perceived) *alignment* was yet another indication of dancing or wanting to dance together.

On this end, participants seemed to take into consideration each other's capacities. This helped reaching a common ground to become more attainable. While explaining how they acted in the interaction, **P10/A** shared *"I tried to, because I know his background, kind of do things that you can expect someone to follow."* This sort of matching was seen as a sort of requirement by **P1/A**. She said *"Yeah, (…) finding the styles that you match, where you both feel comfortable… Otherwise you can't do it together."*

Surprisingly, an aspect that hurts the ability to connect came up in all but one of the sessions: If one person did 'cool moves' belonging to a certain unfamiliar style – in this case, especially breakdance – that made it difficult for the other to 'match' them. **P2**, on the topic, pointed out,

"Yeah, if someone would do full–on hip hop in front of me, like– Cool! Nice! I'm liking this. I have no clue how to do any of this. But cool, nice for you."

This was an experience shared also by **P10/A.** She reflected back that *"When he did cool breakdance things, I was like 'That's cool, I wanna give you the room to do it.' So you wanna give the person the 'shine' (...). [But] for me, it's something that separates."*

This point carried more weight for participants, such as **P2**, that had dance experience of only a specific style: *"Yeah, for example, when you, at some moment, started to do some more 'ballet' movements – at least that's my interpretation of it – it's not my style. So, I'm not that able to match it and join you."*

Some others, on the other hand, pointed out themselves that their capacities in dancing did not begin and end only in the dance style they more regularly practice. This seemed to help to alleviate such disconnection effects to a certain degree. **P4** claimed, for instance,

"I think he also looked if I could follow the moves, so then it was not too hard for me to follow those. I mean if he would have done breakdance, I would be out. But I didn't feel like I missed something, because I mean, I can also dance outside of the ballroom bubble."

Overall, there were convergences and divergences in the ways the participants attempted to dance together. This, as well as their differing perceptions of certain contexts' meanings, made us revisit a question: *What does it mean to dance together*?

P10/A elaborated on the other person doing 'cool moves' that are not easily followable and how this, for them, hurts the connection. Specifically, she explained *"But that's also because I come from partner dancing. That's just slow, total body contact, just feeling each other's energy."* The discussion that followed, involving **P9**, indeed revealed the existence of at least two different frameworks of 'dancing together' within this group of participants:

In ballroom dancing, 'dancing together' is quite literally that you are in close physical hold. In breakdance culture, on the other hand, there exists 'ciphers': Dancers surround each other in a circle-formation and take turns going inside. In this case, 'dancing together' also involves watching each other's personal dance moves and being on the side, supporting one another.

- Physical contact

Despite this being seen only in two sessions (i.e. **W1** & **W3**) (see ***Figure 11***), participants mentioned physical contact as another strong indication of dancing or wanting to dance with someone. Here, they mentioned the way the warm-up exercise, mirroring, impacted them in not opting for this approach. **P1/A**, for example, shared that *"[Physical contact wasn't] necessary actually. During the mirroring, it wasn't necessary for connection."*

This situation manifested in an interesting way in how **P1/A** and **P2** interacted with one another. These participants had former experience in dancing together, also in physical contact and hold. Despite this, during the session, they also did a no-physical-contact version of a Latin/ballroom dance style; cha cha cha.

Later, this interaction pair, as well as those in **W3** were seen dancing around, holding hands – with the former taking part in another established dance style of merengue, due to hearing appropriate counts for this in the music played (see ***Figure 11***).

Figure 11: Example instances of *physical contact* as a strategy and indication to dance together, from two different workshops. The first and second rows correspond to **W1** and **W3** consecutively.

- Exploring other ways

The participants found that there existed more (interesting) ways to dance together, also beyond the mirroring exercise featured during the warm-up. They expressed that they enjoyed exploring these together. **P5**, for example, shared an example, saying *"Because at some points, I tried turning my back to you and you started doing what I was doing."*.

It seemed that this interaction and the following discussion encouraged them in thinking of how they can connect in this sort of a movement context. This was inferred, for example, from **P1/A** ideating and making the observation that *"It [a connection attempt] can also just be someone trying to make a movement in your direction, repetitively."* **[P1/A]**.

5.6.3 Sounds and Sound Changes

In addition to the movement and interaction aspects the system could use as inputs, insights were gathered regarding the sound and sound change outputs it could incorporate. In particular, several sound aspects were observed that made connection-making either easier or more difficult.

- Volume

During the sessions, the built-in laptop speakers, rather than external ones, were relied upon. Furthermore, the participants with the 'technology system'/DJ role were, of course, in charge of the sound change decisions, which also included volume changes. On occasion, if they lowered the volume down too much or if the speakers' quality intervened in this sense, this appeared to pose a challenge for connection-making. On the topic, **P5** pointed out that *"At some points, the volume was quite low. Then, I was mainly just focusing on the music and then, of course, you focus less on dancing together."*

Changes in volume could, however, also play a prompting role. For instance, this sound manipulation was used by **P7** to match the energy levels of the participants in interaction, enhancing these feelings.

- 'Change' and unexpected sounds

Similarly to how the music not being loud enough caused participants to direct their attention and efforts onto this aspect, music changing too frequently also made it difficult for them to focus on each other and the connection-making itself.

This was especially the case in **W1**, also due to this session being the first one. There, the researcher herself played the role of 'technology system'/DJ. She has experience in dance and knowledge in the theoretical bases for the technology system. Regardless, it was obvious that in comparison, the two other 'technology system'/DJs with music therapy expertise knew better certain practicalities, from the get go. This included aspects such as how frequently they should introduce changes in music during the workshops.

Regardless, overall, the participants were not under the impression that the current set-up included too much musical variety. On the topic, **P4** commented: *"I did not feel like there were too many or too few changes, so I think it was appropriate."*

According to one of the 'technology system'/DJ's (**P7**), change could additionally play a connection-prompting role:

> *"I also observed that when I introduced a sudden change in the music, that most of the times, there was a reaction. You also changed something in your dancing. Either just in the dancing or also in the contact."*

This fit the interactors', such as **P5**'s, own observation and insights, whereby a musical change posed an opportunity to dance together: *"When there was a change and we both felt the change, it's also a good moment to start vibing together."*

A specific interesting element here was the sounds played being unexpected. In this way, they could give rise to a raw and authentic reaction from the participants. **P6** shared *"I found it really nice when you consequently put a 'tsk' sound. I didn't expect it and then you get a pure reaction from us, which was really cool."*

As mentioned above, these unexpected sounds could assist participants in getting to know each other in a different sense and to make a connection, due to their reactions' revealing

quality. A specific instance of such was of **P2** and the *'bowling part of the song'*, where he imitated this action. He said that the played sound "*is something typically you are not able to do much with. Because of that specific sound, it really asks for a specific action."* While for him, with this specificness of the sound seemed to *"kill all else off"*, his interaction partner, **P1/A**, had a different perspective:

"So, too quirky sounds are not promoting connection. But on the other hand, I did like that you made that connection. Like, that shows a piece of you. And that made me laugh. It was nice and that made both of us laugh, which made a connection, right? Having a little bit of that and allowing the other person to see how you respond to certain unpredictable cues does show a lot of personality, which I think is very valuable." **[P1/A]**

- Storyline, expectation and anticipation

One other sound aspect mentioned was complementary to the positive impact unexpected sounds can have: If the played music had a sort of a 'storyline', participants could build expectations regarding and anticipate what will come next. This was a point that came up in all but one session. **P1/A**, for example, described what she thought the music lacked as *"longer, interesting beats that return at a certain interval & being able to anticipate them and doing something fun with them"*.

According to **P6**, even with unexpected sounds (like the *'tsk'* mentioned above), it seemed that part of the fun had to do with this tension of whether the participants can anticipate the next instance. She said *"I was kinda waiting [for] when you are gonna do it again."*

Participants, such as **P10/A,** explained how they *"wanted the music to tell them what to do"*. They highlighted the importance of the music telling a sort of story and giving a sort of 'guidance' to their movements. **P9** rationed,

"Everytime, what I rely upon is just the progression of the music itself. As the progression becomes faster, I know like 'Oh, every eight counts, a change will happen.' or I know 'Okay, the song is wrapping up, a drop will be happening soon.' Or 'Oh, it has done this specific thing twice in this music already–'"

With regards to this musical storyline and movement connection, **P10/A** also agreed, saying *"That's why I like choreographing so much. Because when I hear a song, I directly have parts of that already in my mind. What needs to be where, what needs to be happening. Because it's logical, the song asks for it. People don't always understand that <laughs>. But it's just how it is."*

- Musical connection

Regardless of whether or not there existed a storyline within the played tunes, the participants (such as **P4**) highlighted how they "*get a vibe from the music and then [they] either have a feeling [of/for] how [they] can move to it or not really a connection to the music. And then, it's harder for [them] to move."* Having this connection to the music for themselves appeared to be a first requirement for the success of our intentions. Otherwise, the participants

could not connect to the other person present, through the musical interaction. **P1/A** shared, for example, that she *"also felt [that] with many songs, it was difficult to try and connect with your partner, because it was already hard to figure out how you were going to move yourself."*

The participants were asked about when they indicated to one another their interest in dancing together or felt as though they already were moving in collaboration. They appeared to interpret these moments in the following way: These behaviors showed how the other participant and themselves could get 'the same feeling' out of the sounds playing. In the words of **P4**, *"[We were mirroring each other] at some points where the move made sense. (...) It was just feeling it… And obviously you also felt the same."* Hence, the participants' individual connections to the music acted as a bridge for their connection with each other in this way, contributing to a sense of relatedness.

- Upbeat versus slow music

One dimension the sounds played by the 'technology system'/DJ diverged upon was whether they were upbeat or slow. **P8/A**'s observation was that both could serve the intended purpose of allowing for connection. On the topic, she said: *"I thought you were pretty in-sync. Like, you were pretty connected with each other, really seemed so. In both the slow and the upbeat melody."*

The interacting participants themselves, on the other hand, agreed that some song choices made it easier for them to dance (together) and to make a connection. Regardless, their preferences as to what kind of music supports them better in this seemed to differ:

In defense of ***upbeat*** tunes, **P6**, for example, pointed out that something you can easily bounce on makes it easier to *"get in the same groove"*. She likened this sort of music to what is played in clubs. Similar party-context references were also observed in other participants' rationale. **P4**, for instance, claimed *"And with more pop-y songs and a more fun party vibe, it's easier. If you have a strong beat and it gives you a good vibe, you can just party and dance, that's easier."*

What's more, discussions were held with the 'technology system'/DJs, regarding what they thought to be successful sounds, based on their attempts. There, two of these participants mentioned upbeat examples: The *'bongo Latin beat'* and the *'Shaker'* (see ***Appendix C***).

To counter some participants' liking of upbeat music, **P10/A** made a strong statement against this. Specifically, she said *"For me, in fast music, that [the feeling of connection] totally dissipates. (...) I don't think you can make a real connection on fast music."* Other participants also agreed to the benefits ***slow*** tunes had to offer. **P1/A**, for example, explained that *"Slower song– It does give you a bit more time to try out the movements, see how the other person responded. (...) Because then the music itself is also kind of promoting connection rather than quirkiness."*

As mentioned earlier, one participant with a focus on a single dance style, **P2**, displayed another case of preferring one or the other. In particular, he made a connection to a certain slow Latin/ballroom dance style that he especially enjoyed. This seemed to, at least partly, provide a basis for what music he thought best enables connection:

“Whereas, for example, let’s keep it in ballroom terms, more like a rumba… That would be easier to actually connect because it’s a bit slower.” **[P2]**

Also beyond just **P2**, the backgrounds of the different participants informed their opinions and preferences in this sense. They brought this point up in discussion themselves, with **P4** referring to slow music by saying *“Yeah, this, for example, was hard. I think someone that dances, maybe, contemporary or modern dance… That would be easy for those people, because they have a lot of expression on this kind of music. But for me, that was really hard, because it was hard for me to show my emotions or what I’m feeling through dancing.”*

5.6.4 ‘Technology System’/DJ Strategies

Insights regarding movement or interaction inputs, as well as musical outputs the system can feature were informed, also specifically by the actions and rationale of the participants with the ‘technology system’/DJ role.

- Frequency changes

One of the ‘technology system’/DJ’s, **P7**, mentioned changes in frequency as a strategy she tried out and found to be successful. This matched the observation of the researcher herself, who was in the session as a bystander.

Despite this clearly observed effect by both during the session and also in the associated part of the video recording, neither participants were consciously aware of this situation. **P5**, for example, said *"I don't think I was particularly aware of the frequency changes, so [it impacted things] not so much."* They actively noticed this sound alteration only once, when **P7** made a quite slow and gradual change. Regardless, there appeared to be a subconscious impact that they could also attest to. **P6** followed up with the remark *"I don't know if I noticed that kind of stuff. Subconsciously, I'm not like 'oh the pitch changed, so now I should...', but, like, I feel it."*

- Repetition without enjoyment

Yet another strategy **P7** utilized as a movement or interaction input was to use her own sense of time and boredom. She looked for signs that the participants may want or need a change in the ongoing tunes, to keep on being engaged. In her own words, something had to change *"when I saw that you repeat the moves, but had the impression that you are not enjoying repeating it, but just repeating because there is so much repetition [in the music]."*

- Melody versus rhythm

As far as movement-to-sound-mapping strategies go, the other ‘technology system’/DJ, **P8/A**, took a more experimental approach. Rather than making and going with certain (implicit) assumptions, she wanted to first spend time observing how the two interactors would respond to different changes in different or similar ways.

One important difference across different participants related to the upbeat versus slow music dichotomy: **P8/A** explained *“I tried to experiment with different rhythms under the*

melodies, because I wanted to see how much that influenced [their movements and interaction]. (...) some people are more prone to melody, some people are more prone to rhythm."

Some of the differences here were again rationalized by the participants themselves to relate to their individual (dance) backgrounds and what they are used to. For example, **P9** mentioned, in relation to his experience in breakdance, that *“as soon as [they] hear a bass, [they are like] <clap> ‘I can gravitate towards that.’”*.

Despite these predispositions, however, it was highlighted how a given piece of music could make someone dance to either one of these elements: **P10/A** would normally dance to the melody, but she said that *“sometimes in these types of music, there is so much beat that you can’t ignore it”*.

5.5.5 Sound-Related Suggestions

The participants made certain additional suggestions themselves, for example on how the sound material included could be changed and improved.

- Musical layers

Some participants commented on how individual aspects of music, such as rhythm, can sometimes be difficult to dance to, when on their own. **P3** commented that *"Those elements [like rhythm] are separate, so it's a little bit harder to find the way to dance to that."*. While in agreement, **P4** did say that in comparison, rhythm *“was not the easiest one, but easier, because it still has a strong beat.”* Regardless of different participants’ preferences for either rhythm or melody, there was an agreement that more layers and dimensionality in music would be helpful. In this way, they could find some element they can dance to. **P10/A** summed this up very well, mentioning that *“if you don't have the accents in the music you can do something with, that you get inspired by"*, the task was more difficult. One such element that the participants agreed helped to interpret the emotion of the music playing was lyrics.

- Personalization, choices and dynamics

Yet another important theme that came up was making the sound materials better suit individuals. For instance, we could let them get a say and choose certain aspects, according to the participants. **P10/A** found it *“nice to have choice, [because] then you have the feeling that you have influence*.”

Importantly, another discussion point was to what degree or regarding what aspects the participants should be allowed to have an impact. **P9**, for example, mentioned wanting to keep a level of novelty and surprise. He made suggestions based on comparable contexts such as themed parties:

"I kinda like the surprise aspect of it. Reacting to whatever you hear. I prefer giving, like, a general theme of what exactly is to be expected. For example, a lot of parties have things like ‘reggaeton’ or ‘afro beats’. Like, you know what kind of music style you can expect, but whatever comes, that's always surprising. Maybe it's a song you don't know."

P10/A was also in agreement with him that she *“would not like to choose the specific song”*. She suggested, instead, opting for choices such as ‘more slow’ or ‘more fast’; ‘more beats’ or ‘more melody’.

An interesting related interaction dynamic was suggested, as well as utilized during the workshop, by **P8/A**. Specifically, she made use of the differing and shared comfort zones of participants, with regards to movement and sound. She explained this, saying:

"[My strategy was] sometimes taking different aspects that you can both explore together, but also taking aspects that one knows a little bit more than the other, so that you can also play with that. With one person; they have a more dominant role for a part and the other follows. Then you play with the leader-follower roles. I think that would be a very fun interaction."

5.6.6 Context-and-Procedure-Related Suggestions

The participants gave feedback, also in terms of the use context and the workshop procedure.

- Comfort zones and expression

The participants highlighted, as dancers themselves, that they wondered how a non-dancer would experience this sort of an interaction. Even as dancers, they recognized that when you are familiar with only a single or a limited style (e.g. one that relies more on set figures or choreographies), free-styling in this way does not necessarily come naturally. As **P10/A** put it: It *“really depends how comfortable someone is in just free-dancing, because if I only did ballroom, I would not be comfortable doing this.”* **P2**, one participant with a more specific dance experience and background (in ballroom), also mentioned this being out of his typical boundaries.

These comments affirmed, in some sense, the decision to include dancers in the workshops: They could provide certain insights while and through being comfortable in such a context. **P5** commented, also referring to their interaction partner, *“Cause I think for both of us, it was kinda easy to show ourselves in this and let ourselves go.”* Especially this ‘showing themselves’ aspect and, in general, having self-expression in your movements and through dancing is not something everyone possesses to the same degree. Therefore, it makes a (technology-mediated) interaction like the current one both interesting and challenging.

Regardless, a suggestion was made by **P6** to add to the session another part, in order to address this challenge. Specifically, she made a suggestion inspired by her background in physical theater: The interactors could lay on the ground for ten minutes and just check in with their bodies. In this way, they could get in a more focused and grounded state, prior to the interaction.

- Interaction setting

Another point that came up was that the setting itself also played an important role in how the interaction went. The participants pointed out how this context felt different-than-usual

to dance in. Oftentimes in clubs or parties, for example, it is dark, so people feel more comfortable dancing.

Furthermore, it felt more intimate for them to be the only ones dancing in the entire room, as opposed to being in a setting with many others also dancing. This and being watched by the researcher and/or 'technology system'/DJ present influenced in what ways they would attempt to dance together. Physical contact, for example, would be something they are more open for incorporating at a party setting than in a workshop, like this one.

To counter this situation, **P8/A** and **P10/A**, in agreement, suggested that it should be made clear what this space is for; the particular kind of interaction expected to take place. **P10/A** said that it should be highlighted how *"this space is for dancing and it's normal that you dance here."* Similarly, **P9** thought inspiration could be drawn from settings people already enjoy or feel comfortable dancing in. For example, we could also opt for dimmed lights like in clubs.

In terms of ensuring comfort and a safe space, some other points were also mentioned. For instance, a setting could be provided, in which the interactors would not be watched by outsiders. **P10/A** did point out, however, that *"Then* [her] *anxiety is, people don't behave.".*

- Target user group: Children

Last but not least, **P8/A** shared her perspectives on what user group could be a suitable one for this kind of an interaction. She had previously worked as a music therapist, also together with (autistic) children. Thus, she explained that in her experience, *"Anxiety often comes up around age 10-11. Right before puberty. For girls, earlier than for boys. Because they are often more aware of the social norms."* This, in her eyes, made autistic and NA children to be potentially promising participants. Teenagers, in comparison, had a lot of shame, making them a less suitable user group.

Overall, the first iteration workshops and especially the semi-structured interviews conducted gave rise to many insights. These were looked into, through their thematic groupings.

5.7 Results' Interpretation and Discussion

In this section, the raw results from the first iteration and the themes they got categorized into are interpreted and discussed under the main research questions.

5.7.1 RQ 1: Autistic versus NA

Overall, it seemed like the amount of divergences or mismatches between the autistic and NA individuals in interaction – in terms of their perspectives, perceptions and expectations – were a lot less than anticipated. This could be interpreted in a number of ways:

For instance, with this technology-mediated interaction being a novel one for both parties, it is possible that the situation, as intended, brought forth dialogical, rather than social system processes (Steffensen, 2012). The interactors, in this case, may have been prompted to engage in participatory sense-making (De Jaegher & Di Paolo, 2007). They may have established meaning together in the moment, rather than repeating neurotypical normative

patterns of interaction in their actions (De Jaegher & Di Paolo, 2007). This may, in turn, have helped them to be more in alignment.

Another possibility would be related to the heterogeneity of the autistic experience (Lord et al., 2020). This, understandably, cannot be fully captured through the few participants included in the workshops. Especially the involvement of adult autistic individuals – ones that are highly-educated – might have had an impact: These qualities may have minimized or eliminated certain challenges that could have come up in interaction. In particular, it should be taken into account that the autistic participants included in the workshops were all female. Hence, they might have shown better masking qualities, appearing to better fit in with certain NA expectations and behavior patterns.

One line of difference that was observed amongst autistic and NA participants related to a strategy and indication to dance together, that is, eye contact. Interestingly, all participants mentioned this as one strong sign of wanting to dance together regardless of being autistic or NA. Especially **P10/A**'s comment shed a light on how making eye contact, for her, was different in this sort of an interaction and setting, as opposed to during conversation in daily life. She talked about the lessened lines of input in a context like the current one and how eyes allow you to clearly understand someone's energy. This, one could make use of in many ways when relating (back) to their interaction counterpart, rather than keeping eye contact being a mere overwhelming neurotypical-norm-based action as usual.

Eye contact and the divergences of autistic and NA participants' expectations and experiences of it, on the other hand, also led to obstacles in connection-making: In **W2**, **P1/A** had attempted to dance with **P2** by mirroring/repeating his moves. However, due to her lack of (initial) eye contact, **P2** was surprised to hear about this situation after the fact. He had completely missed these intentions and steps taken by his interaction partner. While talking about possible contributions technology could have regarding this challenge, **P2** was in favor of the framing imagined for the current system, where its sound outputs could draw attention to an interactor's connection intention and attempts through its sound outputs, to make these more visible to the other:

"I think if it would actually infuence the music if the other party is seeking contact, in a one-on-one situation, it might infuence it [the challenge/noticing the attempts]. Like, actually making the music adaptive to whether or not you are trying to reach for contact."

In this way, one such main difference across the groups could be identified, which has potential for design opportunities.

5.7.2 RQ 2: Mappings' connection-prompting role

Before talking about movement-to-sound mappings that seem to prompt connection – as also seen from the data itself – we first need to understand what it means to feel connected; how exactly people can be prompted to establish a bond:

'Revealing'/getting to know. There is possibility for connection when the interactors are able to get a sense of 'who this other person is', within this context of the novel technology-mediated interaction. This involves, for instance, the individuals' comfort zones with regards to such a task – their tendencies and patterns and likings in how they move, what they listen to or like in songs, and how they interpret these. When revealed, the participants are able to connect through all these elements that comprise their individuality within the given context. For example, unexpected sounds can make them give an authentic reaction in the moment. Similarly, during mirroring, the participants can take a certain move from their interaction partners and put their own style and twist onto it. When such incidents occur, the participants' interaction partner feels as though something intimate and personal is shared with them.

Agency & reciprocation. There is power in knowing that it is by decision that you connect with the other – and that they are doing the same. Compared to a set-up that inflicts some sort of connection at all times such as the mirroring warm-up, this togetherness is not a guarantee in an open-ended interaction. This, in turn, may make it feel more valuable and real. The participants, then, can interpret someone mirroring or copying their movements as them having the genuine intention to be doing this together. Also when they imitate their interaction partner themselves, the participants communicate that they like how the other moves and that they want to join in. Such exchanges are connecting- and relationship-building.

Physical prompting emotional connection. Interaction set-ups may inflict a kind of a togetherness onto the interactors, through their physicality. As mentioned, these can also enable an authentic emotional bond, in a different way. This holds, even without something formal like the mirroring warm-up – for instance, in a situation where the interactors are in the same groove and are dancing to the same rhythm. Through this 'same' connection they have to the music and its elements, the participant can be helped in establishing a bond. Their individual connections to the shared sound outputs can feed into and act as a bridge for a connection with one other. Such proxy connectors are an interesting theme to build on, when considering how to design for connection.

Interpersonal dynamics & roles. Whether given through a set-up such as the mirroring warm-up or emerging on its own in an open-ended context, the participants having clear expectations regarding the interaction and their roles and dynamics within it is important. This can include, for instance, the ways in which the interactors are to move and work with the other. Such assignment of 'positions' can act as an added layer for connection-building: Taking upon, for example, leading or following roles, the participants can alternate between revealing something of themselves and getting to know the other. Furthermore, while doing so, they may even benefit from the aforementioned physical prompting. They can feel more comfortable and thus, be in a better mental space; able to even think of connection-building in the first place.

There are, however, also certain challenges in connection-making that can come up in interaction. These comprise points that can be addressed by a technology system like the current one:

Synchrony in when (and how) to connect. One challenge a more constrained set-up like the mirroring warm-up alleviates is that of the interactors having to align on their intention and attempts to connect. There, connection is a precursor and is already in existence. However,

when open-ended, interactions pose a difficulty: For any possibility of relationship-building, there must be attention on and intention for this from both sides. Otherwise, failed attempts in bonding may lead to feelings of rejection. This may, on the contrary, hurt a possible connection opportunity and the interactors' openness to it.

Expression of intention. As pointed out by **P1/A**, it may be easier to understand that someone does not want to dance together, rather than to feel confident that they do. How to establish and communicate that we want to, or are already moving with one another remains the core challenge, despite the converging dancing-together strategies observed. This is for the interactors to, in-the-moment, resolve. Here, mismatches, coming from autistic and NA different social expectations and behaviors, are able to intervene and complicate matters.

Meaning of 'dancing together'. Also informed by other aspects such as our cultural or dance background or individual preferences – our comfort zones, we may have very different ideas of what it means to 'dance together', which comprises the 'how to connect' as well. This is where, as **P8/A** suggested, interpersonal dynamics such as leader-follower roles and alternating these can help to balance and validate expectations and needs from both sides. That way, interactors can learn from one another. For instance, this can include figuring out different ways of 'dancing together'. In this way, the participants can together establish a framework for this sort of novel interaction context.

Need for common ground & mutual adjustments. There is still a need for a base the interaction can build itself upon, regardless of however much of this we can develop as we go. Furthermore, as mentioned, a willingness from both sides should be there, to be open and responsive to the other person, their ways. For instance, specialized ways of dancing, such as 'cool breakdance moves', may hurt the connection. This may be the case if these heighten feelings of being unleveled with one another. Such actions may not provide a clear way the other person can join in, as attested to by many participants. In this sense, the challenge of the interaction format itself matches very well the broad goals of the project: Successful neurodiverse dyadic social interactions are, in this manner, made both sides' equal responsibility. This fits with the idea that both parties should get to know or understand, adapt to and appreciate each other's ways. With that said, the technology system can, then, help the interactors in locating these commonalities as possible connection points. It can support the sort of flexibility needed in participants accommodating one another.

Hints from unclear language. The way people talk, when talking about their experiences within the workshops, of their interaction and making a connection provides clues. This points towards a fact: It is not that the connection-making contexts are straightforward and deciphered for themselves either. Fitting with the concept of participatory sense-making (De Jaegher & Di Paolo, 2007), they are constantly learning and developing themselves in interaction. They are not talking about very clear actions that make things successful or cause them to fail. When they talk about connection, they talk about being in the same 'energy' or 'vibe'. They mention that they must 'match' in some sense, even if they themselves also do not know what this may actually mean.

Having to focus on yourself or the input/environment. Last but not least, for connection to be allowed-for in a context, the 'rest' should not be in the way. This involves, for example in a music-based interaction, that the volume should be high enough. In this way, one does not have to worry about understanding what sound is being played. Similarly, the musical

content should support their expectations, provide them with a clear idea as to what they can work with. For instance, this can look like not-too-frequently changing sounds – ones with a clear storyline. The technology-mediated interaction context itself, as much as the system, should help interactors in keeping their attention and capacities for their counterparts, as well as the act of connecting.

All in all, through this first iteration, many interesting insights could be collected, regarding the meaning of connection-making in this technology-mediated interaction, as well as any challenges that may come up, which can be addressed through the movement-to-sound mappings of the system. More practical details of these mappings were described in the following section:

5.7.3 RQ 3: Mappings: Open/adjustable versus pre-determined

Overall, the data obtained from these workshops clearly showed a need for movement-to-sound mappings to be adjustable and personalizable. For instance, participants had different genres of music they enjoyed, despite some commonalities. Similarly, they differed in whether they listened for and responded more to melody or rhythm. There existed strong opinions on both sides, regarding whether an upbeat or a slow melody supports connection-building better.

This need for flexibility also held with regards to the movements that would be sensed by the system. In particular, these appeared to depend on who is present, together with whom else. The participants each had their natural ways of dancing and moving. Within the dialogical system (Steffensen, 2012) that comes to be, the participants' individual approaches would also influence their counterparts' in specific ways. For instance, a different manner of interpreting the musical inputs and dancing to these could be sparked, based on with whom you were joining the interaction. This was similar to how a conversation could be led to different avenues in different cases.

Contrary to initial expectations, the 'movement inputs', thus, also turned out to be better framed as broader themes, rather than very specific actions to be detected. The workshops pointed towards the participants' dancing-together strategies (e.g. mirroring/repetition and physical touch) as what the system could aim to notice and to measure in this sense.

These, coupled with individual perception-related matters (e.g. sound hyper- or hypo-sensitivities), especially in the autistic embodied experience, for sure, require the technology system to suit these different needs of the interactors.

5.7.4 RQ 4: Measuring connection/success

As mentioned above, certain strategies in or indications of wanting to or dancing together were identified through these first iteration workshops, which were the following:

- Eye contact,
- Mirroring,
- Lessened distance,
- 'Matching', and,
- Physical contact.

These could be relevant pointers, especially for the design of the movement-to-sound mappings that attempted to give rise to reactive embodiment (Spiel et al., 2016) through connection-reflecting. The instances of these actions were thought of as possible ways connection or interaction success or progression can be measured during. Similarly, other occurrences such as people laughing inspired a possible measure design. Hence, through this iteration, certain promising indicators for the in-system evaluation purposes were obtained, which needed further empirical validation.

As for the workshops' success and the technology-mediated interactions' evaluations overall, self-reports could be relied upon (e.g. regarding felt connection levels), given that all participants were adults. Specifically, 'success' was defined not through counts of mainly neurotypical expectations such as how much eye contact were made, but rather similarly to the work of Spiel et al. (2016): It was based on the participants' own ideas of whether this was a meaningful shared experience. Here, inclusion of both discussions amongst participants, as well as questionnaires where they could fill in their opinions and feelings without their interaction partner needing to hear this helped with honesty and accuracy. It *was* noted that this overall evaluation of the interaction's success could be more complicated in other cases (e.g. with non-verbal individuals, children, etc.).

5.7.5 RQ 5: Real-life product

The insights that could be derived from this first iteration, with regards to what a good use context, scenario or goal could be for the technology system were limited. The comfort, naturalness and enjoyment ratings participants gave to the interaction were promising, when it came to turning this prototype and conceptual design into a real-life product, albeit keeping in mind the participants' dance backgrounds. It remained a question, how well the findings with this group would translate to other ones, such as those that were not dancers. The participants themselves also pointed this out as an important matter of consideration that needed further investigation.

Conversations with **P8/A**, however, *were* in support of targeting one possible user group that was already common for autism interaction technologies (Oral, 2022): Autistic children. This was the case, for instance, compared to teenagers, who had a lot more shame and would be less free in moving and dancing around.

The workshops informed us, for instance, that a guided and a more open-ended (technology-mediated) interaction could offer different benefits. Therefore, for now, a use set-up was imagined where these would be mixed, to reap both of such positive impacts.

Importantly, inputs from the participants highlighted paying attention also to other aspects such as the interaction setting. The real-life product could, for instance, draw inspiration from successful environments wherein people felt comfortable moving around: Darker environments with dimmed lights could give them a sense of privacy, whereas not being the only ones dancing in a room and not being watched could help the interaction to feel less intimidating. Having this set space for dancing with clear expectations was also noted as a quality that could support (autistic) participants by **P10/A**, with all of these points painting a clearer possible use context for a real-life product.

6. Second Iteration

6.1 Conceptual Design

For the conceptual system design in this iteration, the empirical data from the first iteration was supported by other means such as theory and creative inspiration. As a result, a threefold further-developed design was established (see ***Figure 12***)

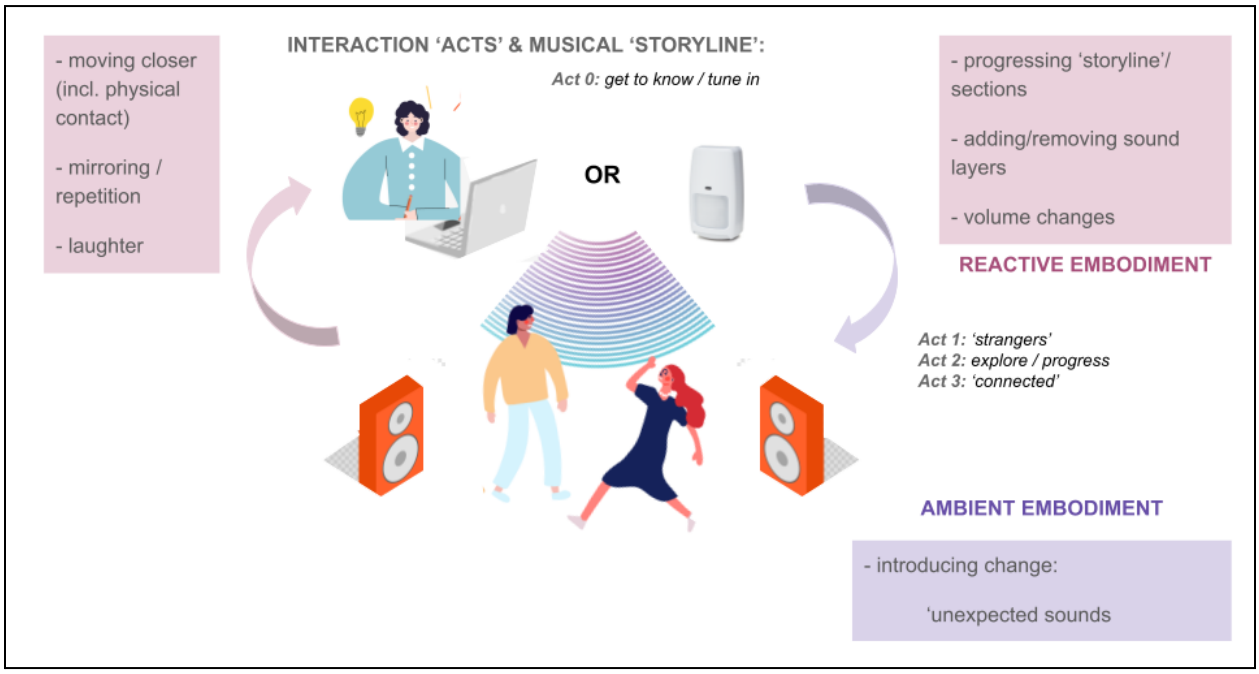


Figure 12: An overview of the conceptual design for the second iteration. The design is threefold, involving *interaction 'acts'* & a *musical 'storyline*, *reactive embodiment through adaptive music* and *ambient embodiment through soundscape evolution.*

1 - Interaction 'Acts' & Musical 'Storyline'. The first iteration findings appeared to favor the system's sound outputs featuring a sort of a 'storyline'. Furthermore, inspiration was drawn from related work – in particular, *Visible Voice* (Monobanda, n.d.) and its act-based structure. As a result, the technology-mediated interaction was designed in a way that included a story-like structure. Specifically, it would have the following acts:

- ***Act 0:*** *Get to know / tune in*

  This pre-act, abiding to the phases-structure design guideline (Oral, 2022) would involve a guided warm-up exercise. The warm-up would be there for the interactors to get accustomed to the movement-and-sound interaction context, as well as to one another. Specifically, it would include the same mirroring exercise as in the first iteration, where the participants would switch their leading or following roles upon cue.

During this stage, the participants' musical preferences would also be learned by the system. This would involve, for example, melody versus rhythm, as well as slow-melody versus fast-melody likings.

- ***Act 1:*** *'Strangers'*

  At this stage, the open-ended music and movement interaction would begin. Then, the interactors would not yet know each other or only know each other to a certain degree, with more to be discovered. They would be in the interaction as two individuals; 'strangers'.

- ***Act 2:*** *Explore / progress*

  Through moving (together) to the music, the interactors would get to know each other in this different-than-usual context and sense. This would be less bound to certain neurotypical norms, as a result of being a novel interaction situation. In this way, the system would prompt the interactors to in-the-moment establish meaning in collaboration, by engaging in participatory sense-making (De Jaegher & Di Paolo, 2007). They would get a sense of what kind of music or musical elements each of them enjoy, in what ways they like to – or feel comfortable in – moving their bodies and similar.

- ***Act 3:*** *'Connected'*

  Hopefully, by the end of the interaction, a connection would be prompted or established between the interactors. As a result, they would feel like they have a fuller picture of one another, their differences and similarities, as well as how to work with these in interaction.

The musical progression throughout this storyline would be similar to how it was in the first iteration. In particular, it would involve sound changes that are either in response to the movement inputs from the interactors (i.e. *connection-reflecting*) or are coming from the system itself, independent of what is happening within the interaction (i.e. *exploration-prompting*):

2 - Reactive Embodiment through Adaptive Music. The first category of sound changes within the storyline would, as mentioned, be tied to interactor actions. Specifically, these movement-to-sound mappings would aim to prompt connection. They would do so, through reflecting certain aspects of the ongoing interaction in and via the music. This would serve the purpose of enhancing interactors' emotive experiences and immersion within the interaction, just as in other adaptive music approaches (Hutchings & McCormack, 2019). In this way, establishing *reactive embodiment* (Spiel et al., 2016) would be attempted at, making use of the sound modality: An interplay would be created between the interactor actions and these outputs. In particular, the together-movement of the interactors would impact the sounds played and the sound changes. This, in turn, would influence how the interactors carry on moving together or their sense of connection. Hence, a chain of opportunity creation for exploration would be established (Spiel et al., 2016).

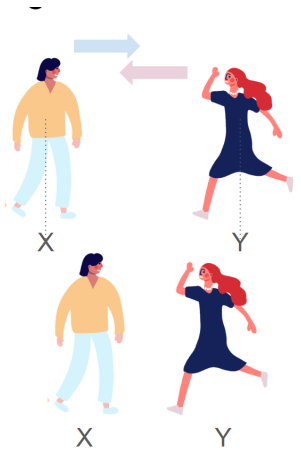
Based on empirical data from the first iteration, a set of promising movement and interaction aspects were gathered. These could be used as inputs for the sound mappings in

this iteration, when attempting to achieve reactive embodiment. Specifically, these movement and interaction aspects involved three main categories: *Moving closer (including physical contact), mirroring/repetition* and *laughter*. The first two of these were also given subcategories, which corresponded to the different degrees to which progression is observed within the interactors' connection:

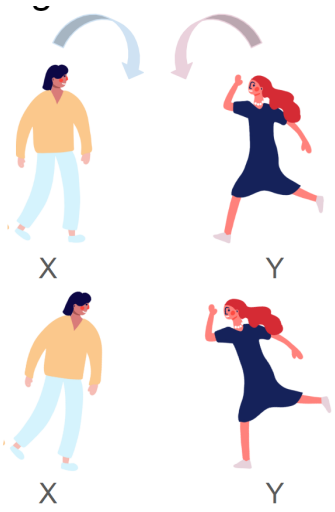
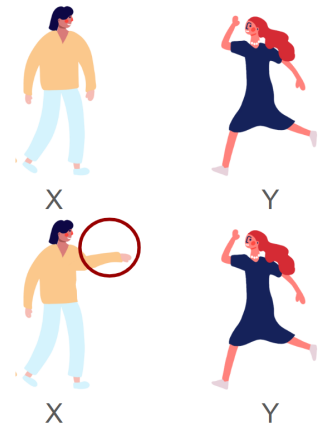
- Moving closer would go from *walking towards* or *leaning* to *reaching* or *dancing around*, to, at last, *physical contact*, and,
- Mirroring/repetition could be of the *body part* or *direction* or *kind of movement* (e.g. speed, number of times, type), showcasing imitation at increasingly higher levels.

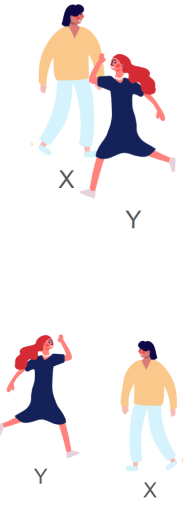
Similarly, ways of musically progressing – through changes and additions in the system's sound outputs – were established, again based on insights from the first iteration. This was a way the connection progression could be reflected in and via the music. These approaches involved *progressing the 'storyline'* (by playing the next section clip, for melodies that were cut into sections), *adding sound layers* and *increasing the volume* – ordered in accordance to the connection progression levels they would be utilized at. As a result, the movement-to-sound mappings to be tested within the second iteration were gathered, as can be seen in ***Table 11.***

Table 11: Movement-to-sound mappings to be tested within the second iteration.

Movement Input Category	**Movement Input Sub-Category Visualization**	**Movement Input Sub-Category**[17]	**Matching Sound Output**
Moving closer		Walking towards	Progressing the 'storyline'/sections

[17]The color fill-ins in this table visually indicate the different degrees of connection progression (i.e. going from yellow to orange to red).

		Leaning	
		Reaching	Adding sound layers
		Dancing around	

			
		Physical contact	Increasing the volume
Mirroring / repetition	-	Body part	Progressing the 'storyline'/sections
	-	Direction	Adding sound layers
	-	Kind of movement (e.g. speed, number of times, type)	Increasing the volume
Laughter	-	-	Increasing the volume

3 - Ambient Embodiment through Soundscape Evolution. In addition to sound changes that would be tied to interactor actions, the system would output sounds independent of these interaction and connection aspects, for the purposes of creating *ambient embodiment* (Spiel et al., 2016). Specifically, this musical layer could be added to the social interaction going on, so that it would be present as another possible source of input for the interactors. In this

way, connection could be prompted indirectly: The soundscape evolution (i.e. sounds and sound changes outputted overtime) could prompt prolonged together-exploration and interaction. This, in turn, could give rise to more connection opportunities for the interactors. The sounds could influence, through their presence, how people interact within the space and with one another, acting as scaffolds (Fletcher-Watson et al., 2018).

Based on empirical data from the first iteration, the specific sounds that would serve this purpose would have to do with introducing change. This could, for instance, be by means of playing short and unexpected sounds at random moments. Specifically, such sounds would be incorporated, in order to elicit genuine and authentic reactions from the interactors: Through these reactions, the other person in the interaction could get a sense of an interactor's personality.

6.1.1 Procedural Specifications

According to the described conceptual design, a procedural plan for the workshops was drafted, for this design's evaluation in an intended use context. This was regarding specific sounds and sound changes incorporated for *reactive embodiment* purposes, as well as how the system could be imitated in this sense, through WoZ (Wooffitt et al., 1997).

During the warm-up, first, the participants' preferences would be determined, inspired by **P8'**s strategies as a 'technology system'/DJ within the first iteration:

- The same melody would be played with different accompanying rhythms, in order to observe how much the participants changed the way they moved – whether they focused more on the *rhythm* or the *melody.*

- Similarly, both *slow* and *upbeat melodies*, or a single melody with both slow and fast parts would be utilized, to obtain insights on the participants' likings of either one or the other.

When it came to the main interaction, in which one or two songs in total would be played, the rules below would be followed:

- If a participant prefers *rhythm* and *upbeat melodies*, a monotonous upbeat melody would be played. Musical progression would be achieved through using different rhythms over this base melody, for example, also going back to certain patterns after some time intervals.

- If a participant prefers *rhythm* and *slow melodies*, the same would be done, opting for using a slow monotonous melody instead.

- If a participant prefers *melody*, specifically, *upbeat* ones, the first song played would be a progressable melody (i.e. cut into smaller sections, where the next clip could be played or moved on to) that has both slow and upbeat parts. Then, another upbeat

progressable melody would be used, achieving musical progression through this developing 'storyline'.

- If a participant prefers *melody*, specifically, *slow* ones, the first song played would be a slow progressable melody, which would be followed by another progressable melody with both slow and upbeat parts – the same as the above case.

6.2 Prototyping

6.2.1 Hardware and Software

Also for the workshops included in this second iteration, the Ableton Live 11 (Intro) software (Version 11.3.3; McKenzie et al., 2014) was used for music creation purposes. As opposed to the first iteration, however, in this iteration, there was no need for the Glover software and the Gliss app – and hence, a phone. This was due to the music being controlled only by the researcher herself, who was already familiar with Ableton Live 11 and controlling sound change aspects in this way.

In addition to running the Ableton Live 11 (Intro) software on the same *laptop*, one hardware added in this iteration was the *JBL GO 2 Bluetooth speaker* (180Hz – 20kHz). This aimed to improve the sound output quality, to address the feedback obtained during the previous workshops.

One other intention was to draw less attention to the camera used for recording. This was given that some adult participants in the first iteration were more aware of being recorded than others during the sessions. Hence, a smaller camera was thought of as a better alternative. Additionally, a wider-perspective angle was preferred. As a result, one of the cameras previously used for recording (i.e. *Panasonic HC-V720*) was swapped with a *GoPro Hero 3*.

6.2.2 Sound Design

Once again, in the interest of time and considering the researcher's own musical skill levels, all of the sound materials included in the prototype were pre-made ones. They were either from the Ableton Live 11 software itself or from online sites, selected based on criteria such as the musical materials' genre or mood (for a full list, see ***Appendix K***).

Following the conceptual design of the imagined system and the different procedural scenarios thought of for the workshops, the prototype featured:

- the same rhythms as the first iteration,
- an additional selection of sounds these rhythms could be layered with,
- slow and upbeat monotonous melodies, for easy-layering with the rhythms and the additional sounds,
- more complex or developed slow and upbeat melodies,
- three different melodies (i.e. slow, a mix of containing both slow and upbeat parts and upbeat) segmented into smaller sections and could be 'progressed', and,
- a set of short and sudden sounds, considered to be 'unexpected' (e.g. dog bark).

Just as in the first iteration, the genre the most participants reported enjoying was prioritized in the system's sound outputs. Thanks to the contextual inquiry conducted prior to the workshops, we had insights regarding the children's musical preferences: They all listened to and liked different musical genres. However, all of them had children's songs in their class activities and were used to these. Hence, all the melodies, except for the more complex slow one were ones labeled as 'children's songs' by their creators. The latter, on the other hand, belonged to the instrumental folk genre.

The *sound changes* available included layering possibilities involving rhythms, other sounds and the monotonous melodies; as well as the 'progressable' melodies that could be moved onto the next clips. Additionally, there was one other sound change that depended on movement or interaction aspects: Volume could be increased or decreased, as this was another strategy that came up in the first iteration.

6.2.3 Interface Design

Once again, the non-traditional vertical arrangement interface/view in Ableton Live 11 was used (see ***Appendix B***). In this way, the different sound categories available were organized into different columns (see ***Figure 13***).

Figure 13: An overview of the Ableton Live 11 (Intro) software (Version 11.3.3; McKenzie et al., 2014) interface included in the second iteration, where different categories of sound clips (e.g. 'rhythms/beats') are placed into their own separate columns.

In particular, the following were the visual elements paid attention to:

- The columns' order was organized in such a way that the three layerable categories of sound clips (i.e. rhythms, other sounds and (monotonous) melodies) were together and came first.

- Different/same colors were used to make it easier for 'technology system'/DJ to be able to recognize sound clips as different or parts of the same one (i.e. progressing).

- The ordering of the individual sound clips within the rhythm column, as well as the ordering of the three 'progressable' melody columns matched this progression. For example, in the case of the melody ones, the clips went from slow, to a mix of slow and upbeat, to upbeat; or involved less to more 'complex' rhythms (per subjective perception of the researcher herself).

6.3 Participants

In this second iteration, the aim was to see the refined conceptual system design in use by the target user groups. One such relevant context – also commonly seen in autism interaction technologies (Oral, 2022) and that came up during the first iteration – was an in-school one, involving autistic children and their NA contacts, such as teachers. Thus, a collaboration was made with Kinderdagcentrum De Walnoot[18], for some students' and their teacher's participation. This was on a voluntary basis and the participants were not provided with any additional compensation.

In total, 4 students (3 male, 1 female) took part in the workshops, together with their NA teacher (female). Two individual workshop sessions were held, where in each one, the researcher herself had the role of the 'technology system'/DJ (see ***Table 12***).

Table 12: Workshop sessions' breakdown in terms of participant amounts and gender-, age- and neurodivergence-division.

Workshop #	Participant #	Role	Gender	Age	Neurodivergent?
W1	P1	Interactor	M	6	Yes, autistic
	P2	Interactor	M	9	Yes, autistic
	T	Interactor	F	-	No
W2	P3	Interactor	M	6	Yes, autistic
	P4	Interactor	F	6	Yes, autistic
	T	Interactor	F	-	No

6.4 Procedure

The workshops in the second iteration, as aforementioned, took place at a school. Thus, there existed a key difference regarding the procedure, compared to the first iteration one: Prior to the day of these workshop sessions, a visit was made to the said school, as a sort of a contextual inquiry (Karen & Sandra, 2017). This was so that the school environment, the teachers, the children and how they are already used to doing certain things could be observed and gotten to know.

The visit involved a tour of the space and the different classrooms, as well as meeting the groups of children that study at the school. Additionally, conversations were held with the teachers regarding the childrens' needs, preferences and similar. For instance, a decision was made together with the teacher regarding where the sessions could take place. Specifically, the school gym was deemed suitable to use for the workshops (see ***Figure 14***). This was so that

[18] For the school/organization website, see https://www.gemiva.nl/locaties/de-walnoot

the children would have enough space to move around. What's more, they already had the association with this space that it is a place where they would get to be active in.

Figure 14: A photo of the school gym where the workshops took place.

The time availability of the children and the teachers posed limitations on certain procedural practicalities. In particular, only one teacher would be able to join the sessions. She would, then, be there both as a participant herself and as a supervisor for the children. This meant that any workshop session that took place that day would have to be within a given time slot and at a single location, so that she could be present at all times.

This situation influenced the current work, due to one of the other researchers aforementioned also wanting to conduct her sessions on the same day. She would be evaluating her own prototype that involved haptics, together with the children. Hence, there would be no procedural overlaps with that and the current work. However, the space had to be shared given the limitations, making the workshop sessions in this second iteration combined ones. In particular, these sessions were planned to be carried out simultaneously, with one child joining the workshop of the other researcher and the teacher partaking in the music-movement one, together with the other (for later deviations from this, see ***6.7.1***).

Importantly, the recruitment was handled through the teachers prior to the workshops, following the contextual inquiry. In particular, the children's legal guardians were the ones giving consent on their behalf (see ***Appendix L***).

Rather than distributing a descriptive questionnaire like in the first iteration, insights from the contextual inquiry were made use of: This, including the conversations that took place during, helped to get to know the participants. For instance, their (sound-related) sensory sensitivities and musical preferences could, then, be taken into account for the workshops.

On the day of the workshops, following the set-up of the space and the equipment, the teacher, herself, went and brought the children in. Before doing so, she explained to them what

would happen (i.e. that they would dance and move to the music however they liked), to set fitting expectations before their participation.

In this iteration, the sounds and sound changes during the sessions were controlled only by the researcher herself. She was acting as a 'wizard' to make alterations according to a set of rule-based instructions put together, in order to help the participants to connect. The procedure from the first iteration was planned to be replicated, in the sense of beginning with a mirroring warm-up (for deviations from this, see ***6.7.1***). Later the teacher and her student would, once again, be asked to move in any way that felt comfortable, natural and enjoyable to them. They could listen to the music and/or look at each other's movements if they wanted to, but without any specific instructions or constraints to connect with one another (in a certain way) while and by moving. Both of these parts of the workshop session were video- and audio-recorded, with consent.

After this main interaction had ended and the participants were sufficiently rested, they were shown an age-appropriate three-point smiley scale[19], so that they could express how they liked the interaction (see ***Figure 15***). Then, it was planned that the two children present in the session would switch places to try out the other prototype (i.e. either audio or haptics).

Figure 15: A photo of the age-friendly and color-coded three-point smiley scale used, when assessing how the children enjoyed the interaction.

Again, given time constraints, one overall semi-structured interview was held with the teacher, once all workshops were completed. This part was also audio-recorded.

In total, all workshops combined (including the following discussion with the teacher) lasted for 1 hour. On average, the actual movement-and-sound interaction was 5-10 minutes long. The interview with the teacher, on the other hand, lasted about 10-15 minutes.

[19] This scale was, once again, chosen in consultation with the teacher, who liked the smiley faces idea and recommended the three-point one over the five-point alternative.

6.5 Contextual Inquiry

During the school visit, observations were made and conversations were held with the teachers. Through doing so, the following insights were gathered:

- The children relied heavily on routines and structures, as well as visualized print-outs of these (e.g. of their daily schedules, put on their classroom doors). This included having certain rooms that they knew they would be doing certain activities in. Such separation and assignment helped to set expectations, and to provide them with a sense of comfort and familiarity.

- Along the same line of thought, the teachers highlighted that ‘showing’ rather than telling children things supported them. In this way, they could better understand what was being asked of them, for example during activities.

- The children had different (sensory) needs and ways of communicating. Some of them, for instance, were non-verbal to different degrees. These students utilized tablets to converse with their peers and teachers. In general, such screen-based (social and otherwise) interactions were very familiar to the children. This was due to the widespread use of these inside and outside of the school context.

- The school and the teachers already saw the value in multi-sensory approaches. The facilities were equipped with a variety of these sort of play elements and spaces (see ***Figure 16***), including technological ones.

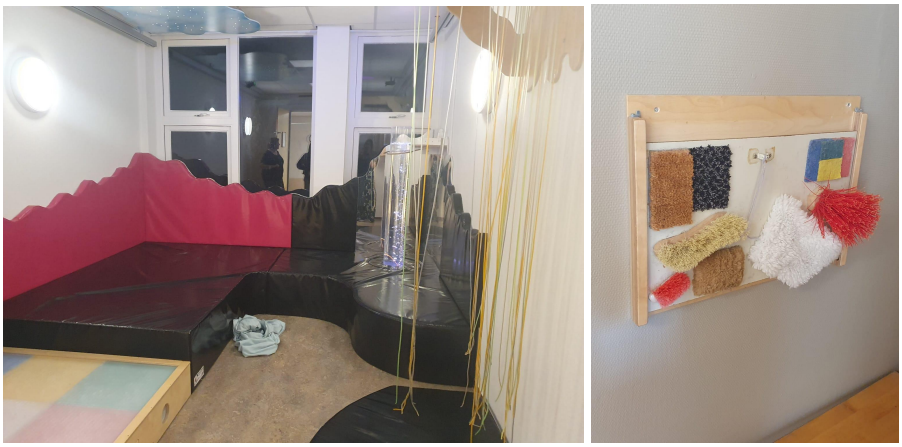


Figure 16: Photos of example multi-sensory facilities the school was already equipped with. Specifically, a ‘quiet room’ the children can make use of, to cool off, enjoying different light settings *(Left)* and a set of differently-textured items hanging on the wall, which the children can touch and interact with *(Right)*.

- At the school, they also had a variety of musical instruments and toys available (see ***Figure 17***). Unfortunately, the children currently did not have a music teacher. Despite this, they would often sing and listen to songs together during their regular classes. These songs would often be children's ones. The teachers did express their reservations when it came to incorporating musical activities, due to feeling as though they did not possess enough knowledge and skill in this area.

Figure 17: A photo of the cabinet that holds the music-related equipment the teachers and the children can make use of. There appear to be a lot of different musical instruments, such as flutes, xylophones and drums.

- Beyond their familiarity with children's songs, what kind of music the different students enjoyed varied, according to their teachers.

- The teacher that joined the workshops in this iteration specialized in 'play'. In this regard, she especially had a room at the school, which was filled with various toys and objects (see ***Figure 18***). The children would get one-on-one time with her there, as well as having other opportunities to play and interact with one another in larger group settings.

Figure 18: A photo of example objects the partaking teacher has available in her playroom. Inside one of the cupboards are these 'sensory bottles' that involve different visual and sound elements. Some of these are also themed (e.g. according to holidays such as Sinterklaas). She creates the 'sensory bottles' herself for her students.

Overall, this information collected through the contextual inquiry helped to better prepare for the workshop sessions. Rapport could be built with the teachers, as well as with some students. In addition, these points helped in better understanding the children and interpreting their actions, later in the said sessions.

6.6 Data Analysis

Beyond the school visit made, as a form of contextual inquiry, the second iteration involved the following data sources that could or needed to be analyzed:

- three-point smiley scale ratings of the interaction by the children,
- workshop video recordings, and,
- audio recordings of the post-interaction semi-structured interview with the teacher.

The interview recording was, again, manually transcribed (see ***Appendix M***). Differently than in the first iteration, more importance was put on the video recordings. This was given the lack of interviews with the children. Specifically, using a free video editing software (i.e. *Kdenlive* (KDE, n.d.)), these were separated into clips. What's more, accompanying event description scripts were created for the workshop recordings per clips (see ***Appendix N***).

Once again, thematic analysis was conducted, using Braune & Clarke (2006)'s methodology as a guide. This was done so, on both the interview script and the workshop event description outputs. Codes were created based on the developed interview questions, as well as any emerging themes. As a whole, these were analyzed based on the main research questions, as well as in relation to the first iteration results.

6.7 Results

On a three-point smiley scale (i.e. going from a sad, to a neutral to a happy face, see ***Figure 15***), three out of four children gave the interaction a happy face, when asked how they enjoyed it. One participant (**P3**), on the other hand, had a neutral face as their answer.

Just as in the first iteration, the outputs were grouped into and analyzed through several themes and sub-themes (see ***Table 13***). These are presented one-by-one in the following sections, also featuring related raw participant quotes or happenings.

Table 13: Overview of themes and sub-themes the second iteration results were grouped into.

Theme	Sub-Theme(s)
Deviations from the Procedure	Workshop compositions
	No mirroring warm-up
Attention	Distracting elements and self-regulation
	Screen affinity
	Social influences
Structure, Expectations and 'Showing	Novelty versus familiarity
	'Showing, rather than telling'
Comparing the Children and Sessions	Modality-related preferences
	Physical-contact affinity
	Success measures
Strategies and Indications to Dance Together	Connected and combined instances
	Additional strategies and indications
Sound and Movement Connections	-

6.7.1 Deviations from the Procedure

During the workshop sessions, partly as expected, certain matters came up that led to changes in the original procedural plans:

- Workshop compositions

The current work focused, overall, on *dyadic* neurodiverse social interactions – that is, ones between an autistic individual and a NA one. In this second iteration, as aforementioned, a teacher (**T**) was included as the NA individual in interaction. She would both participate in the technology-mediated interaction herself, and supervise the children during the session. This was because there were no other teachers that could take upon one of these roles, in order to help her out. On top of this, she and her students had limited availability in terms of time.

Considering these limitations, only two sessions could be held. **T** brought in with her two autistic participants in each of these workshops. Originally, the intention was that one participant would join another prototype's evaluation, as a part of the other researcher's investigation. The other autistic child and their NA teacher, **T**, on the other hand, would be dancing and moving together – making up the desired neurodiverse pairing for the current investigation. The children would later switch their roles within the workshop session.

Despite this, immediately as **P1** and **P2** walked in, they were arm-in-arm and did not seem to want to get separated from one another. With the relevant guideline themes of *flexibility & adjustability* and *comfort & assistance* (see ***Table 2***) in mind, a decision was then made by the researchers: As per **T'**s support and advice as well, all three participants were allowed to join the session together instead. This change made it so that the technology-mediated social interaction took place in a group that consisted of two autistic individuals and a NA one. The sessions were held in this way for both of the workshops, for consistency; with **P1** & **P2** and **P3** & **P4** each joining **T** together.

- No mirroring warm-up

Prior to the workshop sessions, a discussion was held with **T**. There, she was asked whether she thought a more guided warm-up (i.e. mirroring, like in the first iteration) would be beneficial or counterintuitive for the children.

Originally, the guideline from Oral (2022) regarding abiding to the phases-structure was thought to apply here as well. Thus, the warm-up was intended to be carried out once again. However, **T** had her doubts as to how this would be perceived by the (different) children. As a result, a decision was made to judge the situation during the session itself.

In the first workshop, it was observed that once music started to play, it came very naturally for **P1** and **P2** to immediately start moving and dancing around. Hence, the warm-up stage was skipped for them. This was intentionally kept consistent in the second workshop that involved **P3** and **P4**. These children, however, did seem to respond less to the sound modality and were more often distracted (see ***6.7.4***).

Either way, leaving out the warm-up section in this iterations' workshops turned out to have an added benefit: Without already introducing 'mirroring' to the children as a way they could make a connection with one another, observations could be made regarding how often

this strategy still came up in the sessions. Looking into this usage allowed us to get a sense of any priming effects that might have occurred in the first iteration's set-up.

6.7.2 Attention

In both of the workshop sessions, the partaking children appeared to be very curious. They seemed to be prone to wanting to explore their surroundings and what is present and available. Importantly, their attention span seemed to be shorter when compared to the adult participants in the first iteration.

- Distracting elements and self-regulation

In the context of the technology-mediated social interactions themselves, these qualities meant that there existed some elements that acted as distractors at times, grabbing the children's attention:

For example, ***researchers' personal belongings were clustered together and put on the side, on a bench***. Through doing so, it was intended that the items would be away and kept from being distracting – also abiding to *distractibility* guideline theme (see ***Table 2***). However, this turned out not to go as planned. Different children approached these items at different times, wanting to take a look at or to play with them.

Along the same lines, the school gym was previously chosen as a suitable location, also in consultation with **T**. This was given the connection children already had with this space and moving around. Regardless, it was not formerly disclosed that this space also included a **playground set** on the side. The children, thus, did make use of this space on occasion, during the sessions: They played by climbing onto the set or sled down the slide. Beyond the distraction aspect, this situation also presented some positive points. In particular, it seemed to give the children a way they could self-regulate. They could, as a result of having this option available, make a decision regarding the activities they would like to be occupying themselves with at a given moment.

Similarly, the ***other prototype*** present attracted the children at times, causing them to linger around it, curious to know what it was. This situation, just like with the playground set, was not all negative. In fact, it allowed for observing what it would be like to have these sorts of different technology systems or objects available together, in a given space. These varying items or activities the children can choose to engage with at different moments could inspire a possible use context for the current system being developed.

- Screen affinity

Interestingly, a big attraction element for the children was observed: Specifically, as aforementioned, the researcher controlled sounds and sound changes through a ***laptop***. The children seemed to often want to come closer to this ***screen***, starting from the very beginning of the workshops. It was observed how all of the children excitedly checked this out soon, upon entering the playroom. This interest was also sustained throughout the sessions. They, such as **P4,** even sat next to the researcher, to be able to see and watch what is going on on the screen for longer durations.

An interesting difference in the case of children was observed, when compared to the first iteration adults: Despite one ***camera*** even being held by another researcher, who helped to film the session, the children did not seem to pay much attention to or get bothered by these and being filmed. Their affinity seemed to reside with screens, rather than technological objects overall.

- Social influences

Last but not least, it was observed that the children would frequently **influence each other** in what they focused on or what they did in the space. For example, they would follow each other in coming closer to the main researcher, in order to take a look at the laptop. This held also for their trips onto the playground or them joining the dancing and interacting with **T**.

6.7.3 Structure, Expectations and 'Showing'

Once again, it was highlighted during the workshop sessions how important structure and expectation-setting were for the children, even if at different individual degrees.

- Novelty versus familiarity

For instance, **P3** experienced difficulties related to the level of structure available and hit **T** a few times during the session. Regarding his experience of this interaction context and what role his need for more rigidness played on it, **T** highlighted:

"It's difficult for him, because it's an open game and he was not doing here, what he is always doing here. (...) Because it was open, at the end, he touched me and he hit me <laughs>, and uhh, it was difficult for him. He needs more structure in that sense."

However, with this interaction context being new for her and for the children, **T** was not sure where the solution for this challenge would exactly lay. She thought, for instance, that if the setting would be changed into her playroom, where she works with children one-on-one more often, this could help **P3**. Despite this, she pointed out how *"he is always busy with something else"*. She talked about his persistency in wanting to repeat activities the two of them have done together in the past and thus, his affinity with the known and the familiar – perceiving this as a challenge in their interactions:

"When I take him to the play room and, uh– In the past, I played with him with the bottle of fishes. And [when there,] he always asks me for the bottle <laughs>. I cannot do nothing else <laughs>, because it's very difficult to get him on another mindset."

She saw music as an inherently open element, also for herself: *"This was nice. But it was, for me also, an open thing, so I don't know if I do it good, so, <laughs>."* Neither she nor the children were used to making a social connection through their bodies and movements, nor through music. In her daily interactions with the children, **T** often relied, instead, on connecting to them through toys and play. Hence, they did not have a mental or practical framework as to

how they could begin to do this. This was partly as intended and meant that the teacher and the children both had to further explore the social interaction context and establish their roles in it. The language surrounding this situation hinted towards possible dialogical system processes (Steffensen, 2012) at play, which was deemed as promising.

Another interesting thing to note was that **T** assumed this expectation-setting responsibility to be on herself during the interaction, rather than on the technology system. This was a subtle reminder of the power roles at play between her and the students, as well as a hint towards how the system could or should be positioned in such a use context.

- ‘Showing, rather than telling’

In this expectation-setting regard, **T** was observed to apply what she preaches when it comes to showing, rather than telling the children what to do. In particular, this was a strategy she and the other teachers had already mentioned during the contextual inquiry meeting: They had said that it was easier for the children to understand what is asked of them, if you are to model it for them, rather than only rely saying it using words. However, the workshops made it very clear what this would look like in practice (see ***Figure 19***).

She, for instance, often coupled her verbal expectation-setting prompts with non-verbal cues, such as using very happy or excited facial expressions, as if to try and model a certain emotional response in the situation for her students. This interpretation was based also on the times at which she would choose to make comments. For example, in one of the workshop sessions, she said ‘Oh, dit is gezellig!’ *(’Oh, this is cosy!’)* upon entering the room, which was interpreted as her setting the tone for how the children were to relate to the environment and the context. Importantly, this, once again, was hinting towards the power positionings amongst her and the children present.

As another example, when she exclaimed “Muziek maken!”(/”*Music-making!*”) towards the beginning of the workshops, to repeat to the children the type of activity they are about to engage in, she would additionally make use of orchestra-chef-like hand gestures. Similarly, she would point up as she said “Luisteren!”(/”*Listen!*”), to highlight the music starting to play. Crucially, during **W2**, she also started to move her arms and to dance as a way of ‘showing’ the expectations within the interaction to the children. This, in that case, appeared to be what prompted **P3** to come towards her and to mirror her movement, to also start dancing; showcasing the power of this approach.

Figure 19: Some instances of **T ‘**showing, rather than telling’. *(Left:)* She uses orchestra-chef-like hand gestures while repeating information regarding the musical activity to the children that are about to engage in them. *(Middle & Right:)* When music starts to play or a big change happens in it, she points up while prompting the children to listen and to pay attention.

6.7.4 Comparing the Children and Sessions

Across the different children and the sessions they took part in, there existed certain similarities and differences:

- Modality-related preferences

From the partaking children, it was observed that the ones in **W1** seemed to enjoy the sound modality more than the rest. This was the case, especially for **P1**. Specifically, the combined workshop situation, wherein the children could interact with two different prototypes, had its benefits: It allowed for observing the children’s responses to the different modalities. Furthermore, it helped in putting their certain other behavioral patterns into context, regarding this aspect. The first two children that seemed more engaged with the audio modality, for instance, seemed less engaged with haptics. The opposite situation held for **P3** and **P4**.

A very interesting example event in this sense involved **P4**. As aforementioned, she sat next to the researcher to look at the laptop screen and during, she stretched out her leg (see ***Figure 20***). In doing so, she appeared to almost bump into the speaker in front of her. Thus, the researcher interpreted the situation as the speaker being in her way. She tried to move it away, to the front. However, **P4** reached her foot further towards the speaker, keeping the contact as the researcher attempted this. In connection with her enthusiasm and interest regarding the other prototype that involved haptics, this occurrence was interpreted as the following: She enjoyed the feeling of the vibrations she could get, through having her foot against the speaker.

Figure 20: *(Up:)* **P4** sitting next to the main researcher during the workshop session, in order to look at the laptop screen and point towards sound clips on there, to make the researcher play these. *(Down:)* **P4** stretches out her leg, almost bumping into the speaker in front. Thinking that it may be in **P4**'s way, the researcher attempts to move it away. Instead, **P4** reaches her foot further towards it to keep contact, possibly as she can feel the vibrations in this way and enjoys doing so.

Similar instances favoring haptics were observed where, for example, the children looked down at the vibrations they felt from the floor. Especially **P3**, at times, appeared to be completely disconnected from the ongoing sounds. He, instead, only focused on the physical pushing-each-other interaction he was having with **T** (see ***Figure 21****).*

Figure 21: **T** and **P3** engaging in a physical pushing-each-other interaction.

Overall, such variety in modality-related preferences made for a case for the following: Different technology systems should be created that address these different senses, in order to improve the same interaction context. In this way, the children could be working with what works best for them. These similarities and differences observed could also point towards looking into combinations of these modalities in a single design, as a way of providing support for many individuals at once.

- Physical-contact affinity

Also in terms of how much they enjoyed or seeked physical contact, the children appeared to have some differences:

P2 and **P3** seemed to look for this sort of contact at high levels, for example with the former always holding either **P1** or **T** by the arm. Similarly, **P3** got **T** to spend most of their interaction time during the workshop pushing one other – showing sensory-seeking tendencies. In particular, the two of them had contact by the hand and showed resistance against one another, seemingly enjoying this physical engagement (see ***Figure 21***). Importantly, an observation was made once again that **P3**'s persistence in and sustained attention towards this kind of interaction was not matched by how **T** seemed to experience it. She would keep on engaging for the sake of him, but showed some signs of boredom.

If we are to look at the two other children's affinities in this regard, **P1** seemed to not mind physical contact too much, especially with **P2**, with whom he appeared to be good friends. **P4**, on the other hand, was observed to often show a closed body language upon **T**'s attempts in this. That being said, she did not mind being physically very close and almost touching when she sat next to the researcher, to look at the laptop screen (see ***Figure 20***). Hence, for some participants, this could depend on the other person in interaction and the context.

- Success measures

The children showed their likings of or comfort in the interaction in different ways. For example, they differed in terms of how much (or big) they moved to the music. The other researcher present during the workshops, for instance, observed that **P1** appeared to be very engaged and as though he liked the interaction. However, he did not really use his hands that much. In relation to this, the researcher made the comment that *"maybe it is just a personal thing.",* with which **T** agreed. Hence, this quality seemed to be a non-definite success measurement point, as well as an important consideration, in terms of later movement-detecting-related practicalities.

P4, was another good example of this situation: **T** commented regarding her that she was very comfortable in this technology-mediated interaction context. This was shown, also through her sitting next to the researcher during the session (see ***Figure 20***). In particular, **P4** sat by the laptop, looking at what was happening on the screen as the researcher controlled sounds and sound changes for a long period of time. Not only that, but when she came back to do this again, she took initiative by pointing to a clip on the screen, saying "Die! (*/"This!"*): She wanted the researcher to play what she had selected. Interestingly, this verbal cue later also turned into her 'pressing' the screen as an indication, as this was a way of interacting she was used to, thanks to other tablets. She kept engaged for quite some time; curious about the different options, asking for and listening to these different clips.

T said regarding her, *"She is also a girl who needs a lot more structure, but she loves music, she loves to dance, so I think she was very relaxed in this situation."* This fit the researcher's own observation, as **P4** seemed to enjoy the musical context, but perhaps was not in the mood to be as active with her body and movements-wise. It would have been interesting to see if she would react in the same way, had this interaction been repeated and she could be met with the technology system more than just once.

6.7.5 Strategies and Indications to Dance Together

Instances of all the dancing-together indications and strategies derived from the first iteration were also observed during these workshops. Specifically, this included those that could be used in connection-reflecting, in order to get to reactive embodiment (Spiel et al., 2016): The participants moving closer and even making physical contact, mirroring each other or repeating each other's movements, as well as laughing.

Not all instances of these were discussed in high detail in this section, given that an event description script was also provided for the workshops (see ***Appendix N***). Rather, one of these (i.e. moving closer) was exemplified, followed by a showcasing of other patterns noticed amongst these strategy and indications' occurrences. Furthermore, certain new ways seen to be shown or used by the participants were also given mention.

- Moving closer

Throughout the workshop sessions, the participants frequently walked towards one another to interact more, coming to share the physical space. Similarly, instances were seen of them leaning, even if they did not step towards each other to literally close the distance. For example, towards the beginning of the workshop, following the start of the music, **P1** leaned towards **P2**, who was then prompted to hold **P1** on the arm .

Likewise, **T** often leaned forward to get closer to the children's height when trying to connect and dance together with them. This was an interesting way she was interpreted to reduce the power roles at play momentarily, by leveling herself with her students, at least in this physical sense.

Regarding the use of this physical space, dancing around was yet another indication or strategy that could be validated in this session. For instance, when switching from dancing on his own to moving together with **T**, **P1** not only went towards her, but also behind. This was successful in getting her attention and in getting her to dance with him.

Also with regards to the participants reaching towards the other, examples observed included **T** reaching out both her hands towards **P1**, to try and initiate physical contact, in the form of pushing one another. She would also give one hand each to each kid, to come to a holding-hands circle formation at times. Here, again her instinct in doing so was interpreted to be related to her position as an older adult and as their teacher.

Interestingly, physical contact seemed to be a lot more present within this iteration, in comparison to the previous one. Furthermore, perhaps also arising from existing power imbalances of the teacher and student roles, the use of ***physical 'leading'*** by **T** towards some students caught the researcher's attention (***Figure 22***). In particular, she would hold them by

their hands or arms. She did this, for example, when **P2** interpreted her moves to be one of an airplane; holding **P2** by the arms to make him into a plane and to move him, while in this position and hold.

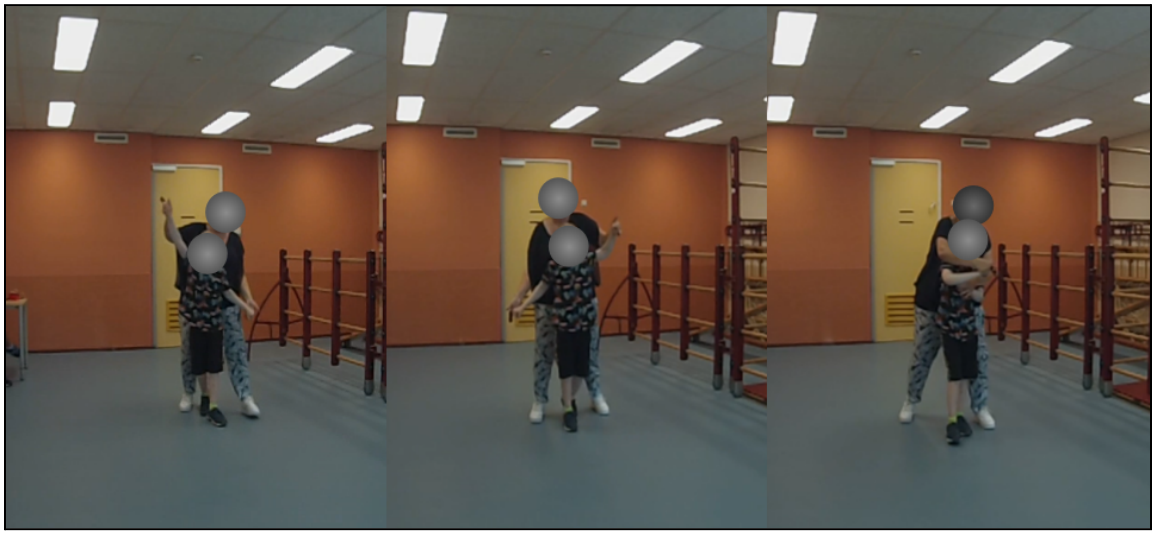


Figure 22: **T** physically leading **P2**, helping him to copy the 'airplane' movement she had previously done herself.

- Connected and combined instances

All of these dancing-together strategies and indications could help to prompt one another or to lead to a progression in connection levels. An example of this was already mentioned, where **P1** leaned towards **P2**, who was then prompted to hold **P1** on the arm. Thus, one interactor leaning could lead to the other initiating physical contact with them.

Similarly, on one occasion, when in physical contact with **P1, P2** mirrored **T** by raising his arm. As this arm was the one he was connected to **P1** by, it led **P1** to also follow suit and mirror this movement.

These strategies and indications could additionally be seen to occur together. For instance, a frequent case of this was events such as **P2** holding **P1** by the arm and the two of them turning together clockwise, with the former going around the latter. In this way, both physical contact and turning around one another would comprise the manner in which the participants danced together.

- Additional strategies and indications

Some other example strategies to dance together included ***different plays on the use of physical distance***. For instance, creating a lead-follow dynamic such as in the game of tag, **T** made use of moving away from **P1**, while still holding her gaze on him. This was as if to get him to follow her as she walks to the rhythm, which was exactly what happened. This movement and body language got **P2**'s attention, who once again faced them, making it a successful attempt to move together and to interact with both children.

Adding onto the existing strategy of mirroring/repetition, ***explicit instructions*** from T were observed at points. This sort of leading dynamic was seen, also across other strategies:

now combining mirroring also with physical contact, **T** took initiative in physically making **P2** do certain moves, together with her. This sustained the questions regarding how existing roles and dynamics play into dancing-together strategies, in different contexts.

Last but not least, adding onto laughter as an indication, it was observed that the children would also otherwise ***scream, vocalize or sing***. These sorts of indications of joy, including ***stimming*** (as **P1** did by flapping his arms can also be interesting success measures the system uses, to detect these instances as those of connection.

6.7.6 Sound and Movement Connections

Just as suggested by Cibrian et al. (2018), it was observed that playing a sudden and random brief high-pitched sound had an attention-grabbing effect. This helped to get participants, such as **P1** and **P3,** back into the main dance area.

Also in general, similarly to one of the 'technology system'/DJ's' observations in the first iteration, a change in sound would, again, prompt a change in movement or connection, bringing the children's attention onto the situation. For instance, they would come by and check the laptop screen or look in the direction of the speakers.

Most importantly, both **T** and the children responded very well to the rhythmic-grounding- and tonal-centering- (Wigram, 2004) inspired baseline chosen, as well as the specific sound contents included. They would match their moves to either of these two dimensions, as well as to the energy. For instance, in response to volume being increased on one occasion, **P1** increased his speed in turning and running around. In this way, patterns similar to former interactive sonification (Frid et al., 2016) literature were observed.

All in all, the contextual inquiry and the workshops included in this second iteration served to support and further extend the insights the first iteration provided us with. Together, these were interpreted through their thematic groupings.

6.8 Results' Interpretation and Discussion

In this section, the raw results from the second iteration and their categorization into themes are interpreted and discussed under the main research questions. These are, at times, added onto what was already gathered through the first iteration.

6.8.1 RQ 1: Autistic versus NA

As highlighted in the first iteration, interestingly, less observations were made than expected during the workshops, with regards to differences in autistic and NA perspectives, perceptions and expectations. While this did not majorly change in the second iteration, more instances could be captured. This was also given a more mixed group of (autistic) participants, for instance, in terms of gender. Hence, one possible interpretation of not seeing as many variances could be supported: Also currently – but especially in the first iteration – it was a struggle to be able to capture the heterogeneity of the autistic experience (Lord et al., 2020) and

thus, the possible challenges that may arise from these not matching NA social expectations and experiences.

Importantly, there additionally existed an age difference between the NA teacher and the autistic children currently partaking. Certain social expectations and behaviors that authentically arose from the autistic embodied experience seemed less toned down in this iteration, given that the participants were younger and less adapted into a society with neurotypical-dominant norms. For instance, events were observed where they clearly showed a 'restricted' interest or a preference for familiar actions: **P1** wanted to repeatedly go back to the movement interaction where he, **P2** and **T** were in physical contact via holding hands, turning and running around. This was regardless of whenever **T** would try and introduce a different move. Similarly, **P3** wanted to continuously engage with **T** through an interaction where they pushed each other by the hands (see ***Figure 21***).

A particular challenge regarding this point was speculated on and highlighted as another avenue the technology system could try and intervene and mediate for: The NA interactor's own sense of boredom – or will to 'move on to something else' – could clash with the autistic interactor's want to keep (sharing) the enjoyment that they are deriving from a particular movement, while in interaction. They could be working with a different set of expectations and a framework in this sense, with NA interactors associating change and adding elements with 'progressing' in the social interaction and connection more. For autistic participants, on the other hand, a possible measure of connection could be related to the amount of fun they might have (together), which in this case could come from sticking with the particular moves they are getting the most out of, while being able to share this with someone else, for a long duration of time.

In the first iteration, mainly, different importances put on eye contact as a dancing-together strategy and indication were discussed. There, a related challenge was highlighted where an autistic interactor made use of mirroring/repetition, without (a first) eye contact with their counterpart. In turn, this NA participant was not aware of these attempts at connecting. Also in the second iteration, it was observed that eye contact was not something the autistic children necessarily relied on in interaction. Therefore, this mismatch of importance or of habituality of making eye contact confirmed the aforementioned possible design direction: Perhaps, it could be signaled through changes in the technology system's sound outputs that there is an ongoing connection-making attempt, in order to help alleviate this challenge.

This idea would be supported by observations in the second iteration, where sound changes were observed to help in drawing the participants' attention. Through such reactive embodiment (Spiel et al., 2016) that is 'adaptive' to and revealing of connection-making intentions, the participants could be supported in not missing these attempts from one another, simply due to them interpreting the ongoing social interaction through their own ideas and expectations and not their counterparts'.

In this way, through the two iterations, some promising differences in autistic and NA social expectations and experiences could be highlighted, as potential addressing points in the future work.

6.8.2 RQ 2: Mappings' connection-prompting role

As far as coming up with movement-to-sound mappings that can prompt connection between autistic and NA individuals goes, the first iteration revealed a set of qualities that define and prompt feeling connected (see ***Table 14***).

Table 14: The set of qualities that define and prompt feeling connected, based on first iteration results.

What does it mean to feel connected? How can connection-making be prompted?
'Revealing'/getting to know
Agency & reciprocation
Physical prompting emotional connection
Interpersonal dynamics & roles

In addition, a set of challenges in the way of achieving connection was highlighted through that same iteration (see ***Table 15***). These were not further and specifically focused-upon in the second iteration, especially given the methods made use of: With the children participants, the same type of lengthy discussions could not be held, which more easily leads to these kinds of conceptual understandings.

Table 15: The set of challenges that get in the way of achieving connection, based on first iteration results.

Challenges in connection-making or prompting this
Synchrony in when (and how) to connect
Expression of intention
Meaning of dancing together
Need for common ground & mutual adjustments
Hints from (unclear) language - Not a straightforward action or concept
Having to focus on yourself or the input/environment

However, these established meanings and prompters of, as well as challenges regarding connection-making could, at times, be supported by observations made in this iteration. For instance, the aforementioned mismatch between autistic participants' tendency to continue and sustain and NA ones' to further develop or, otherwise, feel stuck in interaction was an example of *non-synchrony in how to connect* amongst the two groups. Similarly, the impact of our individual cultural backgrounds and otherwise life-worlds on our *meaning of 'dancing together'*; the types of strategies we can come up with was exemplified again. Specifically, *different plays*

on the use of physical distance were showcased as a possible movement input category, differently than in the first iteration. Interestingly, this seemed to draw from the dynamics of a game of tag, which was more present in these set of participants' daily contexts.

Just as in the first iteration, further details for these mappings could be collected, which are mentioned in the following section:

6.8.3 RQ 3: Mappings: Open/adjustable versus pre-determined

The data from the previous iteration had already shown strong evidence regarding a need in having adjustable and personalizable movement-to-sound mappings. Hence, it was no surprise that the second iteration also supported this claim.

In particular, within the conceptual system design that inspired the set-up of the workshop sessions in this iteration, there existed an act (i.e. act 0 - *get to know / tune in*), in which the participants' individual preferences would be determined. This had to do with the lines of difference, as collected through the first iteration: melody- versus rhythm-, as well as slow-melody versus fast-melody likings. The application of this was observed to be quite successful, as both the teacher and the children responded very well to the sound outputs: These were personalized for them based on a procedure that made use of these noted lines of differences.

Similarly, the individuality of the participants', as well as groups of participants' movements could be highlighted in an even more particular way. This was due to the NA interactor being the same one (i.e. the teacher) in all of the different sessions in this iteration, which allowed for certain unique observations: With whom she was dancing and moving together with seemed to have an impact in the types of moves that come up in the interaction, beyond her own tendencies in moving and dancing.

What's more, cases were noted, such as **P1**, to everyone involved in the interaction, seeming to be very engaged in the technology-mediated interaction and especially the sound modality involved. Despite this, he did not, for instance, use his hands much and had smaller moves he engaged in. Likewise, **P4** seemed not to be in the mood to be very physically active in the interaction. Regardless, she was very relaxed in the interaction and was engaged for longer durations in different ways (e.g. sitting next to the researcher to observe and listen). Hence, suitable 'movement inputs' for these kind of mappings would, indeed, better not be defined in terms of specific actions, but rather broader concepts; showing flexibility based on individual participants' qualities and tendencies.

6.8.4 RQ 4: Measuring connection/success

During the second iteration, the following dancing-together strategies and indications were focused upon and built on within the conceptual design: *moving closer (including physical contact), mirroring/repetition* and *laughter*. Together, they were seen as aspects of interactor actions that could be used as movement inputs, as they indicate different levels of connection. Specifically, these would be used for reactive embodiment (Spiel et al., 2016) purposes, for connection-reflection.

Through the workshop sessions, it could be successfully confirmed that such instances come up, also within this in-school use context, in the NA teacher and her autistic children's

interactions. Some of such strategies and indications from the first iteration were not specifically focused on, but also found their place in these (i.e. eye contact, ‘matching’). Similarly, new possible pointers were observed, which could be further looked into, for in-system evaluation purposes:

- Different plays on the use of physical distance,
- Explicit instructions (i.e. from the teacher towards the students),
- Screaming, vocalizing or singing, and
- Stimming.

When it came to assessing the workshops’ and the technology-mediated interactions’ overall evaluations, age-appropriate three-point smiley scale’s use, as well as making use of the teacher’s expertise and experience regarding the children and reading their emotions, wants and needs were deemed as helpful. However more and better methods can always be used and established, which would allow for a more accurate and equal measure of success.

6.8.5 RQ 5: Real-life product

Perhaps the research question the second iteration efforts could contribute the most to answering were regarding the translation of the conceptual system design into a real-life product. Specifically, useful insights could be gathered for the in-school context, as well as suitable scenarios and goals for the system. This was based on the setting and the participants these workshops incorporated, in addition to the contextual inquiry conducted prior to the workshops.

Schools such as the one collaborated with were highlighted as being very suitable to home this kind of a technology system, for a number of reasons. For instance, these were spaces in which the different sensory needs and ways of communicating of the children were taken into consideration and accommodated.

What’s more, the teachers already valued multi-sensory approaches (especially also music) and had examples of these around. They would even make these themselves, like **T** did with the different ‘sensory bottles’ This included technological objects like the current one, which would have the added benefit of associated costs not getting in the way of the children accessing these products as much. Regardless, especially for this latter sense, there was space to explore the use of such systems further.

As opposed to a one-time exposure like in the current workshops, this product being used across a longer time period and on multiple occasions could make it more beneficial, though further investigations in this regard are needed. This is because, for instance, in the case of getting introduced to other new activities at school as well, the children are given the opportunity to repeat and to get accustomed to these. Furthermore, how they respond to and partake in the technology-mediated interaction can depend on many factors; even their mood on the day of. Therefore, it would be interesting to observe the influences these have on our goals and the success of the system.

Last but not least, through the validation of the conceptual design for the system in this context, the goal for it or the role it could play could also be affirmed, with this bi-directional mediation approach (Oral, 2022) being seen as promising.

7. Discussion

The current work set out to move towards social systems (Steffensen, 2012) that are more understanding, inclusive and welcoming of neurodivergent ways of interacting. The overall aim was to create a *diversity computing (DivComp)* (Fletcher-Watson et al., 2018) device, that is, a technology system that helps autistic and non-autistic (NA) people to better understand, adapt to and appreciate each other's social behaviors and expectations. Specifically, how neurodiverse dyadic social interactions can be supported with multi-sensory embodied interaction technologies was investigated. An overall design-theoretical framework based on *embodied (social) sense-making* (Hummels & Van Dijk, 2015; Van Dijk & Hummels, 2017; Van Dijk, 2022) was applied and a *participatory research-through-design (RtD)* methodology was followed, moving through two reflective design and research iterations. As a result, *Move4Music*, an adaptive sound system that prompts connection during neurodiverse dyadic interactions, was designed (see ***6.1***). Specifically, *Move4Music* would aid in co-located neurodiverse *participatory sense-making*, through the mapping of real-time measurements of body movements onto sound feedback.

The data collected through the two design and research iterations indicated that overall, *Move4Music* can be a promising technology-mediated neurodiverse dialogical system (Steffensen, 2012) example, which can inspire the creation of other such technology-mediated neurodiverse dialogical systems. In particular, one contribution of the current work had to do with this reframing of the challenges experienced in neurodiverse encounters: A point was made, stating that the weight of achieving a mutually satisfactory structural coupling should not be solely put on the shoulders of the autistic participants, nor should this 'satisfaction' be assessed solely from the perspective of NA ones (Crompton et al., 2021; De Jaegher, 2013; Milton, 2012). Rather than replicating existing interventionist approaches and perspectives, a call was made to move towards a different role the technology systems could take upon themselves: They could act as mediators between autistic and NA individuals and thus, help to establish and keep balance points across these individuals' varying, possibly conflicting values and goals in interaction contexts (Oral, 2022). If we are to try and improve these interactions despite existing challenges such as the double-empathy problem (Crompton et al., 2021), technology should be positioned in a way where better grounds for this process to take place could be provided.

Rather than staying at this theoretical level of what such a system should or should not be like or do, the current work took advantage of a key ability of RtD approaches: These could help to produce more concrete knowledge that was applied onto a specific situated context (Stappers & Giaccardi, 2017). Hence, a second practical contribution had to do with the framework that could be provided, regarding how one could create a technology-mediated dialogical system (Steffensen, 2012) for this autistic and NA (dyadic) interaction context:

Prior to engaging in any research or design activities, an initial list of insights were gathered based on existing autism interaction technologies, as well as fields such as music therapy and adaptive music (see ***Table 1***). The outputs of the first and the second iterations were utilized to revisit this list. These teachings from the iterations were looked at from this lens, in order to see per insight, whether and how they were applied, whether they succeeded or else,

how they were altered for the final conceptual design of and recommendations for *Move4Music* (see ***Appendix O***). Hence, a list of design requirements and recommendations were put forward that was supported not only by different theoretical bases, but also the current practical endeavors.

Throughout the RtD process, other such sources of information and inspiration for the creation of other technology-mediated dialogical systems were also presented. These included the set of qualities that define and prompt feeling connected (see ***Table 14***), as well as challenges that get in the way of achieving connection, based on first iteration results (see ***Table 15***). In addition practical insights could be drawn regarding methodological and procedural plans, through which to create these (see ***Table 2***).

7.4 Limitations

As insightful as it was, the current work was not without its limitations. For instance, despite the intention in and the use of participatory methods, the degree to which participants, especially the autistic children within the second iteration could contribute to different aspects was limited. In the end, proxy users such as their (NA) teacher still had to be relied upon for information collection. Furthermore, while in the first iteration and through the in-service WoZ (Boye & Wirén, 2007; Wirén et al., 2007), participants (including autistic ones) could contribute to the conceptual design through their participation, success measures were still derived from these data points that were collected by the (NA) researcher herself. The analyses and the interpretation of the occurrences in the workshop sessions and their translation into design elements could involve the intended user group more.

Regardless of the participation of also autistic individuals in the sessions held, within the first iteration, for example, only female participants were partaking, with no adult male autistic perspectives being represented. In general, due to existing practical limitations such as time and access to the target population, also in the second iteration, only two workshop sessions could be held. This meant that only four autistic children were involved. Autism, however, is highly individual and variable (Lord et al., 2020). Thus, repetition of the procedures, also with larger and more diverse samples are needed.

Last but not least, within the second iteration, the conceptual system design and the movement-to-sound mappings included were evaluated through a WoZ set-up. However, this relied on how well or quickly the researcher could spot the movement inputs/dancing-together (intention/invite) occurrences and make the according sound alterations. Given that this was rather overwhelming, with a lot to be in tune with at all times, it is possible that she was not able to act at a technology system's levels of accuracy. Furthermore, subjective influences that come from human perception abilities are always a possibility within this set-up, not making it a full replacement of testing an actual working (autonomous) technology system.

7.5 Future Directions

Through the current work, several directions came up that future efforts could focus on, regarding the *Move4Music* system itself:

For instance, one such point was regarding how it could be turned into an actual autonomous technology system. Attempts at and investigations into how the dancing-together strategies and indications could be translated into what this system could sense, using what kind of sensors remained a question.

Similarly, further try-outs regarding the movement input and the sound output contents of *Move4Music* would be useful in different ways. The comparison of these alternatives along different directions, for instance, could allow us to better pinpoint their effects and help us to build a strategy regarding how to choose amongst these options. Along these lines, collaborations could be made with musicians to create original compositions that fit the lessons learned from former studies in literature and through the current work, rather than using pre-made audio content. This could give more control and precision into the designs and their intended influences on neurodiverse dyadic interactions.

Another promising direction would be to explore more of the likely contexts for a product such as *Move4Music*. While this was a first look at an in-school one, involving a teacher and her multiple students, also contexts such as home environments, together with family or friends could be suitable for the product. We could gather different insights into opportunities and challenges that can arise through and within these different set-ups that involve different number of peoples, as well as different interpersonal dynamics at play.

Last but not least, looking into also repeated and longer-term use of and interaction with *Move4Music* could enrich our understanding on several aspects. It could, for instance, help to further refine the imagined use context. In addition, we could learn about, for example, how movement-to-sound connections may evolve over time, as the interactors grow more familiar with these, the interaction context and one another.

---

8. Conclusion

As a result of this Master's thesis work, an investigation was carried out regarding how neurodiverse dyadic social interactions can be supported with multi-sensory embodied interaction technologies. Specifically, the capacity of *dialogical systems* in supporting change and transformation in *social systems* was utilized. Because of the existing *double-empathy problem* that complicates neurodiverse interactions (and achieving a mutually-satisfactory *structural coupling*), the following goal was set for the current efforts: Creating a technology-mediated dialogical system that help autistic and non-autistic (NA) people in better understanding, adapting to and appreciating each other's social behaviors and expectations. In doing so, it was intended to work towards social systems that are more understanding, inclusive and welcoming of neurodivergent ways of interacting. Through a *participatory research-through-design (RtD)* methodology that had the *embodied sense-making* design-theoretical framework at its core, two reflective design and research iterations were carried out. A number of different methodical approaches such as literature research, co-design, (contextual) interviews, questionnaires and expert reflections/discussions were made use of, to, in the end, reach a first sketch of the system. This included a list of design guidelines and requirements. Given autism and sound relations, such as within music therapy contexts, the movement and sound modalities were deemed relevant for this piece of embodied interaction technology. Specifically, the design challenge was set as (designing the) mapping (of) real-time measurements of body movements can onto sound feedback, to create an interactive system that aids in co-located neurodiverse participatory sense-making. An *in-service Wizard-of-Oz (WoZ) data collection* set-up comprised the first iteration. This was so that insights and expertise from (autistic and NA) dancers and music therapists can be obtained in an embodied way. Inspiring the design of the intended *Diversity Computing (DivComp)* device, these workshops informed what aspects of the interaction/ (joint) movements the device can aim to perceive, and how these can be mapped onto different sound and sound changes. Consequently, aspects such as strategies participants use to (try and) dance with one another (e.g. mirroring/repetition, lessened distance) and sound aspects that can help to support connection-building (e.g. volume, unexpected sounds) were collected. These were made into a set list of movement-to-sound mappings, that got tested within the second iteration, again through WoZ. Specifically focusing on the intended use context, technology-mediated interactions between autistic children and their teacher were observed for further refinements in the design. The set of dancing-together strategies were affirmed, while the iteration also revealed important different characteristics in the target user group and their interactions (e.g. attention), to give some examples. For future directions/work, attempting at a working (autonomous) technology system, exploring other use contexts and longer-term (repeated) interaction setups were pointed out.

References

Abedi Koupaei, M., Poushaneh, K., Mohammadi, A. Z., & Siampour, N. (2013). Sound therapy: an experimental study with autistic children. *Procedia-Social and Behavioral Sciences*, *84*, 626-630.

Al-Ayadhi, L. Y., Majeed Al-Drees, A., & Al-Arfaj, A. M. (2013). Effectiveness of Auditory Integration Therapy in Autism Spectrum Disorders--Prospective Study. *Autism Insights*, (5).

American Psychiatric Association (APA) (2013) DSM-V Diagnostic and Statistical Manual of Mental Disorders. 5th Edition, American Psychiatric Association, Washington, DC.

Benjamin Leonard, C. (2022). *Wearable device for interactive and collaborative sound making for autistic people* (Master's thesis, University of Twente).

Benton, L., Johnson, H., Ashwin, E., Brosnan, M., & Grawemeyer, B. (2012, May). Developing IDEAS: Supporting children with autism within a participatory design team. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 2599-2608).

Berger, D. S. (2002). *Music therapy, sensory integration and the autistic child*. Jessica Kingsley Publishers.

Boye, J., & Wirén, M. (2007). Multi-slot semantics for natural-language call routing systems. In *Naacl'07 workshop Bridging the gap: Academic and industrial research in dialog technologies. Rochester, NY, USA. April 26-26 2007* (pp. 68-75).

Bratteteig, T., & Wagner, I. (2012, August). Disentangling power and decision-making in participatory design. In *Proceedings of the 12th Participatory Design Conference: Research Papers-Volume 1* (pp. 41-50).

Braune, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Researchin Psychology, vol. 3, núm. 2.

Brbić, I., & Tomić, L. (2020). An integrative review of the effectiveness of the tomatis method in children with autism spectrum disorder. *RAD CASA-Medical Sciences*, *543*, 49-56.

Brown, A., & Kerr, T. (2009). Adaptive music techniques. In *Proceedings of Improvise: the Australasian Computer Music Conference 2009* (pp. 26-31). Australasian Computer Music Association.

Cibrian, F. L., Mercado, J., Escobedo, L., & Tentori, M. (2018, May). A step towards identifying the sound preferences of children with autism. In *Proceedings of the 12th EAI international conference on pervasive computing technologies for healthcare* (pp. 158-167).

Cooke, B., & Kothari, U. (Eds.). (2001). *Participation: The new tyranny?*. Zed books.

Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New directions for evaluation*, *1998*(80), 5-23.

Crompton, C. J., DeBrabander, K., Heasman, B., Milton, D., & Sasson, N. J. (2021). Double empathy: why autistic people are often misunderstood. *Neuroscience*, *9*(554875), 4-11.

De Jaegher, H. (2013). Embodiment and sense-making in autism. *Frontiers in integrative neuroscience*, *7*, 15.

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

Edlund, J., Gustafson, J., Heldner, M., & Hjalmarsson, A. (2008). Towards human-like spoken dialogue systems. *Speech communication*, *50*(8-9), 630-645.

Eklund, R. (2004). *Disfluency in Swedish human–human and human–machine travel booking dialogues* (Doctoral dissertation, Linköping University Electronic Press).

Fachner, J. (2017). Music, Moments, and Healing Process: Music Therapy. In *The Routledge Companion to Music Cognition* (pp. 89-99). Routledge.

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

Fraser, N. M., & Gilbert, G. N. (1991). Simulating speech systems. *Computer Speech & Language*, *5*(1), 81-99.

Frauenberger, C., Good, J., Alcorn, A., & Pain, H. (2013). Conversing through and about technologies: Design critique as an opportunity to engage children with autism and broaden research (er) perspectives. *International Journal of Child-Computer Interaction*, *1*(2), 38-49.

Frauenberger, C., Good, J., & Pares, N. (2016a, May). Autism and technology: beyond assistance & intervention. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 3373-3378).

Frauenberger, C., Makhaeva, J., & Spiel, K. (2016b, May). Designing smart objects with autistic children: Four design exposès. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 130-139).

Frayling, C. (1994). Research in art and design (Royal College of Art Research Papers, vol 1, no 1, 1993/4).

Frayling, C. (2015) Closing Provocations of the 2015 Research through Design conference. http://rtd2015.herokuapp.com/programme/ frayling

Frid, E., Bresin, R., Alborno, P., & Elblaus, L. (2016). Interactive Sonification of Spontaneous Movement of Children—Cross-Modal Mapping and the Perception of Body Movement Qualities through Sound. *Frontiers in neuroscience*, *10*, 521.

Garness, J., Giving, B., Heidebrink, T., Hein, A., Humbert, R., Janorschke, E., ... & Bass, J. D. (2016). Auditory integration interventions for children with autism and developmental disabilities: An evidence-based practice project.

Gaver, B., & Bowers, J. (2012). Annotated portfolios. *interactions*, *19*(4), 40-49.

Hales, C. (1986). Analysis of the engineering design process in an industrial context.(PhD). *University of Cambridge, Cambridge*.

Hayden, C. (2023). *Different, Not Less (A neurodivergent's guide to embracing your true self and finding your happily ever after)*. Murdoch Books.

Heaton, P. (2003). Pitch memory, labelling and disembedding in autism. *Journal of Child Psychology and Psychiatry*, *44*(4), 543-551.

Hess, K. L., Morrier, M. J., Heflin, L. J., & Ivey, M. L. (2008). Autism treatment survey: Services received by children with autism spectrum disorders in public school classrooms. *Journal of autism and developmental disorders*, *38*, 961-971.

Höök, K., & Löwgren, J. (2012). Strong concepts: Intermediate-level knowledge in interaction design research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *19*(3), 1-18.

Hummels, C., & Van Dijk, J. (2015, January). Seven principles to design for embodied sensemaking. In *Proceedings of the ninth international conference on tangible, embedded, and embodied interaction* (pp. 21-28).

Hutchings, P. E., & McCormack, J. (2019). Adaptive music composition for games. *IEEE Transactions on Games*, *12*(3), 270-280.

Isenhower, R. W., Marsh, K. L., Richardson, M. J., Helt, M., Schmidt, R. C., & Fein, D. (2012). Rhythmic bimanual coordination is impaired in young children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *6*(1), 25-31.

Jotform (n.d.). Jotform Form Builder [Online application]. Retrieved from https://www.jotform.com/

Kalas, A. (2012). Joint attention responses of children with autism spectrum disorder to simple versus complex music. *Journal of Music Therapy*, *49*(4), 430-452.

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

Karen, H., & Sandra, J. (2017). Contextual inquiry: A participatory technique for system design. In *Participatory design* (pp. 177-210). CRC Press.

Katagiri, J. (2009). The effect of background music and song texts on the emotional understanding of children with autism. *Journal of music therapy*, *46*(1), 15-31.

KDE (n.d.). Kdenlive [Computer software]. Retrieved from https://kdenlive.org/en/download/

Keay-Bright, W. (2007). The reactive colours project: demonstrating participatory and collaborative design methods for the creation of software for autistic children.

Kientz, J. A., Goodwin, M. S., Hayes, G. R., & Abowd, G. D. (2014). Interactive technologies for autism.

Kim, J., Wigram, T., & Gold, C. (2008). The effects of improvisational music therapy on joint attention behaviors in autistic children: a randomized controlled study. *Journal of autism and developmental disorders*, *38*, 1758-1766.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and human behavior*, *31*(5), 354-364.

Leman, M. (2007). *Embodied music cognition and mediation technology*. MIT press.

Lim, Y. K., Stolterman, E., & Tenenberg, J. (2008). The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *15*(2), 1-27.

Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., ... & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature reviews Disease primers*, *6*(1), 1-23.

Louchart, S., Truesdale, J., Suttie, N., & Aylett, R. (2015). Emergent narrative: past, present and future of an interactive storytelling approach. In *Interactive digital narrative* (pp. 185-199). Routledge.

Lovaas, I., Newsom, C., & Hickman, C. (1987). Self‐stimulatory behavior and perceptual reinforcement. *Journal of applied behavior analysis*, *20*(1), 45-68.

Lyra, L., Rizzo, L. E., Sunahara, C. S., Pachito, D. V., Latorraca, C. D. O. C., Martimbianco, A. L. C., & Riera, R. (2017). What do Cochrane systematic reviews say about interventions for autism spectrum disorders?. *Sao Paulo Medical Journal*, *135*, 192-201.

Malinverni, L., Mora-Guiard, J., Padillo, V., Mairena, M., Hervás, A., & Pares, N. (2014, June). Participatory design strategies to enhance the creative contribution of children with special needs. In *Proceedings of the 2014 conference on Interaction design and children* (pp. 85-94).

Maturana, H. R., & Varela, F. J. (1987). *The tree of knowledge: The biological roots of human understanding*. New Science Library/Shambhala Publications.

Mayo Clinic. (n.d.). Autism spectrum disorder. Retrieved July 5, 2023, from https://www.mayoclinic.org/diseases-conditions/autism-spectrum-disorder/symptoms-causes/syc-20352928

Mazefsky, C. A., Herrington, J., Siegel, M., Scarpa, A., Maddox, B. B., Scahill, L., & White, S. W. (2013). The role of emotion regulation in autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(7), 679-688.

McKenzie et al. (2014). Ableton Live 11 Intro [Computer software]. Retrieved from https://www.ableton.com/en/shop/live/

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

MI.MU Gloves Limited (2010). Gliss [App]. Retrieved from https://www.apple.com/app-store/gliss/

MI.MU Gloves Limited (2010). Glover [Computer software]. Retrieved from https://mimugloves.com/glover/

Moens, B., Muller, C., Van Noorden, L., Franěk, M., Celie, B., Boone, J., ... & Leman, M. (2014). Encouraging spontaneous synchronisation with D-Jogger, an adaptive music player that aligns movement and music. *PloS one*, *9*(12), e114234.

Monobanda. (n.d.). Visible Voice. Retrieved July 5, 2023, from https://monobanda.eu/project/visiblevoice

Mural (n.d.). Mural [Online visualization and collaboration tool]. Retrieved from https://www.mural.co/

Nguyen, T. (2021). *Stim4Sound: a Diversity Computing device helps to alleviate the double empathy problem* (Master's thesis, University of Twente).

Oral, L. (2022). *The State of The Art of Autism Interaction Technologies: From Normative Interventions Towards Bi-Directional Mediation.* [Unpublished manuscript]

Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music perception*, *26*(5), 489-504.

Pares, N., Masri, P., Van Wolferen, G., & Creed, C. (2005). Achieving dialogue with children with severe autism in an adaptive multisensory interaction: the" MEDIATE" project. *IEEE Transactions on Visualization and Computer Graphics*, *11*(6), 734-743.

Rajagopalan, S., Dhall, A., & Goecke, R. (2013). Self-stimulatory behaviours in the wild for autism diagnosis. In *Proceedings of the IEEE International Conference on Computer Vision Workshops* (pp. 755-761).

Ragone, G., Good, J., & Howland, K. (2020, June). OSMoSIS: Interactive sound generation system for children with autism. In *Proceedings of the 2020 ACM Interaction Design and Children Conference: Extended Abstracts* (pp. 151-156).

Ragone, G., Howland, K., & Brulé, E. (2022, June). Evaluating Interactional Synchrony in Full-Body Interaction with Autistic Children. In *Interaction Design and Children* (pp. 1-12).

Roozenburg, N. F., & Dorst, K. (1998). Describing design as a reflective practice: Observations on Schön's theory of practice. In *Designers: The key to successful product development* (pp. 29-41). London: Springer London.

Schön, D. (1983). The reflective practitioner basic books. *New York*.

Shahrudin, F. A., Dzulkarnain, A. A. A., Hanafi, A. M., Jamal, F. N., Basri, N. A., Na'im Sidek, S., ... & Khalid, M. (2022). Music and Sound-Based Intervention in Autism Spectrum Disorder: A Scoping Review. *Psychiatry Investigation*, *19*(8), 626.

Shaker, N., Togelius, J., & Nelson, M. J. (2016). Procedural content generation in games.

Sharma, S. R., Gonda, X., & Tarazi, F. I. (2018). Autism spectrum disorder: classification, diagnosis and therapy. *Pharmacology & therapeutics*, *190*, 91-104.

Sheridan, J., & McAuley, J. D. (1997, September). Rhythm as a cognitive skill: temporal processing deficits in autism. In *Proceedings of the Fourth Australiasian Cognitive Science Conference*.

Simonsen, J., & Robertson, T. (Eds.). (2012). *Routledge international handbook of participatory design*. Routledge.

Simpson, R. L. (2005). Evidence-based practices and students with autism spectrum disorders. *Focus on autism and other developmental disabilities*, *20*(3), 140-149.

Singh, A., Piana, S., Pollarolo, D., Volpe, G., Varni, G., Tajadura-Jiménez, A., ... & Bianchi-Berthouze, N. (2016). Go-with-the-flow: tracking, analysis and sonification of movement and breathing to build confidence in activity despite chronic pain. *Human–Computer Interaction*, *31*(3-4), 335-383.

Sinha, Y., Silove, N., Hayen, A., & Williams, K. (2011). Auditory integration training and other sound therapies for autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews*, (12).

Spiel, K., Makhaeva, J., & Frauenberger, C. (2016, February). Embodied companion technologies for autistic children. In *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 245-252).

Spiel, K., Malinverni, L., Good, J., & Frauenberger, C. (2017, May). Participatory evaluation with autistic children. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 5755-5766).

Srinivasan, S. M., & Bhat, A. N. (2013). A review of “music and movement” therapies for children with autism: embodied interventions for multisystem development. *Frontiers in integrative neuroscience*, *7*, 22.

Stappers, P. J. (2014). Prototypes as a central vein for knowledge development. In *Prototype: Design and craft in the 21st century* (pp. 85-97). Bloomsbury Academic.

Stappers, P. J., & Giaccardi, E. (2017). Research through design. In *The encyclopedia of human-computer interaction* (pp. 1-94). The Interaction Design Foundation.

Steffensen, S. V. (2012). Care and conversing in dialogical systems. *Language Sciences*, *34*(5), 513-531.

Stuart, N., Whitehouse, A., Palermo, R., Bothe, E., & Badcock, N. (2023). Eye gaze in autism spectrum disorder: a review of neural evidence for the eye avoidance hypothesis. *Journal of Autism and Developmental Disorders*, *53*(5), 1884-1905.

Thompson, E. (2010). *Mind in life: Biology, phenomenology, and the sciences of mind*. Harvard University Press.

Torrance, S. (2005). In search of the enactive: Introduction to special issue on enactive experience. *Phenomenology and the cognitive sciences*, *4*, 357-368.

Trevarthen, C., & Daniel, S. (2005). Disorganized rhythm and synchrony: Early signs of autism and Rett syndrome. *Brain and development*, *27*, S25-S34.

Van Dijk, J. (2022, February). Zooming in on embodied social sensemaking: Mapping the design space in the context of videoconferencing. In *Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 1-10).

Van Dijk, J., & Hummels, C. (2017, March). Designing for embodied being-in-the-world: Two cases, seven principles and one framework. In *Proceedings of the eleventh international conference on tangible, embedded, and embodied interaction* (pp. 47-56).

Van Dijk, J., Kopke, M., van Huizen, N., van Uffelen, L., & Beunk, L. (2019, March). Empowering young adults on the autistic spectrum: Reframing assistive technology through design. In *4th RTD Conference: Design United*.

Vargas, S., & Lucker, J. R. (2016). A quantitative summary of The Listening Program (TLP) efficacy studies: What areas were found to improve by TLP intervention?. *Occupational Therapy International*, *23*(2), 206-217.

Varela, F. J. (1991). Organism: A meshwork of selfless selves. In *Organism and the Origins of Self* (pp. 79-107). Dordrecht: Springer Netherlands.

Wensveen, S., & Matthews, B. (2015). Prototypes and prototyping in design research. *The routledge companion to design research*, 262-276.

Whipple, J. (2012). Music therapy as an effective treatment with autism spectrum disorders in early childhood: A meta-analysis. *Early childhood music therapy and autism spectrum disorders: Developing potential in young children and their families*, 59-76.

Wigram, T. (2004). *Improvisation: Methods and techniques for music therapy clinicians, educators, and students*. Jessica Kingsley Publishers.

Wirén, M., Eklund, R., Engberg, F., & Westermark, J. (2007). Experiences of an in-service Wizard-of-Oz data collection for the deployment of a call-routing application. In *Bridging the Gap: Academic and Industrial Research in Dialog Technologies Workshop Proceedings, NAACL-HLT, Rochester, NY, April 2007.* (pp. 56-63). Omnipress.

Wooffitt, R., Fraser, N., Gilbert, N., & McGlashan, S. (1997). Humans, computers and wizards: Conversation analysis and human (simulated) computer interaction.

Wu, B., Wun, S., Lee, C., & Horner, A. (2013, November). Spectral Correlates in Emotion Labeling of Sustained Musical Instrument Tones. In *ISMIR* (pp. 415-420).

Yannakakis, G. N., & Togelius, J. (2015, September). Experience-driven procedural content generation. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)* (pp. 519-525). IEEE.

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., ... & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism research*, *15*(5), 778-790.

Zhang, Y. (2023). *Stim4Sound: a musical interactive system promoting communication between autistic people and society* (Master's thesis, University of Twente).

Appendices

Appendix A: Improvisational Techniques in Music Therapy

Table A.1: An overview and categorization of improvisational techniques in music therapy (Wigram, 2004; *rephrased and made into a table*).

Category	Name	Description	Use & Benefits	Additional Remarks
Empathy - Related Techniques	*Imitation*	The echo/ repetition of the client's response by the therapist, after that response has ended	- drawing attention to the client's own actions - reinforcing their response/ communication - showing that their behavior is accepted - offering turn-based play as an example interaction - ^giving them an opportunity to lead/control	Imitation of, e.g. sound, rhythm, interval, melody, movement, facial expression, verbalization, etc.
	^*variation: Pair imitation*	Two clients imitating one another		
	Synchronization	Doing what the client is doing at the same time	- supporting/ stabilizing/ enhancing client responses - promoting their self-awareness - bringing in more involvement/ intimacy - creating opportunities for a leadership role for them, allowing	Different levels, e.g. modality (uni versus cross), single versus multiple-element (e.g rhythm, melody, dynamics)

			acceptance - developing empathic understanding in them	
	^*variation: Musical mirroring*	The therapist playing each melody/interval in the opposite direction/ movement		
	^*variation: 'Face to face' motion interaction game*	The therapist making the same movements/ activities as the client, but in reverse direction, as in a mirror		
	Incorporation	Using a musical motif or client's behavior as a theme for your own improvisation/ composition and then developing it further	- reinforcing the client to present a musical motif - promoting acceptance of their music - as a model for musical creativity & expression - building a musical repertoire that can be used in therapy	at the moment in which the motif is presented or at a later time
	Exaggerating	Expressing something magnified that differentiates or makes the client unique in their behavior	- bring to their attention the quality of the client	e.g. rhythm, timbre, melody, interval, etc. NOT to imitate/ridicule

Structuring Techniques	*Rhythmic grounding*	The therapist maintaining a basic beat/ providing a rhythmic one ostinato for improvisation by the client	- helping the client organize their improvisation according to an underlying pulse - stabilizing their pace - promoting feelings of security and stability - providing the physical and psychological support to the client's performance and keeping them 'on track'	NOT to control/restrict the client's improvisation by the rhythmic basis, remaining flexible
	Tonal centering	The therapist providing a tonal center, scale, or harmonic basis for the client's improvisation	- to ensure client success - helping them to organize their melodies - stimulating musical thinking - evoking certain emotions and moods - providing structure and reassurance	e.g a restriction on the tones the client may use in a particular key, giving a bass line, adding a pedal point/melodic ostinato
	Shaping	In a musical context, the therapist helping the client to determine the length and shape of a complete musical idea	- if the client is unable to translate their feelings into musical ideas/phrase units - developing thematic material - the therapist leading the improvisation	e.g. chord accompaniment, target point (climax tone in the beginning, middle or the end of a phrase), energy (crescendi, diminuendi, accelerandi, ritardandi), form (rondo -

				ABCADA, chorus, verse (ABABAB)
Elicitation Techniques	*Repeating*	The therapist offering the same rhythm or melody several times, either sequentially or at short intervals	- repetition exactly the same way leading to predictability and satisfaction, reassurance, stability and consistency; clear expectations	- ending at a 'resting point' versus not and musical question asking & answering - engaging the client in a musical interaction and helping them to notice and remember the musical motif
	^*<u>variation</u>: with (the same) changing elements*	Making changes so the client expects not just the next step, but also how the whole thing will end	- creating feelings of: predictability, premonitions, encouragement, adventure, stability, change, expectation, built-up attention	e.g. a melody repeated several times, louder and louder each time
	^*<u>variation</u>: with (different) changing elements*		- a clear expectation that *something* will happen, but the result cannot be foreseen - creating feelings of: anticipation, inevitability, unease	
	Modelling	The therapist exhibiting or demonstrating a target behavior, quality, feeling, trait/property, etc.		e.g. a musical motif or movement to imitate

	Making space	The therapist improvising and creating space within the structure of the patient improvisation, to allow the client to answer or fill in sounds		i.e. - empty silence (e.g. a rest) - filled silence (e.g. sustaining tones/chords) - providing a musical background that invites participation
Techniques to go in a different direction	*Introducing change*	The therapist trying to improvise the client by introducing a new thematic material or by adding a new section to the start improvisation	- if the client is stuck in his actions and does not want to go in a different direction - if the client wants to do something new, but has reached an impasse	i.e. stopping the ongoing musical process and making noticeable changes in the rhythmic or melodic motifs
	Differentiating	When the therapist improvises simultaneously with the client, distinguishing and dividing the two sides, with theirs being very different from what the client is playing, but go along with it/are compatible	- differentiating between the two musical identities and establishing independence between the roles while still maintaining a relationship	e.g. rhythms, melodies, timbres, movements, registers, forms, etc.
	Modulating	The therapist gradually changing the time/key signature	- to move from one mood to another	~musical equivalent of changing your perspective/ attitude

Intimacy-Related Techniques	*Playing the same instrument/ sharing instrument*		- greater intimacy - developing coordination and a working relationship - creating reciprocity	
Procedural Techniques	*Experimenting*	In a musical or movement context, the therapist providing a structure, a procedure or an idea, to guide the client's improvisation and asking the client to experiment with all their possibilities of expression	- helping the client test various alternatives or choices - promoting creativity	
	Conducting/ Leading	The improvisation being led by one person through expressive gestures, musical symbols, verbal messages or other signs	- providing opportunities to learn leader and follower roles - building self-confidence - building group cohesion	
	Playing Back	The therapist recording and playing the client's/group's improvisation afterwards/at a later time	- promoting self-awareness	tangible result of their efforts as a source of pride
	Reporting	(Immediately after an improvisation or during/after playback) asking	- cultivating self-awareness - aiding interpretation	especially useful if with someone else and we want to know more about their

		the client to report on various aspects of the improvisational experience, including about what actually happened musically, any thoughts they had at any point, etc.		relationship and the underlying theme
	Reacting	^The client narrating what they felt about the result	- obtaining information about the client's self-image - clarifying attitudes and feelings towards the assignments	e.g. what they liked/disliked
Techniques for Exploring Emotions	*Interpreting*	The therapist offering possible explanations or meanings for the experiences of the client		
	Distancing/ Metaprocessing			e.g. asking the client to listen and observe what is going on in the improvisation, or 'Where is the group standing at the moment?'; going from spontaneous to thoughtful consciousness

Appendix B: Prototype Software/App Interfaces

Figure B.1: A snapshot from the **Glover software** (Version 1.1.2; MI.MU Gloves Limited, 2010). There, you cn example set of inputs' (i.e. different sliders from an app interface) representation *(Left)*, as well as their connection to various sound change/effect outputs to a software that allows for music creation *(Right)*. The below part shows real-time data from the app that movements inputs come from.

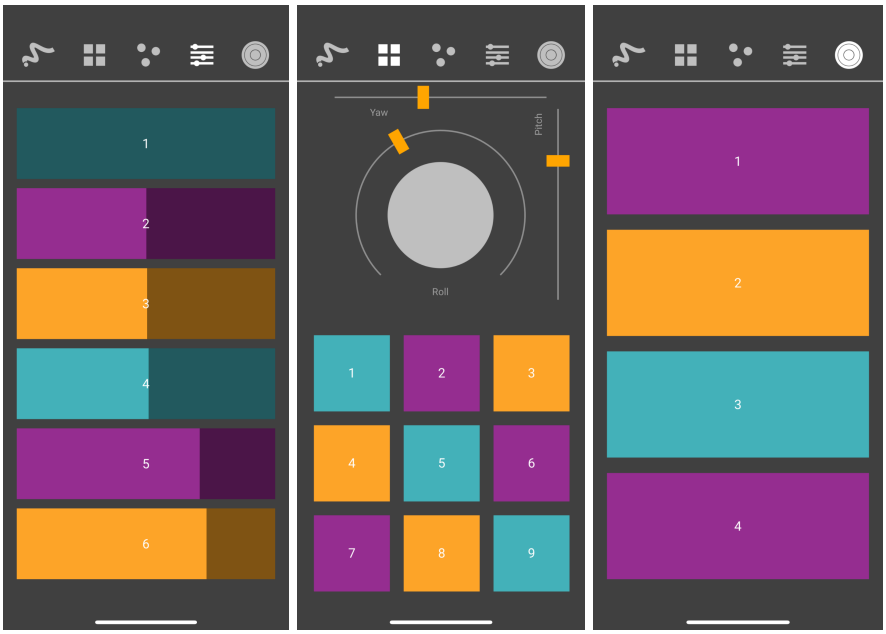


Figure B.2: Example different input options included in the (free) **Gliss app** (Version 1.0.0; MI.MU Gloves Limited, 2010) interface. Specifically, sliders (*Left)* pads and phone orientation captured along three axes (*Middle),* larger buttons (*Right)*.

Figure B.3[20]: A visual comparison of the traditional linear *(Up)* versus vertical *(Down)* interface views the Ableton Live 11 software offers.

[20] [Untitled screenshot of Ableton Live 10/11 linear and vertical arrangements, taken from a YouTube video.]. Learn Live 10: Arrangement View. https://www.youtube.com/watch?v=riOD-fnyCsg.

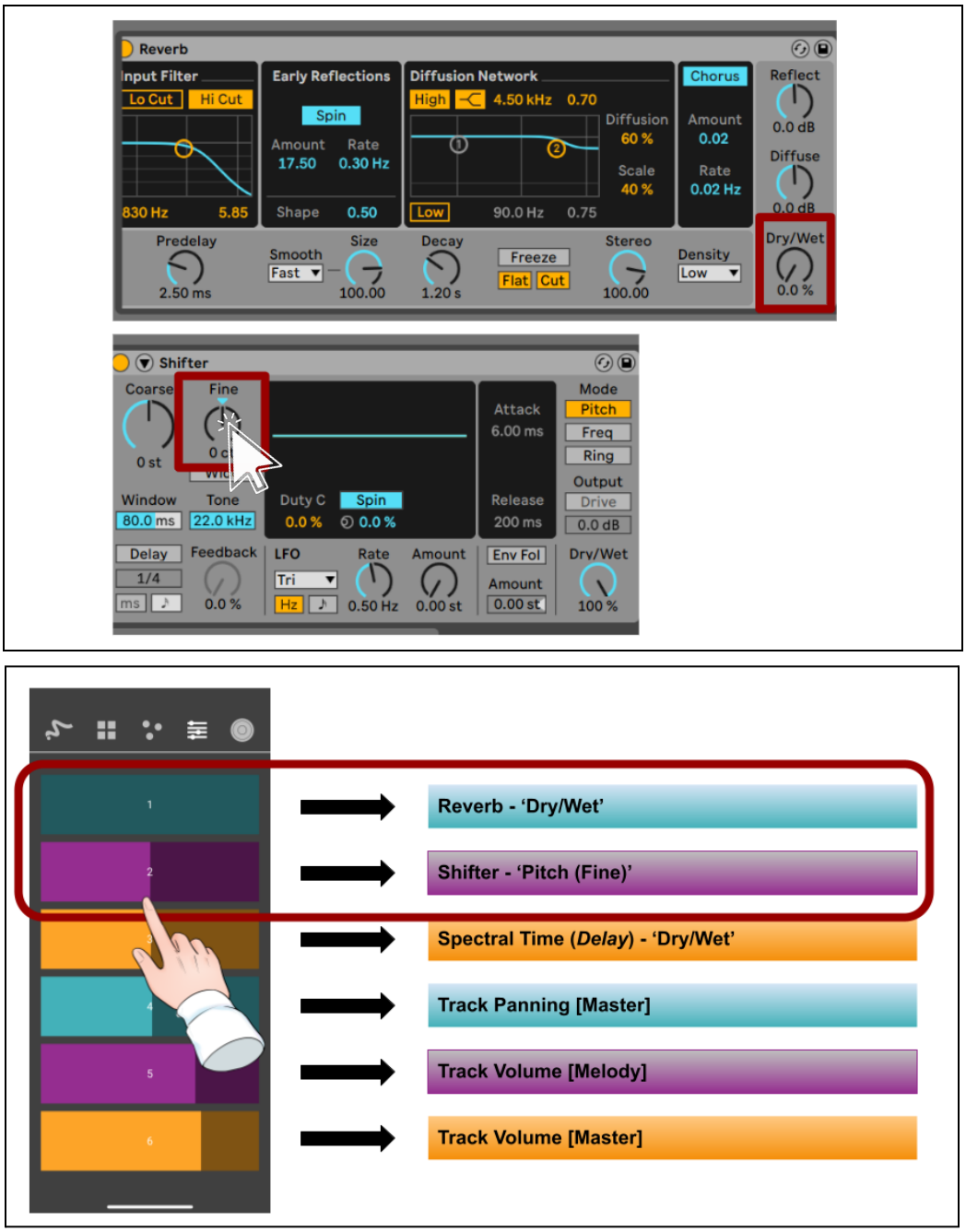
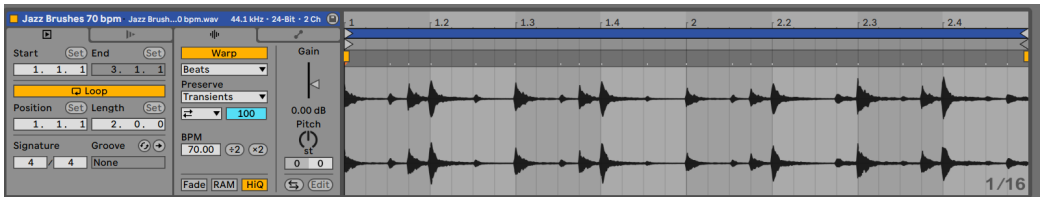


Figure B.4: A visual comparison of the use complexity of the Ableton Live 11 computer *(Up)* and the Gliss app phone *(Down)* prototype interfaces. The latter has a much more intuitive and direct way of manipulating different sound changes' levels, using touch-based sliders, and features a lot less additional and technical language, details and options.

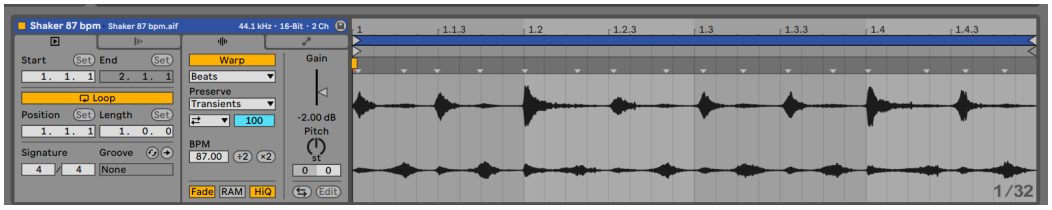
Appendix C: All Sound Material Used *(First Iteration)*

Rhythm/Beats:

- (r1) Jazz Brushes 70 bpm

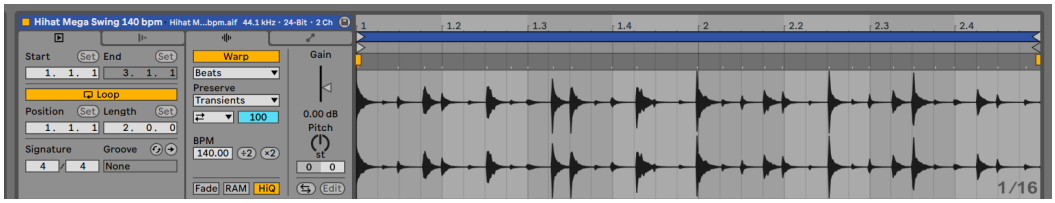


- (r2) Shaker 87 bpm

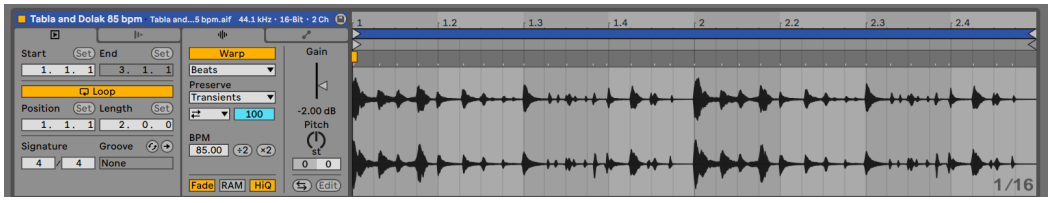


- (r3) Garage Room 135 bpm

- (r4) Hihat Mega Swing 140 bpm

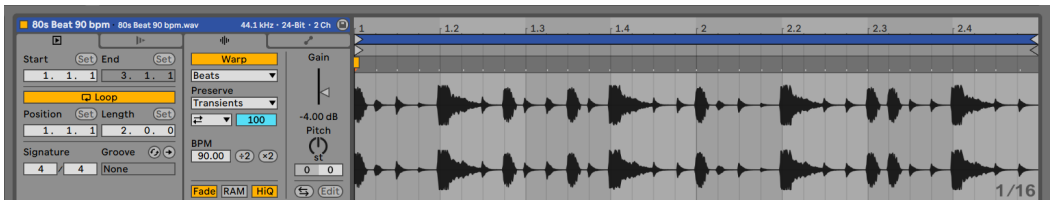


- (r5) Tabla and Dolak 85 bpm[21]

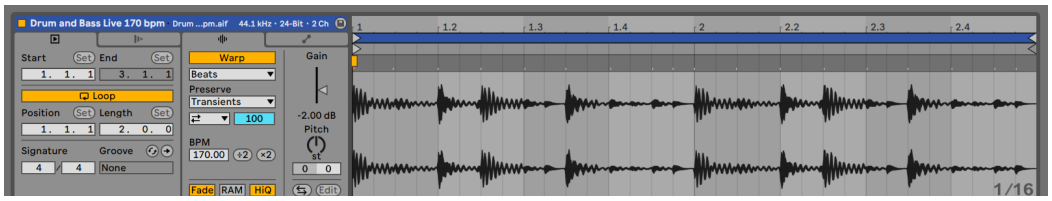


[21] Also previously referred to in text as the ‘bongo Latin beat’.

- (r6) 80s Drum Machine Wingman 120 bpm



- (r7) Drum and Bass Down 170 bpm



Upbeat Melody & Effects:

- Loveline TL E 100 bpm
- (1) energetic-funky-groove-138634
- (2) energetic-funky-groove-138634
- (1) jazzy-abstract-beat-11254
- (2) jazzy-abstract-beat-11254
- (3) jazzy-abstract-beat-11254
- (4) jazzy-abstract-beat-11254

Upbeat Melody 2:

- (1) upbeat-disco-pop-145047
- (2) upbeat-disco-pop-145047
- (3) upbeat-disco-pop-145047
- (4) upbeat-disco-pop-145047

Slow Melody:

- Organic Bells TL Amin 115 bpm
- (1) cosmic-glow-6703
- (2) cosmic-glow-6703
- (3) cosmic-glow-6703
- (4) cosmic-glow-6703

Sound Change Possibilities:

- Reverb[22] effect could be added, using the associated 'Dry/Wet Mix' parameter on Ableton Live 11. This was to control how much reverb is present in the sound output", by controlling the balance between the input signal (dry) & the processed signal (wet).

- 'Shifter' and specifically, its 'Pitch (Fine)' parameter on Ableton Live 11 would allow for changing the amount of transposition (i.e. the amount of pitch changes/the notes going up or down).

- 'Spectral Time'[23] and once again, the associated 'Dry/Wet Mix' parameter on Ableton Live 11 would allow for adding delay effects.

- The sound outputs (in their entirety) could be played to sound as though they are coming more from the left or the right side, using 'Panning' in Ableton Live 11 on the 'Master/Mixer Track'.

- The volume of the sound outputs could be increased or lowered in their entirety, using the 'Master/Mixer Track' in Ableton Live 11.

- The volume of the melodic sound outputs could be increased or lowered, using the associated track in Ableton Live 11. For example, this could decide how predominant rhythm/beats versus melody contents would be, when played/layered together.

All Combined:

Ableton Live project file

---

[22] An audio effect that can be applied to a sound signal to simulate reverberation, which is when sound waves are sent out in all directions in a space, and may be reflected upon surfaces and sustained for some time in the space, in lowered amplitudes.

[23]When using effects featured in the 'Spectral Time' feature, the sound keeps going in this 'frozen' state and gets sustained.

Appendix D: Demographics *(First Iteration)*

Table D.1: The division of participants' level of experience in and state of playing a musical instrument/making music.

	No Experience	In the Past	Actively
# Participants	5	3	2
Participants	P1, P2, P3, P6, P10	P4, P5, P9	P7, P8
# Years	-	(all) 1-3	(all) 10 or more

(within those that do have musical experience:)

Table D.2: The division of the type of instrument(s) musical participants play.

	Keyboard	String	Woodwind	Percussion	Songwriting
# Participants	3	2	1	1	1
Participants	P5, P7, P8	P7, P8	P4, P7	P7	P9

Table D.3: The division of musical participants, in terms of whether they sing or not.

	Yes	No
# Participants	3	2
Participants	P4, P7, P8	P5, P9

Table D.4: The division of participants, in terms of how many years of dance experience they have.

	1-3	4-6	7-9	10 or more
# Participants	4	1	2	3
Participants	P3, P6, P8, P9	P4	P1,P7	P2, P5, P10

Table D.5: The division of participants, in terms of their dance habits/frequency.

	Not in the past 3 months	**Once a week**	**2-4 times a week**	**5-7 times a week**
# Participants	1	2	6	1
Participants	P8	P7, P10	P1, P2, P3, P4, P5, P6	P9

Appendix E: Information Letter *(First Iteration)*

05/04/2023

Participant Information Sheet

~

Music Technologies for Neurodiverse Interactions

Dear participant,

I would like to invite you to take part in a **workshop** for my **Master's thesis**, as a part of the MSc Interaction Technology study program. Before you decide, you need to understand why the research is being conducted and what it would involve for you. Please take your time to read the following information carefully. Ask questions if anything you read is not clear or if you would like more information. Take time to decide whether or not to take part.

Who I Am & What This Study Is About

My name is **Lara Oral** and I am a second-year MSc Interaction Technology student. I am conducting this research as a part of my thesis, and will be using the results I obtain for the write-up of my final report.

The overall aim of my thesis research is to build **technologies that support social interactions between different groups**, such as autistic and non-autistic individuals, and do so in a way that is not an 'intervention' that aims to change the ways of one party (e.g. 'teach' autistic children to make more eye contact using robots), but that mediates a **two-way** process of better **understanding, adapting to and appreciating** each other's social expectations and behaviors.

Specifically, this technology is to incorporate **movement and sound**, and we are to look into how the (joint) movements of a duo can be mapped onto sounds/sound changes, in a way that prompts a sense of **connection** between them as a result of the interaction that takes place.

What Will Taking Part Involve?

Taking part in (this section of) my research will involve joining a *music & movement workshop*:

- The (in-person) workshop will take place (most likely) somewhere between **April 10th and 21st**, at the **University of Twente** campus.

- It will take a maximum of **1 hour**.

- You will participate together with **two other participants** in your workshop, one of which may be the researcher herself (Lara Oral).

- During the workshop, you will be given **guided and more free** music & movement exercises and possibly the role of **controlling the music/music elements**.

- You will be asked to **answer questions and discuss opinions** with the researcher and the others, regarding your workshop experiences.

- With your consent, the music & movement exercises will also be **video and audio-recorded**, for later analysis.

- Through this workshop, the aim is to get inspiration on the design of the technology system to be designed.

Why Have You Been Invited to Take Part?

You have been invited to take part in this study as you - as a dancer/dance teacher, musician/ music therapist or similar - **put thought into communicating through movement and sounds on a regular basis**, and have shown interest in helping me with my research (thank you!).

Do You Have to Take Part?

Please note that participation is **completely voluntary** and that you have the right to refuse participation, refuse any question and withdraw at any time **without any consequence** whatsoever.

What Are The Possible Risks & Benefits of Taking Part?

There are **no possible risks** of participating in this research, aside from possible **sound sensitivities**, which you will be asked about and accommodated for, so that you would not be put in any uncomfortable situations.

The possible benefits of taking part include having a pleasant experience with others where you get to interact in a different way than you normally would, through music and movement.

Will Taking Part Be Confidential?

As I will use a WhatsApp group chat with the other participants for general information and instructions, the study will not be anonymous. However, if I need to contact you personally, I will not do so via that group chat for your privacy. The personal data collected will remain confidential as no one other than me (and the UT staff involved in the thesis) will have access to it.

How Will Information You Provide Be Recorded, Stored & Protected?

The workshop itself (with your consent) will be video and audio-recorded. The data will be recorded/stored for research purposes on a **laptop, password-protected** and accessible only by the researchers themselves. This data will be saved in a safe research drive after the research has been conducted for **at most 10 years**.

The identities of the participants will be protected by the use of **participant numbers** in any analysis conducted or any reporting. Any visuals that will be used in our reports will ensure **anonymity (e.g. by blurring faces)**.

All participants are voluntarily participating, ensuring that the participants' guardians are in full knowledge of de research, privacy of data and data storage. You reserve the right to request access, rectification or erasure of personal data.

What Will Happen to The Results of The Study?

The data collected will be used for the **Master's thesis and its possible publications**, where the (anonymized) results, transcripts and the filled-in forms will be included in the final report.

Who Should You Contact for Further Information?

For any further information you can contact the student researcher Lara Oral through the email l.oral@student.utwente.nl.

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact dr. Jelle van Dijk, the supervisor for the thesis, through the email jelle.vandijk@utwente.nl.

Additionally, you can contact the Secretary of the Ethics Committee of the Faculty of Electrical Engineering, Mathematics and Computer Science at the University of Twente through ethicscommittee-cis@utwente.nl.

THANK YOU!

Appendix F: Consent Form *(First Iteration)*

Consent Form
~
Music Technologies for Neurodiverse Interactions

*YOU WILL BE GIVEN A COPY OF THIS INFORMED CONSENT FORM

Please tick the appropriate boxes	**Yes**	**No**
Taking part in the study		
I have read and understood the study information dated [DD/MM/YYYY], or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	□	□
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and can withdraw from the study at any time, without having to give a reason.	□	□
I consent to be audio-recorded	□	□
I consent to be video-recorded	□	□
I consent that the recorded video files may be used in the report, with the visuals blurred to ensure the participants' anonymity.	□	□
I agree that my information can be quoted anonymously in research outputs	□	□
Use of the information in the study		
I understand that the information I provide will be used for graduation reports and possible future publications.	□	□
I understand that any personal information collected about me that can identify me, such as [e.g. my name], will not be shared beyond the study team.	□	□

Future use and reuse of the information by others

I give permission for the anonymised transcripts, audio and video *recording* that I provide to be archived in a secure file on a computer, so it can be used for future research and learning. After at most 10 years, the files will be destroyed from the safe research drive. □ □

Signatures

_____	_____	_____
Name of participant	Signature	Date

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

_____	_____	_____
Lara Oral	Signature	Date

Study contact details for further information:

Lara Oral *(student researcher)*, l.oral@student.utwente.nl

dr. Jelle van Dijk *(supervisor)*, jelle.vandijk@utwente.nl

Contact information for questions about your rights as a research participant

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee Information & Computer Science: ethicscommittee-CIS@utwente.nl

Appendix G: Procedure Script *(First Iteration)*

(as preparation:)

to the interactors: “Let’s first listen to some sound options <’technology’ person> will have, to see if everything is okay for you all, for example in terms of any sound sensitivities.”

/

to the ‘technology system’/DJ: “I will first show you the app/software, through which you can play sounds and make changes in different sound aspects like <frequency, …>. We will take a look at the different options (and how the interactors like them, considering sound sensitivities, etc.), to get you more familiar with these. You can also make notes here and look at this paper throughout the session, to remind yourself of the options.”

(before the warm-up and main interactions’ start:)

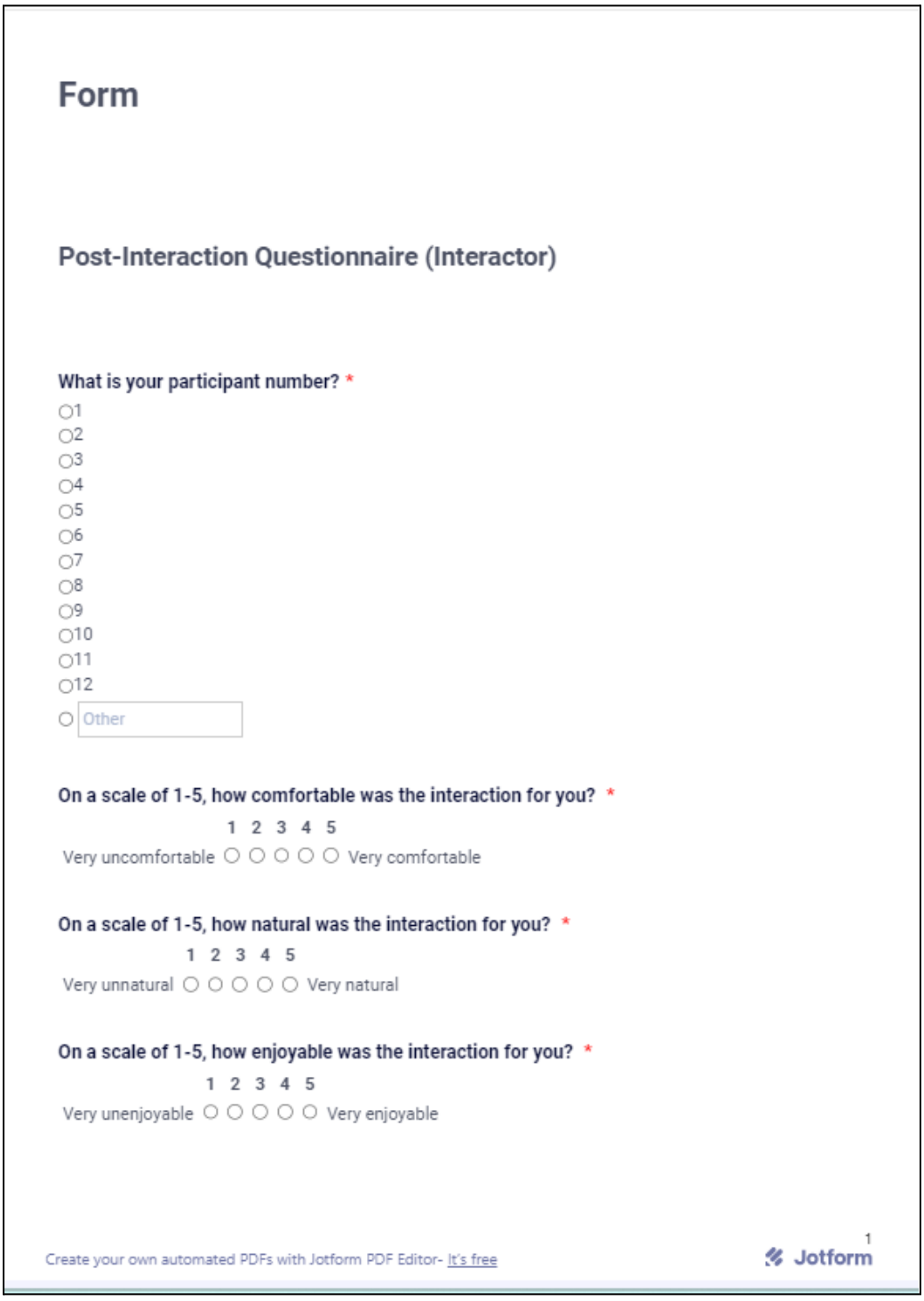
to the interactors: “We will first warm-up through a more guided activity. I will play some music (rhythms/…) and I would like one of you to move/dance to it in a way that you find comfortable, natural and enjoyable. The other should try to copy/mirror you and try and do the same movements the best they can. There is no right or wrong way to move, whether you initiate the movements or try to copy the other’s. I/the ‘technology system’/DJ will then say ‘Switch!’ and you will switch roles. Now, the other will initiate the (dance) movements and be copied. We will do this a few times.”

/

to the ‘technology system’/DJ: “During this main interaction, I would like you to play music/make changes in sound, to which the other two can move/dance to. (...) The goal is for them to build a connection through moving together, so I want you to watch them to see if you think this is happening or not and to try and help, through the sounds you play, the changes you make. You can try out different things (like <...>), there is no right or wrong; I want to see how you approach this.”

Appendix H: Post-Interaction Questionnaire *(Interactor)*

Online Fillable Form Link:
https://form.jotform.com/loral/post-interaction-questionnaire-inte



Form

Post-Interaction Questionnaire (Interactor)

What is your participant number? *

- ○1
- ○2
- ○3
- ○4
- ○5
- ○6
- ○7
- ○8
- ○9
- ○10
- ○11
- ○12
- ○ Other

On a scale of 1-5, how comfortable was the interaction for you? *

	1	2	3	4	5	
Very uncomfortable	○	○	○	○	○	Very comfortable

On a scale of 1-5, how natural was the interaction for you? *

	1	2	3	4	5	
Very unnatural	○	○	○	○	○	Very natural

On a scale of 1-5, how enjoyable was the interaction for you? *

	1	2	3	4	5	
Very unenjoyable	○	○	○	○	○	Very enjoyable

Please feel free to use this textbox to explain your answers.

What, if anything, did you enjoy about the interaction? *

What, if anything, did you NOT enjoy about the interaction? *

What, if anything, would you change about the interaction (e.g. sounds used, set-up, duration), and in what way? *

On a scale of 1-5, how much of a connection did you feel with the other person, through the interaction? *

	1	2	3	4	5	
No connection at all	○	○	○	○	○	A very strong connection

Please feel free to use this textbox to explain your answer.

The End

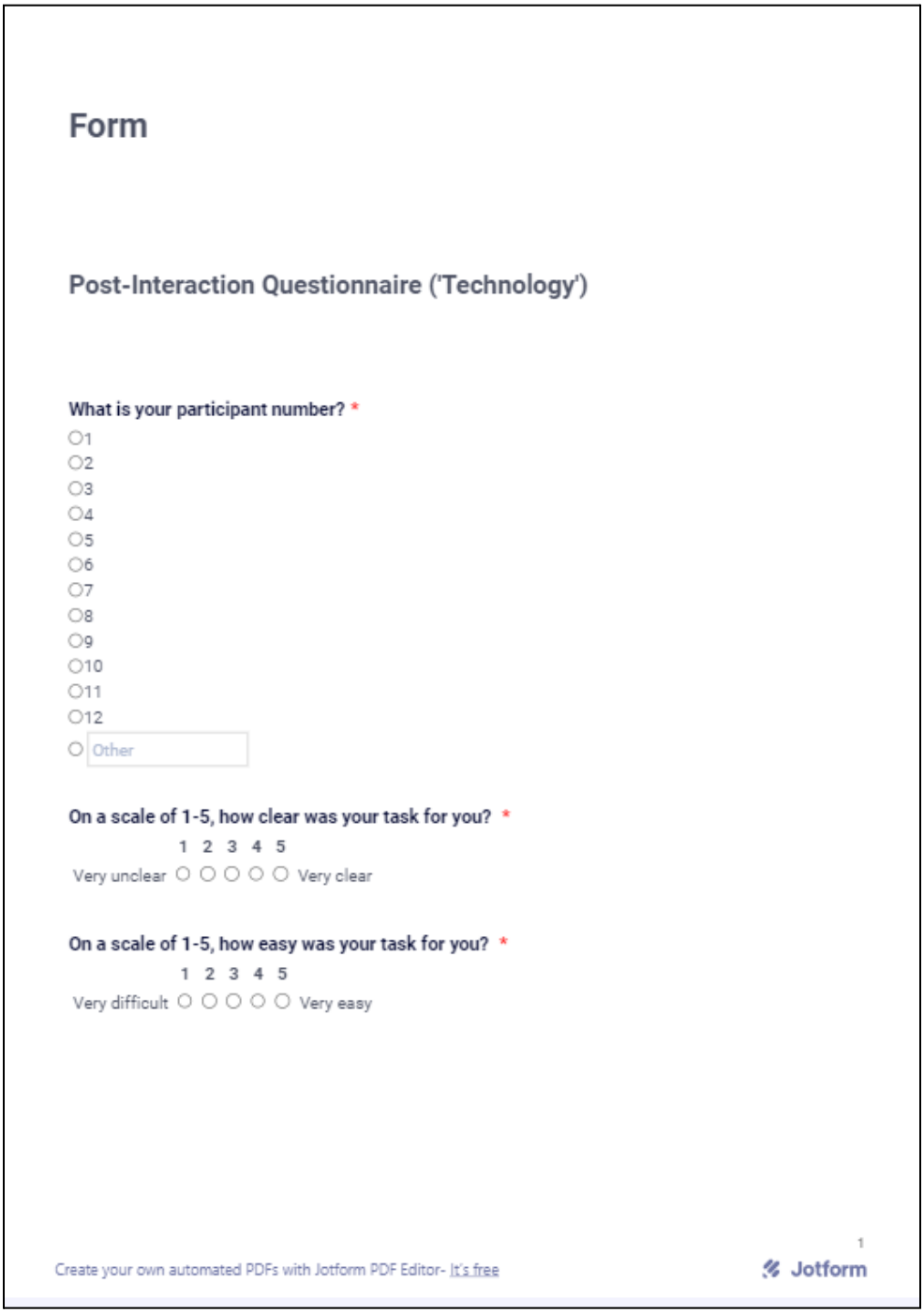
Thank you so much for your participation :) Please remember to submit the form!

If there is anything you would like to add/clarify regarding any of the questions or your answers, or further comments on the interaction, feel free to use this text box.

Submit

Appendix I: Post-Interaction Questionnaire *('Technology System'/DJ)*

Online Fillable Form Link:
https://form.jotform.com/loral/post-interaction-questionnaire-tech



Form

Post-Interaction Questionnaire ('Technology')

What is your participant number? *

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7
- ○ 8
- ○ 9
- ○ 10
- ○ 11
- ○ 12
- ○ Other

On a scale of 1-5, how clear was your task for you? *

	1	2	3	4	5	
Very unclear	○	○	○	○	○	Very clear

On a scale of 1-5, how easy was your task for you? *

	1	2	3	4	5	
Very difficult	○	○	○	○	○	Very easy

Please feel free to use this textbox to explain your answers.

What, if anything, did you find easy about the task? *

What, if anything, did you find difficult about the task? *

What, if anything, did you enjoy about the provided sound options? *

What, if anything, did you NOT enjoy about the provided sound options? *

What, if anything, would you change about provided sound options, and in what way? *

Please briefly explain what your strategy was in trying to make the others feel a connection through the changes you make in sounds. For example, what observations did you make and how did you act based on these? What kind of sounds/sound changes did you try? Were any particularly helpful or unhelpful? *

On a scale of 1-5, how successful do you think you were in making the others feel a connection, through the interaction? *

	1	2	3	4	5	
Very unsuccessful	○	○	○	○	○	Very successful

Please feel free to use this textbox to explain your answer.

The End

Thank you so much for your participation :) Please remember to submit the form!

If there is anything you would like to add/clarify regarding any of the questions or your answers, or further comments on the interaction, feel free to use this text box.

Appendix J: Data Analysis Using MURAL *(First Iteration)*

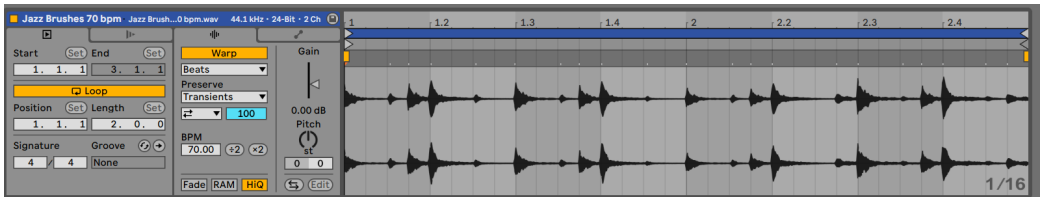
Raw Results Link:
Study 1 - Raw Results.pdf

Thematic Mapping Link:
Study 1 Results - Thematic Mapping.pdf

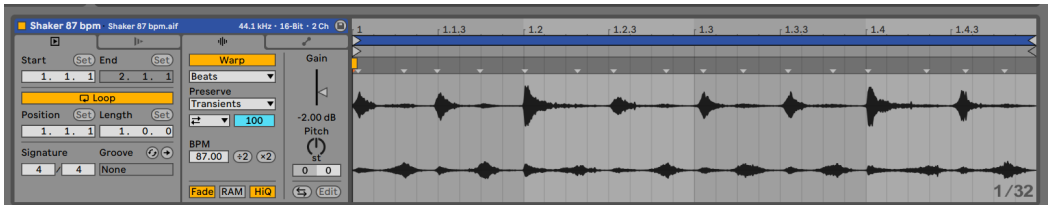
Appendix K: All Sound Material Used *(Second Iteration)*

Rhythm (Prog.):

- Jazz Brushes 70 bpm

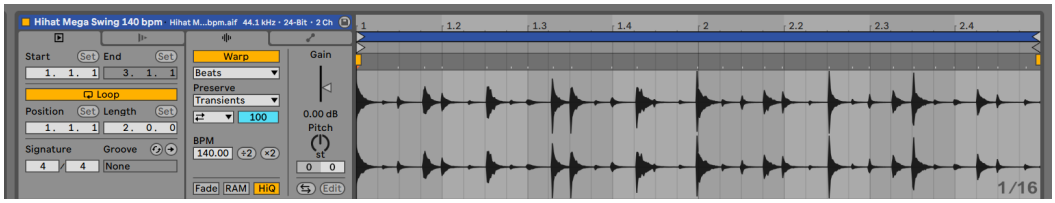


- Shaker 87 bpm

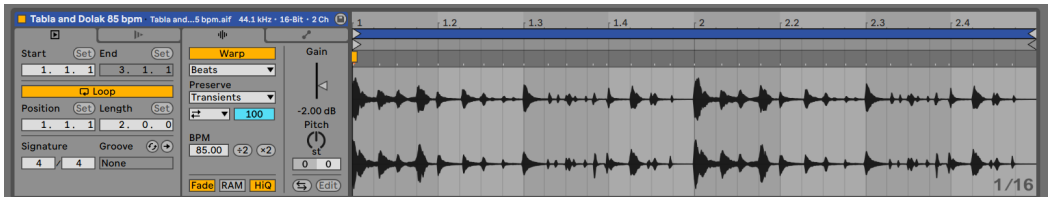


- Garage Room 135 bpm

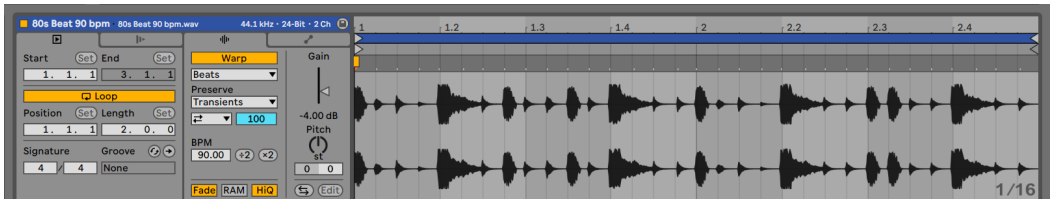
- Hihat Mega Swing 140 bpm



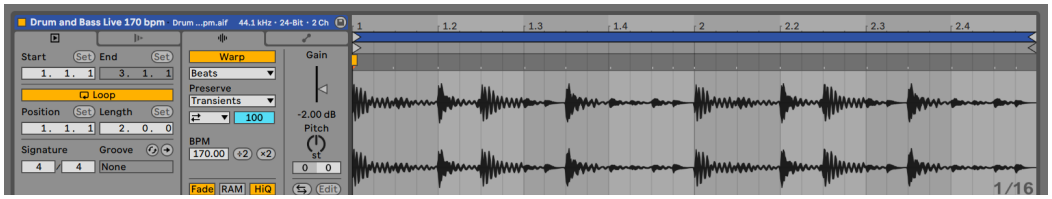
- Tabla and Dolak 85 bpm



- 80s Drum Machine Wingman 120 bpm

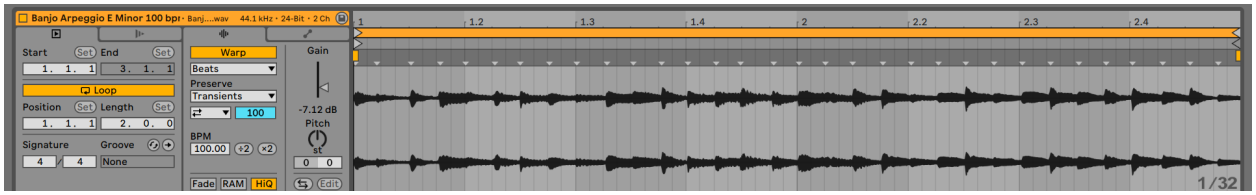


- Drum and Bass Down 170 bpm

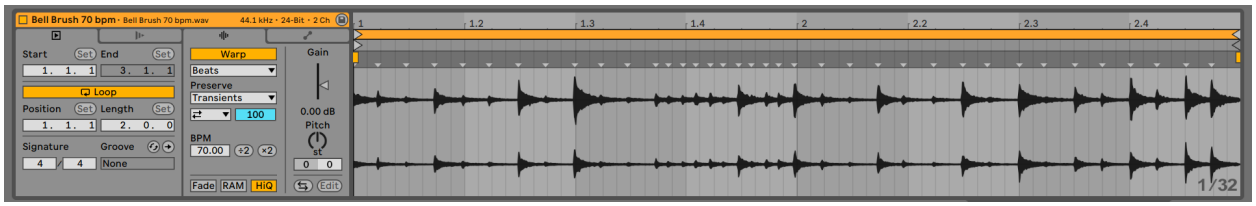


Layers:

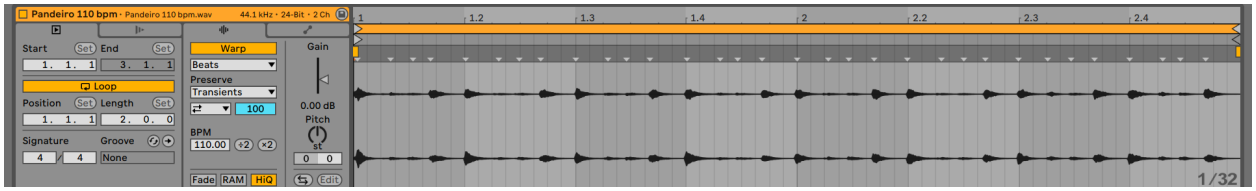
- Banjo Arpeggio E Minor 100 bpm



- Bell Brush 70 bpm



- Pandeiro 110 bpm

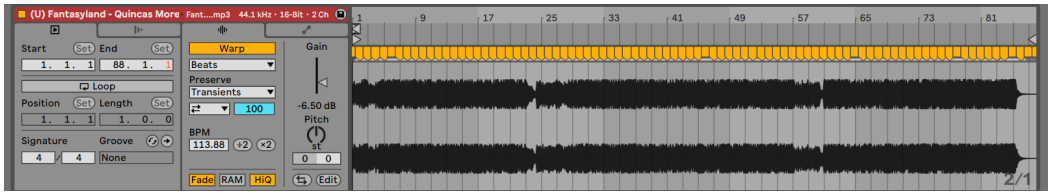


Melody:

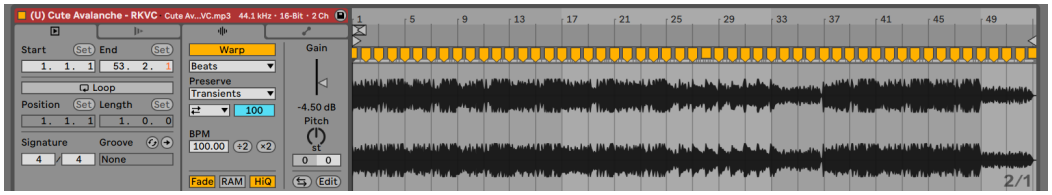
- (S[24]) Xica Xica

- (U) Try To Catch Me - Reed Mathis

- (U) Fantasyland - Quincas Moreira

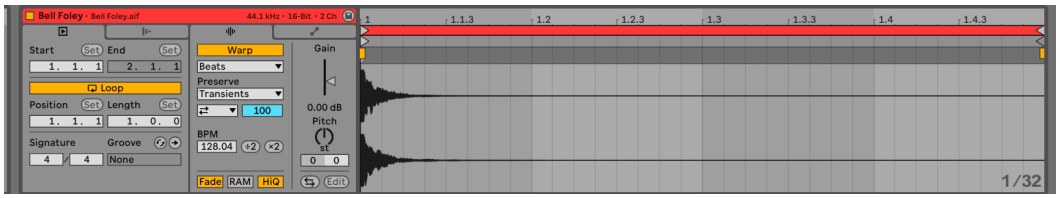


- (U) Cure Avalanche - RKVC

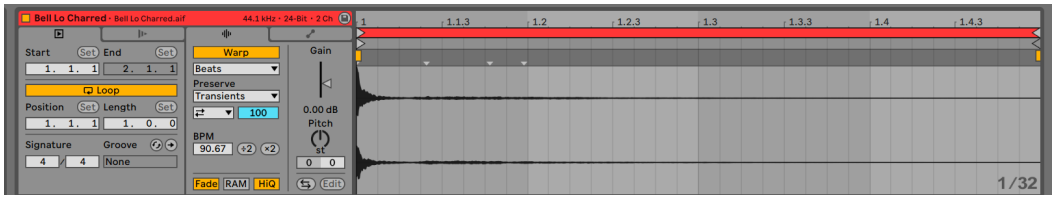


Unexpected:

- Bell Foley

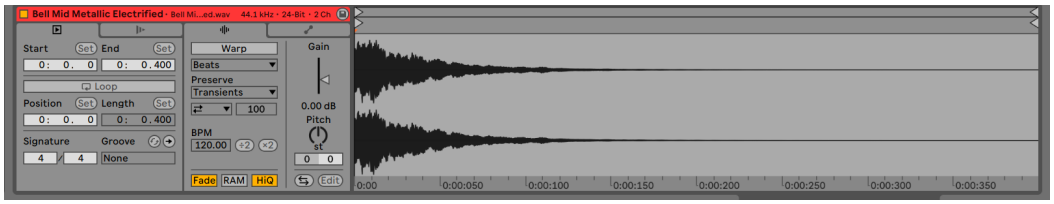


- Bell Lo Charred

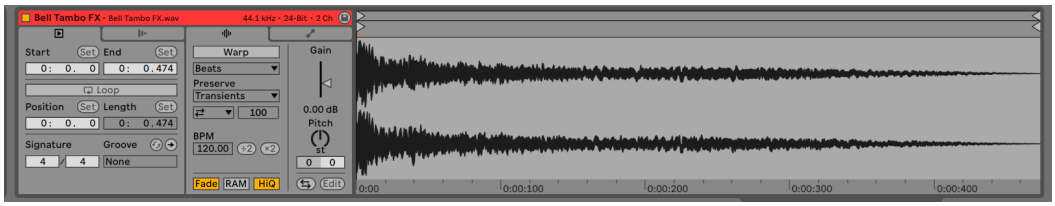


[24] 'U' is short for 'Upbeat' and 'S' is short for 'Slow'.

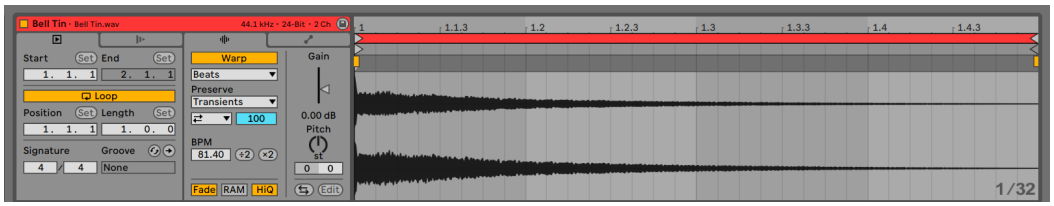
- Bell Mid Metallic Electrified



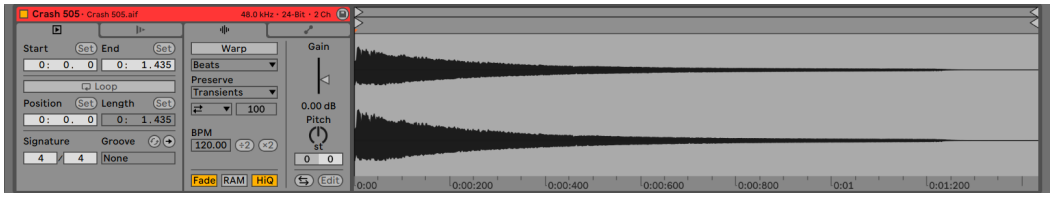
- Bell Tambo FX



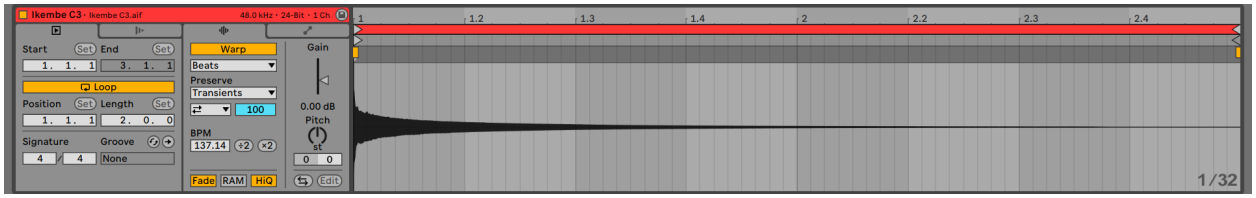
- Bell Tin



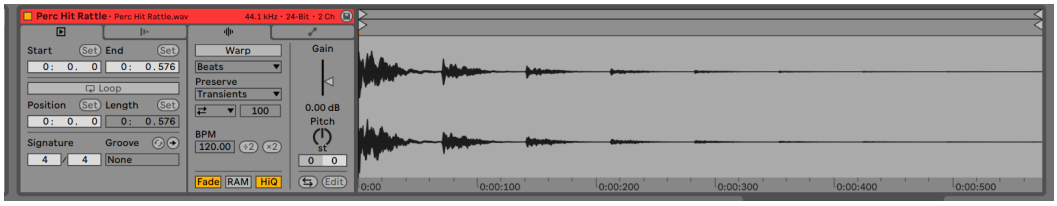
- Crash 505



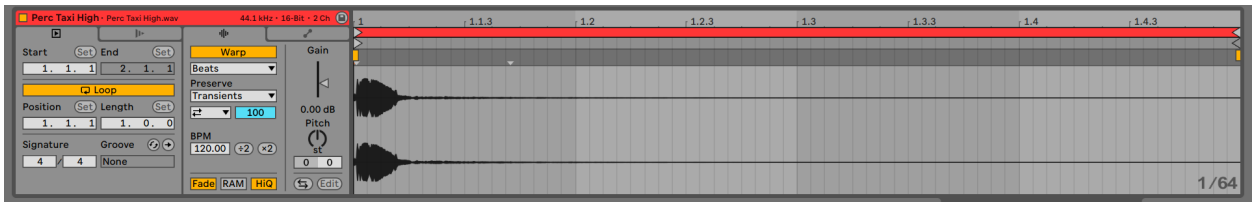
- Ikembe C3



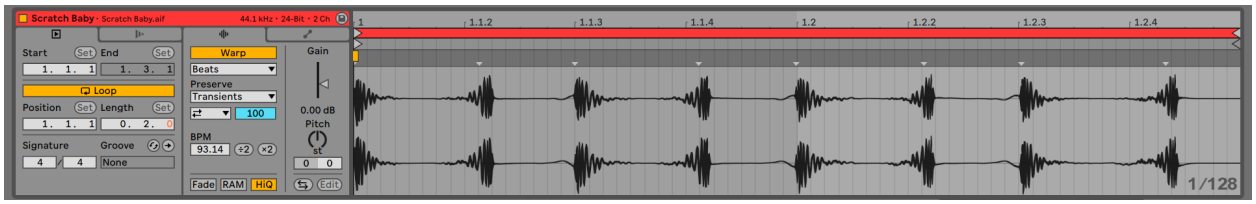
- Perc Hit Rattle



- Perc Taxi High

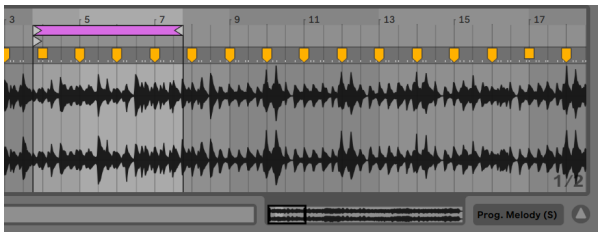


- Scratch Baby

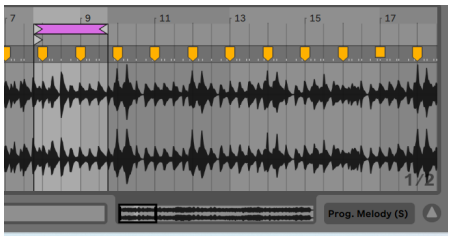


Prog. Melody (S):

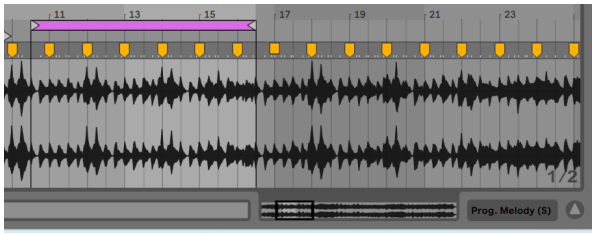
- (1) Twirly Tops - The Green Orbs



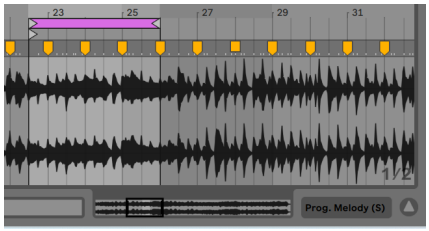
- (2) Twirly Tops - The Green Orbs



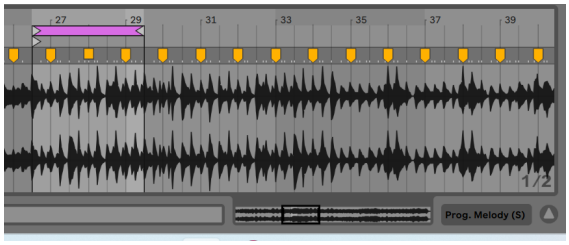
- (3) Twirly Tops - The Green Orbs



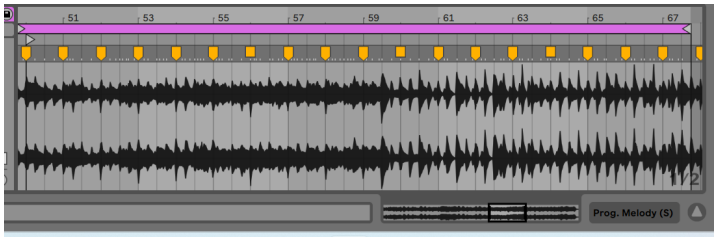
- (4) Twirly Tops - The Green Orbs



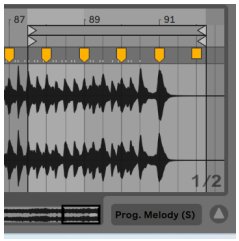
- (5) Twirly Tops - The Green Orbs



- (6) Twirly Tops - The Green Orbs

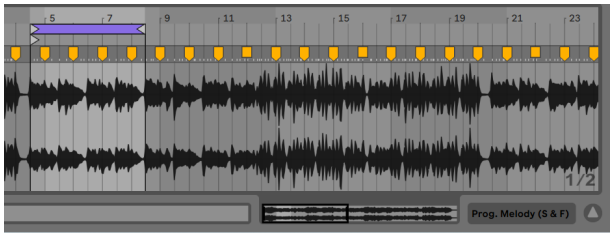


- (7) Twirly Tops - The Green Orbs

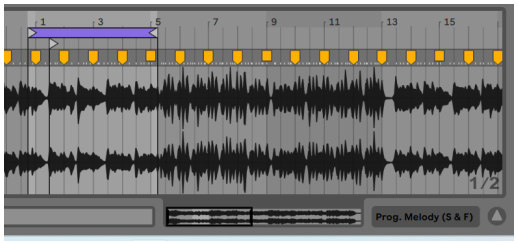


Prog. Melody (S & U):

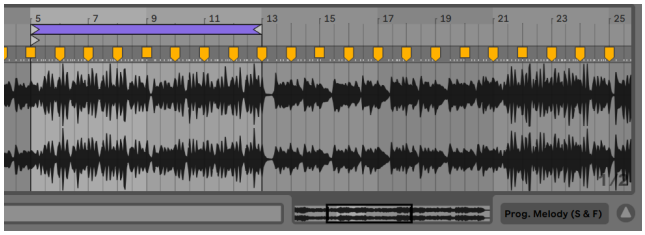
- (1) My Dog Is Happy - Reed Mathis



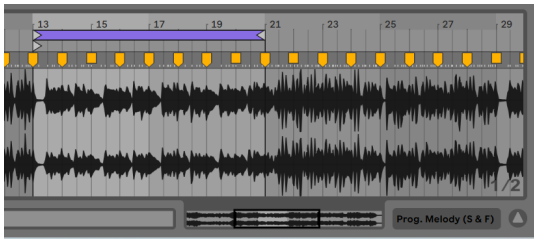
- (2) My Dog Is Happy - Reed Mathis



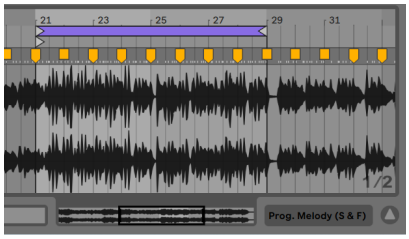
- (3) (U) My Dog Is Happy - Reed Mathis



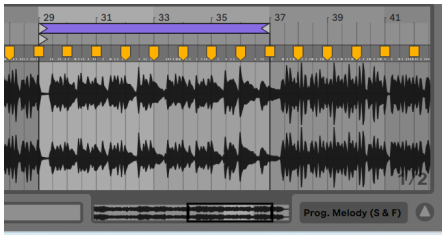
- (4) My Dog Is Happy - Reed Mathis



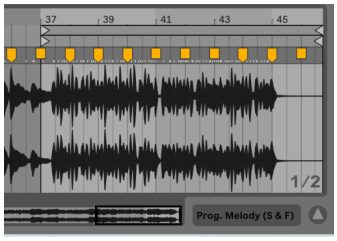
- (5) (U) My Dog Is Happy - Reed Mathis



- (6) My Dog Is Happy - Reed Mathis

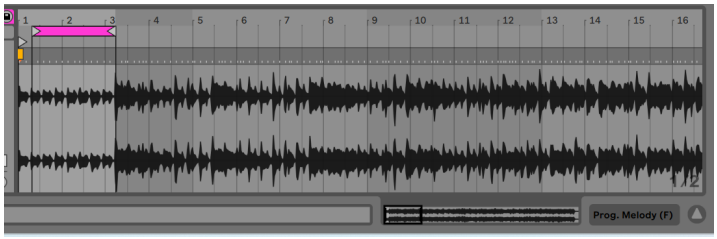


- (7) (U) My Dog Is Happy - Reed Mathis

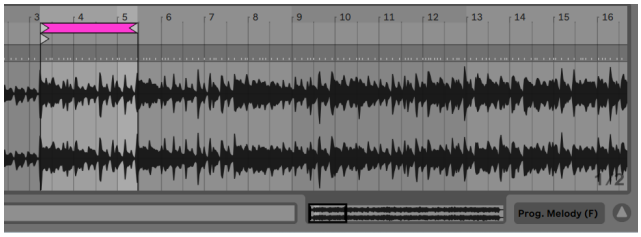


Prog. Melody (U):

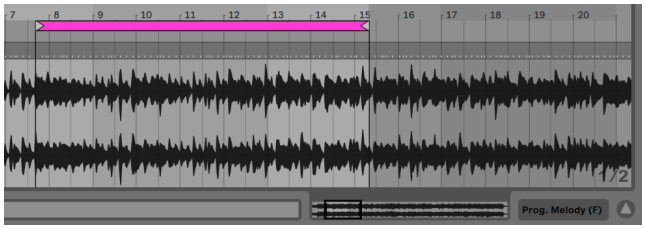
- (1) [(Put On Your) Dancing Pants - Reed Mathis]



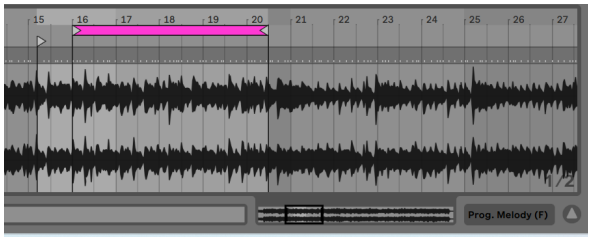
- (2) (Put On Your) Dancing Pants - Reed Mathis



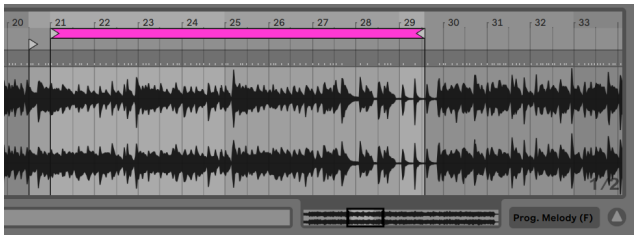
- (3) (Put On Your) Dancing Pants - Reed Mathis



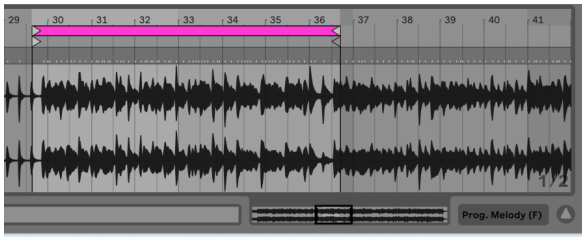
- (4) (Put On Your) Dancing Pants - Reed Mathis



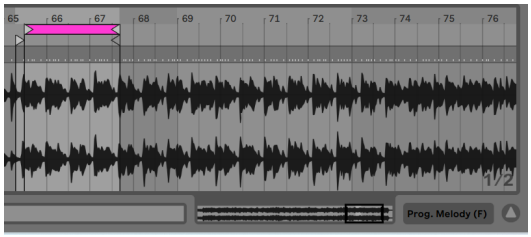
- (5) (Put On Your) Dancing Pants - Reed Mathis



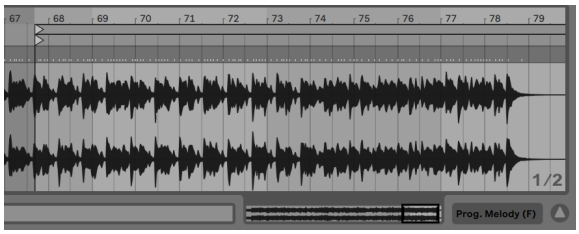
- (6) (Put On Your) Dancing Pants - Reed Mathis



- (7) (Put On Your) Dancing Pants - Reed Mathis



- (8) (Put On Your) Dancing Pants - Reed Mathis



All Combined:

[Ableton Live project file]

Appendix L: Information Letter & Consent Form *(Second Iteration)*

<date>

Information Letter (EN)

Dear teachers and parents,

Thank you, again, for your interest in partaking in our study. In this letter, we would like to inform you about what your students'/children's participation in our workshop sessions will entail, for you to have a clearer understanding.

Our Aim

Our goal is for your students/children to have playful interactions with one another though the different games they get to play. Besides having fun, they would be helping us, as a part of our final research projects, to design a technological system that can improve collaboration and enjoyment in neurodiverse settings in the near future.

What?/Workshop Contents

Specifically, the research we will carry out during these workshops is about childrens' experiences of different sensory inputs (e.g. sounds and touch). We would like to create products that best suit their needs in this sense and therefore want to get valuable insights through these sessions.

The participants will each be in max 3 workshops. In total each of these workshops will take a maximum 90 minutes long, between now and summer 2023. For each workshop, there will be sufficient breaks in between. The participant can always choose to join only one or several of the three workshops. In the workshops, the children will work in pairs. The sensory inputs mentioned will be included in objects (like a bracelet) the participants can hold or wear and the participants. The participants do not need to touch things they insist on not touching and can leave at any time. The sounds that the participants are objected to will be in normal hearing range and will be discarded/changed in case of any discomfort.

Informed Consent & Withdrawal Possibility

As the participants will be minors, the formal consent will be given by the parents and/or the school (representatives)/teachers, alongside their voluntary participation, for them to participate in our study.

In case of a wish to withdraw from the study, the participants will need to inform the workshop facilitators (Lara or Purna). If one does not want to participate anymore, this can be done during the sessions at any point. If the participant wishes to withdraw from the study after it has been conducted, the participants' guardian can email Purna Bishas, stating this situation.

The Nature of Research & Potential Risks

,

The research is not medical, it is not about a medical therapy or medical procedure of any kind.

The sessions are completely safe without any risks for the participants personal/mental health, as the research project has been reviewed by the Ethics Committee Information and Computer Science to ensure this.

Data Privacy & Protection

The data will be recorded/stored for research purposes on a laptop, password-protected and accessible only by the researchers themselves. This data will be saved in a safe research drive after the research has been conducted for at most 10 years. The identities of the participants will be protected by the use of participant numbers in any analysis conducted or any reporting. Any visuals that will be used in our reports will ensure anonymity (e.g. by blurring faces). All participants are voluntarily participating, ensuring that the participants' guardians are in full knowledge of de research, privacy of data and data storage.

Study contact details for further information:

Purna Bishas, s.a.bishas@student.utwente.nl

Contact Information for Questions about Your Rights as a Research Participant

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee Information & Computer Science:
ethicscommittee-CIS@utwente.nl

Consent Form for *neurodiverse sensory preception*

YOU WILL BE GIVEN A COPY OF THIS INFORMED CONSENT FORM

Please tick the appropriate boxes	**Yes**	**No**
Taking part in the study		
I have read and understood the study information dated [DD/MM/YYYY], or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	☐	☐
I consent voluntarily for my child to be a participant in this study and understand that they can refuse to answer questions and can withdraw from the study at any time, without having to give a reason.	☐	☐
I consent for my child to be audio-recorded	☐	☐
I consent for my child to be video-recorded	☐	☐
I consent that the recorded audio and video files may be used in the report, with the visuals blurred to ensure the participants' anonymity.	☐	☐
I agree that my children's information can be quoted anonymously in research outputs	☐	☐
Use of the information in the study		
I understand that the information I provide will be used for graduation reports and possible future publications.	☐	☐
I understand that any personal information collected about my child that can identify them, such as [e.g. my name or where I live], will not be shared beyond the study team.	☐	☐
Future use and reuse of the information by others		
I give permission for the anonymised transcripts, audio and video *recording* that I provide to be archived in a secure file on a computer, so it can be used for future research and learning. Afterthe study has been finalized, the files will be destroyed from the database.	☐	☐

Signatures

_____	_____	_____
Name of legal representative	Signature	Date

_____
Name of participant

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

_____	_____	_____
Purna Bishas	Signature	Date
_____	_____	_____
Lara Oral	Signature	Date

Study contact details for further information:

Purna Bishas, s.a.bishas@student.utwente.nl Lara Oral, l.oral@student.utwente.nl

Contact Information for Questions about Your Rights as a Research Participant

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee Information & Computer Science: ethicscommittee-CIS@utwente.nl

___/___/_____

Informatie brief (NL)

Geachte leraren/verzorgers en ouders,

Nogmaals bedankt voor uw interesse en deelname aan onze studie. Met deze brief hopen wij u beter te informeren over wat de workshop inhoudt en waar uw kinderen aan deelnemen.

Ons doel

Onze intentie is om uw kinderen te laten spelen met elkaar op speelse, bewegelijke manieren. Naast het hebben van plezier dragen ze mee in ons afstuderen, dat gaat over het ontwerpen van technologische systemen. Deze toekomstige ontwerpen hebben als doel om samenwerking te versterken en plezier te hebben voor neurodiverse kinderen in sociale situaties.

Wat houdt de workshop in?

Het onderzoek zal worden gedaan tijdens de speelsessies, wij kijken naar ervaring die de kinderen hebben wanneer zij met de verschillende zintuigen stimulerende objecten gebruiken. (zien, voelen en horen) Wij willen graag ontwerpen wat kinderen nodig hebben met betrekking tot de zintuiglijke waarnemingen. Door deze workshops krijgen wij waardevolle inzichten om voor te ontwerpen voor de zintuiglijke stimulans.

De deelnemers zullen maximaal in 3 workshops deelnemen in de periode vanaf nu tot aan de zomer 2023. De workshops bestaan uit verschillende sessies die ongeveer 10-15 minuten duren. In totaal zal dat maximaal 45 minuten zijn per workshop. Er zijn voldoende pauzes tijdens de workshop en die kinderen zijn vrij om zich te verplaatsen en verwijderen van de spelomgeving. In de workshops zullen de kinderen deelnemen in duo's of trio's. De bewegingssensoren bevinden zich in objecten die de deelnemers kunnen vasthouden, dragen (zoals een armband) of waarnemen(zoals bij schermen).

De deelnemers zijn niet verplicht tot bepaalde handelingen en kunnen zich zonder reden afmelden voor de studie. De objecten hoeven de kinderen niet aan te raken indien niet gewenst. De geluiden die de deelnemers zullen horen worden ook in normaal gehoorbereik en veranderingen zullen worden gemaakt in het geval van ongemak.

Formele toestemming & annuleringsverzoek

Gezien de deelnemende leerling minderjarig is dient het toestemmingsformulier te worden getekend door een ouder en/of de school leraren, dit met de toestemming van de deelnemende leerling om vrijwillig mee te doen met de studie.
In het geval dat de leerling niet meer mee wil doen aan de studie en daarmee zijn/haar deelname wil annuleren kan dat door een van de organisatoren te informeren (Yifan, Lara of

Purna). Dit kan tijdens en na de studie. Als de deelnemer zich alsnog wil terugtrekken nadat de studie is voltooid kan (de ouder/verzorger) dit doen door een mail te sturen naar Purna.

De aard van het onderzoek en mogelijke risico's

Dit onderzoek is in geen enkel opzicht een medisch onderzoek. Deze studie heeft ook niets te maken met medische therapie of andere medische procedures.

De sessies zijn volledig veilig zonder enige risico's voor de deelnemer zijn mentale of fysieke gezondheid. Deze studie is goedgekeurd door de ethische commissie van informatie en computer science van de Universiteit Twente.

Data Privacy & veiligheid

De data wordt bewaard op een laptop voor studie doeleinden. Deze laptop is voorzien van een wachtwoord en is alleen toegankelijk door de onderzoekers. De data zal na de studie worden bewaard op een beveiligde research drive voor uiterlijk 10 jaar. De identiteiten van de deelnemers zullen worden beschermd door genummerd te worden in de verdere analyses of rapportages voor de studie. Beelden in de rapportages zullen anonimiteit waarborgen (door gezichten vervagen). Alle deelnemers doen vrijwillig mee aan de studie en zijn zich bewust van de studie in de details en de privacy van de data en data opslag.

Studie contact details en voor meer informatie:

Purna Bishas, s.a.bishas@student.utwente.nl

Informatie voor contact voor vragen over uw rechten als onderzoeksdeelnemer

Als u vragen heeft over uw rechten als onderzoeksdeelnemer of informatie wilt verkrijgen, vragen wilt stellen of zorgen wilt bespreken over deze studie met iemand anders dan de onderzoeker(s), neem dan contact op met de secretaris van de Ethische Commissie Informatie & Computer Science:
ethicscommittee-CIS@utwente.nl

Toestemmingsformulier *neurodiverse zintuiglijke waarneming*

U KRIJGT EEN KOPIE VAN DIT TOESTEMMINGSFORMULIER

Kruis aan wat van toepassing is	**Ja**	**Nee**
Deelname in deze studie		
Ik heb de informatie gelezen en begrepen. Deze brief kan ook aan mij zijn voorgelezen. Ik heb de mogelijkheid gekregen om vragen te stellen over de studie en de antwoorden voldoen tot voldoening.	☐	☐
Ik geef toestemming dat mijn kind vrijwillig meedoet aan de studie en begrijpt dat hij/zij de opdrachten en/of instructies mag weigeren, ook mag hij/zij zich volledig terugtrekken uit de studie zonder enige reden.	☐	☐
Ik geef toestemming voor mijn kind om opgenomen te worden (audio)	☐	☐
Ik geef toestemming voor mijn kind om opgenomen te worden (video)	☐	☐
Ik geef toestemming dat de audio en video bestanden gebruikt mogen worden in de rapportage, deze beelden zullen visueel onherkenbaar zijn voor de privacy van de deelnemer.	☐	☐
Ik geef toestemming dat mijn kind zijn/haar information anoniem gequote mag worden in de rapportages	☐	☐
Gebruik van de informatie in de studie		
Ik begrijp dat de informatie die is verstrekt gebruikt wordt voor afstudeer rapportage en mogelijke toekomstige publicaties.	☐	☐
Ik begrijp dat alle persoonlijke data verzamel van mijn kind wat hem kan identificeren (bijv. naam en leeftijd), zal binnen dit team blijven en niet verder worden gedeeld.	☐	☐
Toekomstig gebruik van de research data voor anderen		
Ik geef toestemming voor geanonimiseerde transcripties, audio en video bestanden die ik beschikbaar stel, te worden gearchiveerd in een beveiligde map op een computer. Zodat deze kunnen worden gebruikt voor toekomstige studies en onderwijs. Na de studie zullen de bestanden worden vernietigd van de database.	☐	☐

Handtekeningen

_____	_____	_____
Naam van de wettelijke vertegenwoordiger	Handtekening	Datum

_____
Naam van de deelnemer

Ik heb zo accuraat mogelijk de informatie voorgelezen van de informatiebrief voor de potentiële deelnemer, naar mijn beste vermogen, de deelnemer geïnformeerd tot waarin hij/zij vrijwillig toestemming voor geeft.

_____	_____	_____
Purna Bishas	Handtekening	Datum
_____	_____	_____
Lara Oral	Handtekening	Datum
_____	_____	_____
Yifan Cheng	Handtekening	Datum

Study contact details en voor verdere informatie:

Purna Bishas, s.a.bishas@student.utwente.nl Lara Oral, l.oral@student.utwente.nl

Informatie voor contact voor vragen over uw rechten als onderzoeksdeelnemer

Als u vragen heeft over uw rechten als onderzoeksdeelnemer of informatie wilt verkrijgen, vragen wilt stellen of zorgen wilt bespreken over deze studie met iemand anders dan de onderzoeker(s), neem dan contact op met de secretaris van de Ethische Commissie Informatie & Computer Science: ethicscommittee-CIS@utwente.nl

Appendix M: Teacher Interview Script *(Second Iteration)*

R[25]: What did they think of it, do you think?

T: Uhh, at first they are 'Oh! Okay.', that you are in here. It's difficult for him (**P3**), because it's an open game and he was not doing here, what he is always doing here. It was difficult, but he was enjoying it, so– Well, he loves to have contact, but because it was open, at the end, he touched me and he hit me <laughs>, and uhh, it was difficult for him. He needs more structure in that sense.

R: And how could we do that? Like, umm, just to–

T: No, I have to do that I think. Because– Yeah, or you can do it also, but, music is an open thing, so…

R: Does he like music normally?

T: Hmm, I don't know, I can ask the group. Because he was so busy with (...), so he did not really listen to the music.

R: It was also that there were a lot of people and–

T: Yeah, this area is used as a gym and now he has to do different things.

R: Would this be better in a music room or something, you think? Or, like, just, what kind of setting/room would help them?

T: Yeah, the playroom, I think, would help him, but he is always busy with something else. (...) Because I take him to the playroom and, uh, in the past, I played with him with the bottle of fishes and [when there,] he always asks me for the bottle <laughs>. I cannot do nothing else <laughs>, because it's very difficult to get him on another mindset.

R: And he was also saying 'Oma!'(*'/Grandma!'*).

T: Yes, he compares me to his grandmother I think, always calls me that <laughs>. So, I don't react [to] that.

R: And he does that when he is overwhelmed or–

T: No, all the time, when he sees me. In a group, [also] when I am doing my own play activity with the suitcase…

[25] '**R**' is short for 'Researcher' and '**T**' is short for 'Teacher'.

R: And how do you understand if the kids are enjoying themselves or not? Like, could you just give me some examples from these kids? Some behaviors and…

T: Uhm, I saw a smile, so– And **P4**, she was very relaxed, also sitting next to you. She is also a girl who needs a lot more structure, but she loves music, she loves to dance, so I think she was very relaxed in this situation. Uhm, it was nice to see. And I follow the children, so it's not that they 'must' do something. And, yeah, I see that they move freely, make use of the space.

R: Yeah, and for you, was this a different way of connecting with them or how did you experience it?

T: Uhh, yeah it was connecting [to them] with my body and the music. And normally, I have toys also, to connect. Now it was only observation, to look, what they like and (...).

R: Do they do the same for you or not really, you think?

T: Uh, yeah, they follow my movements a little bit, sometimes. I saw that they– Yeah. Or they follow what they find nice, for example **P3** liked pushing. But I can, again, follow them. So it is a bit back-and-forth, actually.

R: And for you, how enjoyable was this?

T: <smiles/laughs> This was nice. But it was, for me also, an open thing, so I don't know if I do it good, so, uhh <laughs>.

R2: Some more structure would have been…

T: Yes, or, you told me of before [,speaking of the possible mirroring warm-up,] that I follow the hands or–

R: But I also didn't know how to instruct them on that.

T: No, it's difficult, so.

R: Cause I guess they don't have similar gym classes or something where they copy moves, do they?

T: No and if they have, they do it over and over again, so they learn repeating classes.

R2: The first kid (**P1**) was also… I thought he found it nice.

R: He liked it a lot, I think. <**T** also confirms.>

R2: He was very engaged but he also didn't really use his hands that much. So maybe it's just a personal thing.

T: Yes, it can be.

R2: He really enjoyed it I think. If I looked at him.

T: Yeah and they are not used to moving their bodies lose <laughs>.

R: Do you think if we would do this multiple times-

T: Yes! It would get easier.

Appendix N: Workshop Recordings' Descriptions *(Second Iteration)*

(Workshop 1)

[C1[26]]

- The two children (**P1** & **P2**) walk into the room, accompanied by their teacher (**T**).

- **P2** shows some initial surprise and briefly stops.

- Then, they both look around and together start walking to the front, towards the main researcher (*myself*).

[C2]

- **P1**, specifically, walks a bit to the side, to check out what is on the bench.

- He also takes the initiative to come closer to the researcher. He leans in and tries to see what is behind the laptop screen.

[C3]

- **P2** is also curious and follows **P1** in doing the same. He holds his gaze for longer, but he appears more hesitant.

- The researcher tells them that she will be playing music for them today.

[C4]

- **T** repeats this in a more child-directed and high-pitched tone, with a happy and excited facial expression (i.e. "Muziek maken!" (*/"Music-making!"*), coupled with orchestra-chef-like hand gestures, to showcase the meaning).

[C5]

- **P2** is, at first, heading towards the playground. However, when **T** calls out his name, he faces her and walks closer.

- In the meantime, **P1** is still looking around a bit, facing them.

[C6]

- **P2** comes over and takes another look at the set-up including a laptop.

[26] '**C**' is short for 'Clip'.

[C7]

- **P2** holds onto **P1**'s arm and they walk around together. **T** laughs and points this out ("Kijk!" (/"*Look!*")).

- *initial plan x change:* They seem to want to stick together, so a decision is made to include all three together in the workshop.

[C8]

- **T** repeats what we will be doing a few more times, again with animated gestures with her hands, or while shaking her arms to imitate dancing.

- Upon this, the children come together with **T** in the middle.

[C9]

- Music starts playing, which **T** highlights by pointing up and saying "Luisteren!"(/*"Listen!*").

- The children both stop and are listening, appearing focused.

[C10]

- **P1** leans towards **P2**, to which he responds to by holding onto **P1**'s arm.

[C11]

- When the music starts developing, **P2** pulls onto **P1'**s arm. **P1** moves towards the other direction, walking backwards/sideways for them to turn together clockwise.He does so, while looking at **T** and smiling.

- They start turning together in this way, gazing at each other at times. **T** is smiling and gives positive affirmation, saying "Goed zo <their names>!" (/"*Good job!*").

[C12]

- The music is getting louder as a response to this connection, which seems to prompt **P2** in increasing his speed of running around (so, **P1** on the inside and **P2** on the outside, clockwise running in a circle in place).

- They are both smiling and laughing, **T** is also watching them with a smile on her face.

- They seem to experiment/change a bit the speed at which they run.

- **P1** looks like he might want to stop turning, because he tries to slow down/stop a few times. **P2**, however, keeps hold of him.

[C13]

- With **P2** holding him on the arm, **P1** starts to move around differently, stopping with the turning interaction.

- Perhaps as **T** also noticed this, she raises and moves her arms and says “Arms up!”, which **P1** briefly follows, with a small gesture of raising his hand up (that is not held by **P2**).

[C14]

- She tries to get their attention(, repeating this a few times,) and get them to join as **P1** walks the other way with his back towards **T**, but **P2** is still facing her, despite holding onto **P1**.

- She also uses both their names, which seem to work in getting **P1** to pay attention.

[C15]

- In the meantime, **P2** finally copies **T** by moving his arms up and down.

- Perhaps also because he feels the movement through their physical connection, **P1** also joins in.

[C16]

- **P1** then leans in to look at the laptop screen again, prompting **P2** to do the same as they walk by.

[C17]

- **P2** actually finally lets go of **P1**’s arm to get closer and look at the screen for longer.

- Also when brief high-pitched random sounds are played, **P1** turns towards the system, despite walking away.

- In the meantime, **P1** starts dancing to the music himself, jumping around. He is going around and behind **T**, getting closer to her, which she notices.

[C18]

- **T** starts walking to the back (with her back turned to **P1**) while still looking back at him, as if to get him to follow her as she walks to the rhythm.

- This movement/body language also seem to get **P2**’s attention, who is once again facing them, instead of the researcher/laptop.

- **P1** and **T** starts walking around side-by-side with exaggerated stomping movements as **P2** also goes running towards them, moving in a similar way.

- They seem to enjoy this exaggerated movement, smiling, perhaps also the sound it makes on the floor.

[C19]

- **P2** holds onto **P1**'s arm again as they step in synchrony and to the rhythm, now with **T** behind them and following in the same way.

- **P1** looks back as if to check her involvement.

[C20]

- The children then together walk towards the playground, leaving **T** on the side and start to climb the slide one after the other.

- **T** is still facing them and bouncing to the rhythm, trying to get their attention using exaggerated facial expressions and hand/arm movements.

- She also dances to the melody, moving side-to-side to the flute notes.

- Melody volume is turned down to reflect this disengagement from the dance interaction and high-pitched short sudden sounds are repeated while doing so.

[C21]

- This seems to prompt **P1** in going back to the middle area.

- As a different, only-rhythm sound starts to play, he runs towards **T**, who is in turn leaning down to get closer to his level.

[C22]

- **T** repeats their former movement of stomping, this time to even clearer beats, as **P2** is also joining/moving closer to them.

- The two children seem to be looking at each other.

- In response to T, **P2** jumps to the side to the beat. He is also looking down at the floor, possibly as he feels vibrations of their movements through the floor.

[C23]

- **P1** is smiling and copies **P2'**s movement by jumping towards him, to the front.

- They are all running, jumping around, laughing to the music, during which **P2** also makes brief physical contact with **P1** and appears to be gently pushing him.

[C24]

- **P2** then goes back to the slide to climb up, during which **P1** and **T** move towards each other.

- They both briefly look at **P2**, but then let him be and face and focus on one another.

- As they alternate between moving closer and apart as they jump around, **P1** briefly looks at the direction of the music/speakers when another layer (of bells (?)) is added, as to reflect the progression in connection.

[C25]

- He then goes right back, moving closer. He starts to laugh as he and **T** move together.

[C26]

- Changing from the ongoing jumping move, he starts walking/running around as he laughs (while **T** is also smiling).

[C27]

- **T** seems to anti-mirror him in this, copying the movement style of walking/running, but doing so backwards to get away from him as he is going towards her.

[C28]

- As his laughs are getting louder and he seems to be enjoying this interaction, **T** also 'pretend'-moves to the left and right a few times.

- As they fall out of the dance interaction and **P1** also heads towards the playground, the ongoing music volume is lowered.

[C29]

- To better assess rhythm or melody preference of the interactors, now one of the (upbeat) monotonous melody options is played, to which **P1** immediately responds to, by heading to the middle again.

- **P1** is laughing and running towards **T**, to which she responds by also leaning in/down again and also running towards **P1**.

[C30]

- They seem to play around with their movement directions and doing the same or the opposite responses.

- In the meantime, **P2** also joins back in and copies **P1** by running towards **T**.

- They all continue this interaction in different pairings (e.g. **P1** and **T** or **P2** and **T**).

[C31]

- The whole time, **P1** is vocalizing/laughing loudly and even hitting his upper leg, as if to show his excitement/joy.

- He also flaps his arms as he runs in circles.

- **P2**, in the meantime, is back at the playground.

[C32]

- When another layer with rhythms are added to mirror this progression in interaction, **P1** comes over again to look behind the screen.

[C33]

- After watching for a bit, he moves back and closer to **T** again, who is modeling a different dance of having her arms raised and moving them (like an orchestra chef), also leaning her head side-to-side.

- **P1** seems to be bouncing to the rhythm as he walks and then joins by raising his arms, except he flaps them up and down instead.

- In turn, **T** changes her movement to copy **P1**.

- She also offers variations that build onto this, whether moving her arms in circular motions or one up and one down.

[C34]

- **P1** has brief moments of distraction, first towards the music/speakers again, then to the other prototype on the side.

[C35]

- With the rhythm change, he seems to get prompted back into the interaction and they continue flapping their arms together.

[C36]

- Matching the melody's energy, **P1** increases the speed at which he moves his arms.

- Importantly, as they move together in this way, they are not always looking at each other, but **P1** is also with his back against **T**.

[C37]

- **P1** seems to enjoy these hand flaps for himself for some time (despite **T** copying him in the back and trying to connect) and comes over and looks at the laptop screen again.

[C38]

- When **T** calls out **P2** by name, **P1** also attends to her again.

- Importantly, up until this point, **P2** was left to what he wanted to be doing, which was to sit/slide down the slide.

[C39]

- **T** makes a slightly different arm/hand gesture to the beat, which **P1** mirrors as he gets closer to her. They go back-and-forth with some variations, which makes **P1** laugh.

- They then also repeat the alternation they had of the movement directions/getting closer or apart.

[C40]

- Now, with a better sense of their musical preferences, one of the progressing melodies are played.

- **P1** first looks in the direction of the music and then towards **T**, following this change.

- They gaze at each other as they adjust to and figure out how to dance to this new song (which is also lower energy).

- **T** initiates a new figure, alternating leaning to the sides to match the melody, to which **P1** responds to by laughing, moving closer and mirroring her.

[C41]

- **P1** also seems to get closer to and vocalize towards **P2**, who is at the playground.

[C42]

- **P1** makes a brief gesture, stretching one of his arms down and perhaps prompted by this, **T** reaches out her arms/hands, for him to take.

[C43]

- However, as melody has moved to its next section (because of the ongoing connection), **P1** first takes a look in the music direction again, then makes a turning movement.

- **T** joins and copies him in this.

[C44]

- For a moment, **P1** looks as though he would stop turning, but he notices **T** doing the same, so he continues.

- Perhaps noticing **T'**s arms, he also raises his, which T again copies. They continue turning together and in different directions.

[C45]

- **P1** repeats looking at the system whenever there is a change and vocalizing/hitting himself on his upper legs as an expression of some emotion.

- This time, it may be that he is feeling done with the interaction(, which the researcher and **T** do not see as such at the time), because he moves his arm side-to-side, as if to say no and gets closer to and looks at the laptop screen.

[C46]

- In the meantime, **T** has explicitly went up to **P2,** calling him by name and inviting him, so they now hold hands.

- She gives her a positive affirmation (“Goed zo!” (/”Good job!”) and also calls **P2** by name, so she has **P1** on one hand and **P2** on the other.

- She prompts them to also hold hands, so they are now in a circle.

[C47]

- She starts moving and then moving them all clockwise. The children increase their speed and are laughing.

[C48]

- **T** tries to move in a way different then just running in a circle, perhaps stomping to the random dog bark sound, but **P1** goes back into the circular motion.

- Here, **P2** seems to follow **T** more.

[C49]

- They (interestingly, seems to be initiated by **P2**) release hands and **T** attempts to raise her arms again.

- **P1** vocalizes and also moves his arms, hits himself on his upper legs.

- **P2** goes back to the playground.

[C50]

- Now another progressive melody is played, so **P1** checks out the screen again and then faces **T**.

- She starts moving to the new song and as **P1** moves closer, he joins in by mirroring her movements.

- When he seems to stop and stare down, she tries to get involved by moving closer.

- At some point, he introduces his own movements (drawing form before), turning with raised hands, in which **T** follows him. They are both also smiling.

- Based on this progressing connection, the next (upbeat) section of the song is moved to. However, this seems to fit the turning motion less, so **P1** slows down and stops, whereas **T** goes to try and get **P2** involved again.

- When unexpected (‘bong’) sounds are played, **P1** smiles and seems more excited as he also walks towards **T** and **P2.**

- **P1** holds **T’**s arms, as if to recreate the all-together turning interaction, as they are all in physical contact.

- **T** tries to move in a way that matches the playing sound, but **P1** wants to do the same turning motion, so they do.

- At some point, **T** lets go of their hands and **P1**, wanting to continue, holds **P2** in the arm instead to keep running in circles.

- **P2** seems to go with it, but also calmer/lower energy.

- Upon the change into the next upbeat part of the song, **P1** lets go of **P2’**s hand and does the same turn with his arms raised on his own again.

- This time, **T** does not mirror/follow him, but tries to do other moves she sees fitting to the music, including clapping.

- **P2** goes back to the playground.

- When the connection seems monotonous, **T** does mirror **P1**, also when he moves to the front with his back against her. He looks back and is surprised to see her follow , smiling.

- He looks at the screen again and says <something> with a hands going down gesture, to indicate that he wants the music to come to an end.

- The researcher and **T** confirms this by repeating/asking if it is enough, so he repeats himself.

- The researcher then comes to the last clip and slowly reduces the volume to come to an end.

(Workshop 2)

[C1]

- The two children (**P3** & **P4**) walk into the room, accompanied by their teacher (**T**).
.
- They immediately start looking around and exploring the contents of the room, appearing to be curious.

[C2]

- **T** says ‘Oh, dit is gezellig!’ *(‘Oh, this is cosy!’)*. This is interpreted as her setting the tone for how the children are to relate to the environment.

[C3]

- **P3**’s attention is drawn towards the other prototype in the room, next to which another researcher is sitting.

- **P4**, on the other hand, walks towards the main researcher *(myself)* and the set-up involving a laptop.

[C4]

- **P3**, referring or talking to the main researcher, says ‘Barman!’ *(‘Bartender!’*)

[C5]

- **P4** heads to play at the playground.

[C6]

- **P3** comes running towards the main researcher, to see behind the laptop screen.

- In the meantime, **T** introduces the children to the researchers by their names.

[C7]

- **T** says ‘I prepared them for dancing, so maybe we can first (...)’.

[C8]

- **T** repeats to the children (**P3** & **P4**) that they will be dancing, in an inviting way.

- Matching this invitation from **T**, the music starts to play, starting from a lower volume.

[C9]

- **T**’s invite and the music starting to play appears to prompt **P4** to come over, into the ‘dance area’ – so the middle of the room where **T** is also standing.

- To ‘mirror’ this interest of hers in the musical interaction, volume is increased.

[C10]

- **P4** also comes over and checks behind the laptop screen, to see what's going on, as the volume increase continues.

- In the meantime, there is another verbal prompt from **T** towards **P3**.

[C11]

- Crucially, **T** also 'shows' the children what to do, by starting to move her arms and to dance. This appears to prompt **P3** to come towards her and to mirror her movement.

- As he approaches, **P3** is also singing, imitating the melody of the music.

- **P4** also heads over towards **P3** and **T**. It appears to be the case that the two children generally influence each other in what they do in the space and where their attention goes.

- Then, however, **P4** goes towards the other prototype again.

[C12]

- **P3** touches and pushes **T**, who afterwards continues the same movement they were both doing previously.

- **P3** mirrors her in this movement once again.

[C13]

- **P3** initiates physical contact again, by holding onto **T**'s arm.

[C14]

- **P3** says '<**T**'s name>! Oma!' *(/Grandma!)*.

- **P4** heads to the playground to play.

[C15]

- **T** makes use of 'showing' once again, raising her hands. She has them almost in a double-high-five position and invites **P3** to put his palms onto hers.

- **P3** follows this lead and while holding hands, they shortly move their arms together to the sides.

[C16]

- **P3** then follows **P4** in going onto the playground instead.

[C17]

- **T** observes the children and gently tries to initiate interaction with **P4** in the same way.

[C18]
- **P4** mirrors **T** in terms of also moving her arms – so the same body part as **T** just did.

[C19]
- In return, **T** mirrors her in her specific movement or way of swinging her arms. Afterwards, she also repeats some of her own movements again.

[C20]
- **P4**, instead of responding or interacting, goes through her and towards the playground.

- She moves in a way where she is distancing herself and moving away, with a body language that appears to be closed off.

[C21]
- Both children play in the playground and are distracted from the dance activity for some time, so the music volume is lowered to 'mirror' this disconnection.

- A musical switch is made to play rhythms, rather than melody, as well as some high-pitched irregular sounds.

[C22]
- The high-pitched irregular bell sound played appears to prompt **P4** to go back in the 'dance area'.

[C23]
- **T** – literally, through a gesture – points the music out to the children (**P3** & **P4**).

- She often uses their names when addressing them, assuming, also to get their attention.

[C24]
- Keeping this pointing gesture, **T** moves her arm up and down to show or mark the rhythm.

- She also follows **P4** as she walks towards the main researcher, lowering the distance in between.

[C25]
- **P4** appears to walk to the rhythm, with her toes pointed up and her weight mainly laying on her heels. In this way, she is also making a sound that matches the rhythm herself and is smiling.

- **T** mirrors her in this.

[C26]

- **P4** appears to notice **T** mirroring her, as she looks at **T** and holds her gaze or attention.

- Regardless, she is briefly distracted by the things on the side.

[C27]

- **P4** gets prompted to dance, sliding her foot back while keeping it on the floor, turning and jumping.

- She also holds onto or up her dress, smiles and vocalizes/sings.

- **T** affirms her participation, saying ‘Jaa!’(/*‘Yess!’*).

[C28]

- **T** mirrors **P4**’s turning movement, slightly changing it by having her arms out instead. She also turns in the opposite direction, to face her, smiling. This alteration appears to make her movement seem bigger and more attention-grabbing.

[C29]

- Music is progressed by changing to another rhythm, to ‘mirror’ the social progression.

- **P4** again gets distracted by the items on the side.

- **T** walks towards **P4** and extends her hand, saying ‘Kom!’ *(/’Come!’)*.

[C30]

- Despite this, **P4** moves away and has a closed body language again.

[C31]

- **P3** approaches the screen again.
- **P4** seems to be prompted to do the same: **T** continues her attempts in getting **P4** to dance together, but she (quite literally) turns and walks away.

[C32]

- **P3** mirrors **T**’s extended-arms/plane move and walks towards her.
- To imitate/’mirror’ this progression, a different rhythm is now playing.

[C33]

- **P3** interprets **T**’s dance figure (where she has her arms out, in a T-shape) as a plane flying.
- **T** repeats this verbal comment, perhaps as a validation.
- **T** also raises her hands (in a double high-five) again, to initiate physical contact.

- It sort of works, after which she takes more initiative to actually hold **P3**'s hand to turn him around (so his back faces her front), guiding his arms to do the same movement, while in physical contact.
- **P3** seems to let his balance onto **T** as they walk together to the front, heavily stepping/ stomping to the rhythm, as well as tilting towards left and right according to it.

[C34]

- **T** also tries putting **P3**'s own arms around himself, wrapping him up as if he is held in a hug.
- They together bounce to the rhythm while in this position.

[C35]

- **T** and **P3** continue to dance together to the rhythm, holding physical contact.

[C36]

- Reflecting this continuation and progression in social connection and interaction, melody is added on top of rhythm.
- This seems to prompt a change in movement and connection for the pair: **T** lets go.