EXPLORING TIMBRE TRANSFER AS A CREATIVE MUSIC PRODUCTION TOOL TO AID AUTISTIC STIMMING

By

Aarav Pillai

THESIS

submitted in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in Data Science and AI for the Creative Industries

Creative Computing Institute, University of the Arts London

2022

Supervisor: Prof. Mick Grierson

Word Count: 12,291

"...i have swarms of fire ants just under my skin
they are never not moving I am never not
aware of them most of the time it's just
something I've gotten used to until something
touches me the wrong way and
makes all of them bite my skin burns
spreading out in ripples from the
epicentre of me all of my cells are screaming
and the only way to quiet them is to
flap rock twist spin drag shoe along floor
rub hand across hair
notice the release of it
do it again."

- Gabe Moses, Spoken Word Poem titled 'Stimming' (2013).

Abstract

Restricted repetitive behaviours, or 'stimming' in particular, are routinely described as disruptive and problematic in autism literature, and are often the target of suppression or elimination in popular therapeutic modalities. However, the neurodiversity movement encourages people to look at stimming as a presentation of difference that has a productive and pleasurable purpose, instead of seeing it as a symptom of deficit - with a growing body of contemporary scientific research attesting to the same. A small but rising cohort of disability and musicology studies have also suggested that autistic people can hear and perceive music in ways that might significantly differ from neurotypical people, and that it could have tremendous benefits on autistic wellbeing. Therefore, in this thesis, the researcher examined whether a deep learning technique called timbre transfer could be utilized as a creative music production tool to aid certain auditory stims. Through in-depth interviews with six autistic adults, the researcher tried to gain detailed insights into their relationship with stimming, music and technology, and explored if timbre transfer could be a beneficial technique to support sensory dysregulation via user testing Google's publicly available timbre transfer tool. A thematic analysis was performed on the interview transcripts to explore patterns and find common pain points in the collected data. Using these results, the researcher then worked on two different approaches to timbre transfer, and tested both the models with the same six research participants, followed by a user experience evaluation using the system usability scale. Findings from this study indicated that participants stimmed for various reasons i.e., to self soothe or to communicate intense emotions, and found the concept of timbre transfer for stimming intriguing, but worth iterating further.

Acknowledgements

I would like to thank my supervisor, Prof. Mick Grierson, for all the guidance I've received from his end to help me materialise my aims for this thesis, and for taking extra time out to support me whenever I had any doubts. Furthermore, I am also thankful to the following people: My course leader, Dr. Louis McCallum, for teaching me the basics of programming and machine learning for audio; my mentor, Vanessa Maugey, for helping me manage my time efficiently; my friends, Nivedita Bansal, Sarah Cooper, and Tejie Xu for spreading the word about my study amongst their circles; and most importantly, my six research participants, for voluntarily offering their invaluable time to me for this study, and for being as enthusiastic as me about the interviews and the usability tests. This thesis is dedicated to all six of them.

Table of Contents

1. INTRODUCTION

For decades, autism has been defined as a lifelong developmental disability, typically characterized by varying degrees of hindrances in the areas of social, cognitive, and motor functioning (Faras, Al Ateeqi and Tidmarsh, 2010). The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013) outlines perceptive peculiarities and strong emotional reactions to sensory overload as few of the traits for a formal diagnosis (Chamak, Bonniau, Jaunay and Cohen, 2008). These sensory sensitivities potentially predisposes autistic individuals to higher levels of anxiety if restricted repetitive behaviour, or “stimming”, is discouraged (Kapp et al., 2019). However, to date, most research focuses on curbing autistic stims through psychological interventions due to them being viewed as a volitional act that is meaningless, embarrassing, or something that could be distracting for the people around the individual (McCormack, Wong and Campbell, 2022).

1.1. Funding Patterns for Autism Research:

While scientific research so far attests to some of the broad range of deficits brought on by autism as a neurological difference (Hill and Frith, 2003), little attention is paid to improving an individual’s wellbeing and personal development upon receiving such a diagnosis, let alone focusing on all the positive attributes that the individual might possess by virtue of being autistic. In the United States, $950 million was invested in autism research over five years via the ‘Combating Autism Act’ in 2006 to develop early screening programmes, prevention strategies, and diagnostic tools for children (Singh, Illes, Lazzeroni and Hallmayer, 2009). The US funding portfolio analysis (IACC, 2010) highlighted 37% of research funding being directed towards identifying risk factors for autism, 24% being focused on discovering treatments and early interventions for the same, 18% on understanding the underlying biology of the neurotype, and 13% on diagnostic tools. Only 6% of research funding was allocated to lifespan issues and wellbeing services for autism post diagnosis (IACC, 2010), which illustrates a massive research gap in this area.

1.2. Importance of Participatory Design in Autism Research:

Interpretive approaches to priority setting i.e., actively seeking out and including views of all stakeholders for efficient rationing and resource allocation (Sibbald, Singer, Upshur and Martin, 2009) is a popular practice in health-related research (e.g., Lloyd and White, 2011; Oliver, Mossialos and Robinson, 2004). However, this is not a common occurrence in the field of autism research as most funded projects fail to include autistic researchers to help design the motivations and goals of the study (Askham and Dattaro, 2022), or highlight lived experiences of the group in the research through participatory design methods (Fletcher-Watson et al., 2018), which could render the study impertinent if it's not making a positive difference to the lives of autistic people.

Several research projects that revolve around finding the genetic markers of autism have been criticized by autistic people (Chapman, 2021) due to the probability of pre-natal testing leading to a rise in abortions if autism is suspected (Schneider, 2022). Since the introduction of prenatal tests for Down's Syndrome in Iceland in the early 2000s, only about two children with the syndrome have been born each year, with the country's government celebrating that they've almost "eliminated" Down's Syndrome (Quinones and Lajka, 2017). This raises serious concerns of future eugenicist attempts to discard an entire subgroup of people that are neurologically different, instead of accepting and celebrating the strengths of neurodiverse people.

A study in 2014, focused on gathering views on priority setting for autism research in the United Kingdom, analysed primary data via in-depth interviews and focus groups with autistic people, and other stakeholders such as mental health practitioners and family members of autistic individuals (Pellicano, Dinsmore and

Charman, 2014). The study highlighted a clear disparity between the United Kingdom's focus on funding projects, and the needs of the stakeholders - with a consensus amongst the participants on prioritizing research that improves the lives of autistic people, instead of finding the cause or a cure for the neurotype.

1.3. Aim of this Study:

Since stimming usually serves as a coping strategy to manage emotions in response to an overwhelmingly sensory environment, this research study aimed to support that activity via a music production tool utilising a deep learning technique called timbre transfer. As autistic people are the primary stakeholders with regard to autism research, in-depth interviews and usability tests with autistic adults were conducted to help inform the design process of the tool. The researcher hypothesized that this music tool could be used to aid certain auditory stims such as echolalia (i.e., repetition of speech or sound), whilst also proving to be an excellent tool for creative music production.

This thesis also includes a thematic analysis of all the interviews, with the participants' lived experiences quoted using pseudonyms to safeguard their identity. Moreover, despite there being no correct or universally agreed upon usage of terms when referring to autism, this paper utilizes identify-first language i.e., "autistic people" or "autistic traits", instead of person-first language i.e., "people with autism" or "autism-related traits" as it is recommended by psychologists (Dunn and Andrews, 2015), and also preferred by a lot of autistic people (Kenny et al., 2015) since it is prone to less stigma (Gernsbacher, 2017).

2. LITERATURE REVIEW

2.1. Autism and Stimming:

Restricted repetitive behaviours (RRB) are a core diagnostic feature of autism (Cooper et al., 2021), and are usually described as strong preference for structured routines, repetitive motor movements, an enveloping preoccupation with interests, and a fascination with sensory aspects of an object or the individual's environment (American Psychiatric Association, 2013). The RRB criteria also includes having a hyper or hypo-reaction to various sensory stimuli (Charlton, Entecott, Belova and Nwaordu, 2021), and these traits are exhibited differently amongst people on the spectrum.

2.1.1. Types of Stimming:

Self-stimulatory behaviour, or "stimming" as reframed by autistic self-advocates (Orsini and Smith, 2010), is a particular type of RRB that can be more easily recognisable by others due to it being more explicit in nature (McCormack, Wong and Campbell, 2022) and is something that involves at least one sensory domain i.e., touch, sound, smell, movement, or vision (Davidson, 2010). Whilst even neurotypical people have tendencies to stim, the frequency and duration for which autistic people stim is much higher (Smith, 2022) and can sometimes be uncomfortable or painful to suppress (Crosman, 2022).

Some examples of stimming include; (1) motor stims such as hand-flapping, finger waving, running, jumping, or rocking back and forth; (2) auditory stims such as humming, screeching, covering and uncovering of ears, listening to the same song over and over, or various forms of echolalia i.e., repetition of speech or sound; (3) visual stims such as repeated blinking, fascination with spinning items and bright patterns, or intently gazing at objects for a long time; (4) tactile stims such as squeezing objects, playing with slime and playdough, rubbing hands on different types of surfaces, or touching a vibrating object; and (5) oral and olfactory stims such as licking, biting, grinding teeth, sniffing particular items such as food, essential oils, clean laundry, etc. (Robson, 2022; Johnson, 2022; Experia USA, 2022). The aforementioned list is not exhaustive, and autistic people stim in variously unique ways that often support them with emotional or cognitive self-regulation, along with helping them to focus better (Kapp et al., 2019).

2.1.2. Auditory Processing Disorder and Stimming:

Along with sensory sensitivities, autistic people also tend to have auditory processing disorder (APD) - a fairly common co-occurrence wherein sound and speech can be heard, but perceiving the same becomes difficult (Moore and Ferguson, 2014) i.e., they might have a hard time understanding verbal speech if there is background noise, leading tothem missing out on some words in a sentence being spoken to them, or verbal speech taking longer to process (Haigh et al., 2018). Sensitivity to background noise due to APD also means that autistic people can hyperfocus on local sound sources that others might miss out on, and some might stim to them too if there is a repetitive, predictable pattern to it (Baron-Cohen et al., 2009).

2.1.3. A Cognitive Defence of Stimming:

In 1998, Roy Baumeister conducted a series of historical experiments to prove his ego depletion theory (Baumeister, Bratslavsky, Muraven and Tice, 1998), an idea that he introduced in the field of psychology that hypothesized that the brain’s decision-making capabilities have limited power when they contradict our inclinations, and that we have a finite amount of energy to devote to cognitive tasks if we’re suppressing temptations (Hagger, Wood, Stiff and Chatzisarantis, 2010). His preliminary experiment design included asking two groups of participants to solve challenging puzzles, with one group being given some chocolate cookies prior to the task, and the other group being asked to resist the urge to eat the cookies and instead focus on completing the puzzle. He found that participants that were asked to withhold the urge to eat were quicker to give up on solving the puzzle, as opposed to the group that had the opportunity to indulge in the cookies before the task.

This experiment tested various skills related to executive function i.e., the brain's ability to focus, manage several tasks simultaneously, and initiate or inhibit actions (Kim, 2013), which are skills that autistic people usually struggle with (Demetriou, Lampit and Quintana, 2022). Baumeister's study concluded with the claim that it could be ineffective to expect optimum productivity when one is suppressing an urge or temptation to do something, which can be correlated to the social expectation from autistic people to suppress stims in public environments such as classrooms or offices where productivity is expected. Therefore, it could be hypothesized that the cost of inhibiting stims could not only emotionally drain out autistic people, but also make it harder for them to focus on tasks at hand as they are devoting most of their brain reserves on suppressing their stims.

2.1.4. Self-Injurious Stims:

It is important to acknowledge that not all kinds of stimming may be positive though. Certain stims, such as pinching, head banging, or ingestion of non-edible substances could be acutely harmful to the person (McCormack, Wong and Campbell, 2022), and the occurrence of such stims may depend on the individual or certain overwhelming situations that trigger meltdowns (Bennie, 2019).

2.1.5. Applied Behaviour Analysis:

Regardless of all the benefits that stimming has on autistic individuals, it is often viewed as something that needs to be fixed in autism research literature, due to the activity looking odd or disruptive from a neurotypical perspective. Stimming has been subjected to various controversial therapeutic interventions, one of the most popular ones being Applied Behaviour Analysis (Matson et al., 2012), which is a teaching method heavily reliant on operant conditioning. It makes use of both positive rewards for good behaviour (e.g., verbal praise, receiving tokens, etc.) and punishments for bad behaviour (e.g., physical restraint, losing tokens, etc.) to teach desirable skills to children (Lynch, 2019). Essentially similar to dog-training and with over 40 hours per week of one-on-one sessions (DeVita-Raeburn, 2022), the ethics of conducting such arduous and sometimes dehumanizing interventions on children is debatable, with "appropriate behaviour" being a very subjective notion - leaving the child with diminished critical thinking skills uncomfortably parroting socially acceptable norms.

Interventions that make use of punishments and rewards without outlining the purpose of the task have been proven to have detrimental effects on an individual's self-esteem, and to make sound judgements independently (Sandoval-Norton and Shkedy, 2019). Presentation of post-traumatic stress disorder symptoms in some autistic individuals who underwent ABA therapy has also been observed (Kupferstein, 2018). Given the rising amount of studies on the negative mental health impacts of masking in order to pass as neurotypical (e.g., Cage and Troxell-Whitman, 2019; Mandy, 2019), other strategies that help autistic people better manage sensory stressors, instead of forcing them to camouflage stims, needs to be researched.

2.2. Autism and Music:

Autism was first academically mentioned in 1943 by Leo Kanner in a scientific paper entitled "Autistic Disturbances of Affective Contact." (Chamak, Bonniau, Jaunay and

Cohen, 2008). Despite being a contentious paper in other respects, Kanner reported an interesting finding: six of the eleven autistic children that he studied had a strong aptitude and interest in music, with musical processing abilities being equivalent to typically developing children - despite having language processing deficiencies (Neumärker, 2003). Since then, several studies have shown that music can serve as an excellent therapeutic medium for social engagement and emotional growth for autistic people as there is no active need for verbal communication (Quintin, 2019) and due to multiple senses being involved at the same time, the listening experience can be very stimulating (Vines, Krumhansl, Wanderley and Levitin, 2006). Some caretakers have also reported that their non-verbal autistic children would have no problem singing and would prefer doing that instead of speaking (Molnar-Szakacs and Heaton, 2012), which probably is because verbal language and music abilities develop in separate regions of the brain (James et al., 2014). Moreover, due to its rhythmic structure, music also has an element of predictability to it which makes it ideal as a form of therapy for autistic people to self-soothe (Geretsegger, Elefant, Mössler and Gold, 2014).

2.2.1. Music Therapy - Active or Receptive Approach?:

The World Health Organization has authorized music therapy as one of the best approaches to support the wellbeing of autistic people (Fancourt and Finn, 2019), with some studies having concluded that music has immense potential to even encourage speech production in non-verbal autistic children (eg., Wan et al., 2011; Janzen and Thaut, 2018). Contemporary music therapy literature outlines two different approaches for the same, namely active and receptive (Alpers, 1941). The active approach to music therapy involves the client interacting with a musical instrument themselves, alternating between methodical and spontaneous patterns. In contrast, the receptive approach revolves around the client listening to a set of live or recorded music to either elicit a particular response, or to alleviate certain undesirable emotions (Hashemian, Mashoogh and Jarahi, 2015).

While both approaches are beneficial and functional, the former potentially serves as a better method for autistic people due to sustained music creation being proven to enhance executive function (Shen et al., 2019; Williamson, 2022), which is something most people on the spectrum typically struggle with (Demetriou, Lampit and Quintana, 2022). Moreover, due to autistic people usually having stronger pitch perception and pitch memory compared to neurotypical people (Stanutz, Wapnick and Burack, 2012), along with the ability to hyperfocus on local sound sources (Baron-Cohen et al., 2009), music creation also presents itself as an innately sensory and fun way to creatively communicate and express ideas in a world where they might feel socially excluded. Interacting with musical interfaces also increases activity levels in the thalamus (Mendability, 2022), a brain region responsible for voluntary motor functions, balance, and transmitting 98% of sensory information to the cerebral cortex (Robertson, 2022). Therefore, the active approach to music therapy, or music making in general, emerges as a potentially strong device to act as an aid not only for sensory and emotional regulation, but also for motor coordination for autistic people with dyspraxia - a common co-occurrence associated with the neurotype (Cassidy et al., 2016).

2.2.2. Accessibility and Music Technology:

Despite all the benefits that music making can have on autistic people, most acoustic instruments, such as the guitar or the piano, can be too complex to navigate for some people with cognitive or physical impairments (Cibrian, Peña, Ortega and Tentori, 2017), resulting in a restricted capacity for creative self-expression. However, with the advent of technology and its applications in music, it has now become possible to create and develop interesting and novel musical instruments with multisensory surfaces, making them very easy to be used by anyone, including autistic people with delays in cognitive or motor development (Kossyvaki and Curran, 2018).

Some studies also hypothesize that autistic people are naturally drawn towards technological devices, possibly due to its predictable functions and outputs, and the fact that there is no expectation to possess strong social or language skills to meaningfully engage with it (Kossyvaki and Curran, 2018). Therefore, collating and applying research on assistive technology with music could be very advantageous in understanding how to support autistic people.

2.3. Timbre Transfer and Neural Audio Synthesis - Background:

In music, timbre refers to the distinctive tone and quality of a musical instrument or human voice (Chase, 2022). Also known as tone colour or tone quality, it is one of the basic elements in music which helps people distinguish between two different instruments or voices, even if the same note is being played or sung. The flavour of different apples is a suitable analogy for timbre in music. Since green and red apples have their own individual flavour, despite both being apples essentially, people can tell them apart after tasting it. Likewise, the character of the sound produced by playing a particular note on a violin would be very different from that of a guitar, despite both being string instruments.

Interestingly, a deep learning technique called timbre transfer could enable a user to synthesize perfect digital reproductions of one instrument through any audio input on a standard laptop or personal computer e.g., using human speech as an input to make it sound like a piano, or using a synthesizer as an input to make it sound like a violin - essentially operating as a tool that reimagines a musical performance of one sound source as another through trained data (Adkins, 2022).

2.3.1. Parallels to Style Transfer Models:

The underlying principle behind most timbre transfer models is slightly similar to how a style transfer model works (TensorFlow, 2022), which utilizes an optimisation technique by capturing high level features such as the basic structure of the content image (e.g., a cat), along with the low level features such as the colour palette and brush strokes of the style reference image (e.g., Van Gogh's Starry Night). The model then blends both these features together and outputs an image (e.g., modernized portrait of a cat that resembles the artistic style of the said painter's magnum opus).

This could be applied to audio in a very similar way by training a model on an instrument's timbre. A model trained on low level features (i.e., usually the harmonic distribution) of a piano would emulate the timbre of the same, and one could pass on an audio recording of a trumpet as input audio to make the output sound like a piano instead, whilst keeping the high level features of the trumpet such as pitch or volume (Adkins, 2022).

2.3.2. Differentiable Digital Signal Processing (DDSP):

Presently, there are a wide range of neural audio synthesis techniques that creatively employ various machine learning architectures for timbre transfer - with one of the most novel approaches being the Differentiable Digital Signal Processing (DDSP) library by Google Magenta (Google, 2020) which integrates the expressivity of deep learning models with archetypal digital signal processing elements such as oscillators, filters, and synthesizers - reaching state-of-the-art results in various tasks such as dereverberation and timbre transfer. Most popular neural networks for audio synthesis, such as WaveNet (Oord et al., 2016) or GanSynth (Engel et al., 2019), often overfit details of a dataset and operate as black boxes, which makes the output very difficult to interpret (Shiledarbaxi, 2022). By tweaking the frequencies and responses of sinusoidal oscillators and linear filters, the DDSP library allows the user to synthesize or reconstruct the sound of an acoustic instrument such as a violin or flute with ease (Google, 2020).

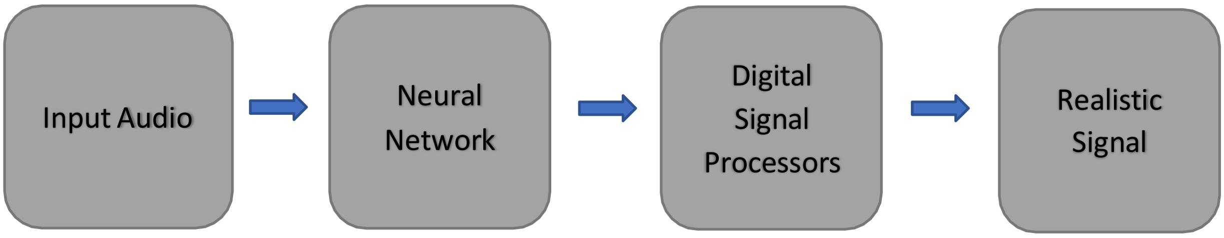


Figure 1: Basic DDSP Workflow

As illustrated in Figure 1, the DDSP library makes use of a neural network to transform a user's input audio into DSP controls, in order to produce more realistic signals. The beauty of this library is that one need not require massive amounts of data to train a model, with an audio file less than thirteen minutes in duration being more than sufficient to get interesting results after ten hours of training (Google, 2020). For timbre transfer in the context of violin as the target audio, a variational autoencoder with an encoder model and decoder model that's trained on violin audio is utilized. If an input audio of a child singing is sent across the model, the fundamental frequency (F0) and loudness features are picked up from it, with the former also being shifted up by two octaves to fit a violin's usual register (Shiledarbaxi, 2022), whilst the decoder is used for resynthesis. The end result is a resynthesized audio capturing diminutive details of the child's singing, with the timbre of the violin through the training data.

In 2020, Google released a public tool called Tone Transfer (Google Magenta, 2020) based on the DDSP library, which allows users to change the timbre of an input audio using intuitive controls - either by choosing offered audio samples by the program, or by uploading their own audio files. The user has the creative freedom of transforming the input audio to make it sound like any of the following instruments: Flute, Saxophone, Trumpet, and Violin. The tool also allows the user to finetune parameters such as mix, octave, and loudness via a Slider option (Basu, 2022). A music producer that tested the program online observed that it worked best on simple, monophonic sound samples, as opposed to polyphonic textures that resulted in really harsh noises (Huang, 2020). Regardless, this tool demonstrates how far deep learning has come in order for an AI to have the capacity to recreate the tone of musical instruments on its own. In 2022, Google followed up by releasing a VST Plugin for the same (Magenta DDSP, 2022), allowing users to produce creative music with this tool directly from their preferred digital audio workstation, along with training their own models.

2.3.3. Strided Convolution Models:

Symbol-to-Instrument Neural Generator, or SING (Défossez et al., 2018), is a type of strided convolution model that generates musical notes given a desired velocity, pitch, and instrument by the user, serving as a lightweight neural audio synthesizer. The model is trained to generate notes from almost 1000 instruments, and architecture comprises an LSTM based sequence generator, along with a single decoder. The output from the former component is plugged to the decoder which transforms it into a waveform that the user might desire.

Fast Spectrogram Inversion using Multi-head Convolutional Neural Networks (Arik, Jun and Diamos, 2019) is another method to generate waveforms via spectrograms, the output being speech in this case. The model is trained on a huge speech recognition dataset called LibriSpeech (Panayotov, Chen, Povey and Khudanpur, 2015), which contains 1000 hours of speech from audiobooks by 2484 speakers.

WaveGAN (Donahue et al., 2019) is another interesting approach to neural audio synthesis wherein GANs, that are popularly applied to image generation tasks, are utilized to generate raw-waveform audio. Based on the DCGAN architecture (Radford, Met and Chintala, 2022), this model can synthesize audio from various domains such as drums, piano, or bird vocalizations, along with intelligible words.

2.3.4. Autoregressive Models:

WaveNet (Oord et al., 2016) is a probabilistic raw audio generating model that is based on the Pixel Recurrent Neural Network (Oord, Kalchbrenner and Kavukcuoglu, 2016) architecture, and is said to yield excellent results as a use-case for text-to-speech conversion technology, as opposed to the traditional approaches to speech synthesis such as concatenative systems and parametric systems - the former being time consuming due to the need of voice actors, and the latter resulting in lower audio quality that sounds robotic (Sciforce, 2020). WaveNet is also capable of generating new and interesting music samples, but as previously discussed in section 2.3.2, it tends to overfit details of the given dataset.

2.3.5. Fourier Based Approaches:

Tacotron (Wang et al., 2017) is a Fourier-based generative model that is capable of producing speech directly from text at a significantly faster level than autoregressive approaches, and bypasses the need for spending long hours on building traditional text-to-speech components, which often requires expert domain knowledge.

3. RESEARCH QUESTION(S):

Based on the literature review, two research questions were formulated for this study:

RQ1: Can Timbre Transfer models be used as a music production tool to aid autistic sensory dysregulation?

The researcher decided to make use of in-depth interviews and usability tests to explore if stimming to background noises, as a form of echolalia due to auditory processing disorder, could be supported via music production through timbre transfer (as discussed in Section 2.1.2, 2.2, and 2.3). Finding inspiration in everyday sounds for music making is not an uncommon practice: David Gilmour from Pink Floyd composed the song 'Rattle That Lock' after being inspired by a French metro train jingle (Euronews, 2015), and Scriabin wrote his 10th Sonata inspired by insects (Judd, 2021). A lot of musicians also make use of field recordings in their musical practice. Given that autistic people typically tend to be more sensitive to background noises, designing a timbre transfer model for them to produce music using repetitive natural sounds could possibly yield some interesting results, and also help them to stim and self-soothe. A thematic analysis on their relationship with stimming, technology, and music was done to better understand motivations and pain points.

RQ2: Between the DDSP approach and the Fourier-based approach, which model is more effective to aid autistic stimming?

Building up on the findings gathered from the first research question, along with the background knowledge about different types of timbre transfer models from Section 2.3, the researcher decided to test out a DDSP-based model and a Fourier-based model to examine which approach worked better for the purpose of music-assisted stimming. Usability tests and survey methods for user experience evaluation were utilised to address this research question.

4. RQ1 – METHODOLOGY:

4.1 Overview:

The first part of the study employed in-depth semi-structured interviews or IDIs (Harvey, 2022) with autistic adults with a special interest in music, along with a usability test on an existing product by Google Magenta called Tone Transfer. As semi-structured interviews are a gold standard for qualitative interviews that require open-ended data on sensitive topics (DeJonckheere and Vaughn, 2019), the researcher chose this method to explore the first research question, along with a usability test on a timbre transfer model to collect insights on what the participants thought of the technology as a use case for stimming, whilst looking for any pain points or difficulties that they faced during the test.

The focus of these interviews and usability tests were to center lived experiences of autistic people in the design process by (i) exploring the participants' personal experiences with stimming, (ii) their perception of music through a phenomenological approach, (iii) their opinions on technology and AI in particular, which was followed by (iv) a usability test to better understand how they interacted with the design of Tone Transfer, and to gather opinions on timbre transfer as a technique for autistic stimming, or just music production in general.

4.2 Participants:

Following the approval from the university's ethics committee, participants were recruited through purposive sampling via various autism advocacy groups, as well as the researcher's personal networks. The study was introduced as exploratory research on whether music production could aid stimming. Gender and ethnic minorities were encouraged to apply to ensure that the participant pool was diverse. An infographic was designed to support the recruitment process for the study, along with a brief write-up about the same (see Appendix A).

Out of the 83 autistic adults that indicated interest, 6 were chosen ensuring that the participant pool was diverse of cultural and musical backgrounds. The participant's ages ranged from 22 and 37 years (median 25 years) and the genders were a balanced mix of male (n=2), female (n=2), and non-binary (n=2). The participant pool consisted of a mix of ethnic groups; namely White British (n=2), African American (n=1), Indian (n=1), Chinese (n=1), and Hispanic (n=1).

The participant pool also comprised of a balanced mix of formally diagnosed (n=3) and self-diagnosed (n=3) autistic adults. So far, research into autistic stimming has been majorly focused on participants with a formal autism diagnosis. However, autism may be under-diagnosed or simply overlooked in certain racial and gender-minority populations (Kim and Bottema-Beutel, 2019; Gould, 2017; Yuan, Li and Lu, 2021), which could be due to differences in presentation, cultural or class factors affecting social interaction differently, or difficulties with getting a referral for a diagnosis (Furfaro, 2022; Gould, 2017). Therefore, this study also included a balanced number of self-diagnosed autistic participants, along with formerly diagnosed autistic participants, to ensure that under-researched autistic populations were given fair representation. Lastly, to assess whether all the participants were autistic, a 14-item Ritvo Autism and Asperger Diagnostic Scale Test (Ritvo et al., 2010) was administered (see Appendix B), with all of them scoring above the suggested cut-off of 14 which is considered indicative of autism (Eriksson, 2013).

With regard to musical backgrounds, most participants were musicians utilizing either acoustic instruments such as bass guitars, cello, drum kit, kalimba, etc. or using technological devices such as autotune, TC electronic stereo looper, digital audio workstations, mobile apps, sound analysing softwares, etc. in their creative practice. One non-musician who used music to stim and had a special interest in the same was also selected. Participants did not receive any compensation for being part of the research, and were aware that it was voluntary. The table below illustrates all the basic information of each participant that was collected prior to the interviews:

Name	Age	Gender Identity	Ethnicity	Country	Relationship with Music	Autism Diagnosis	Chat Option
Olivia	23	Female	White British	United Kingdom	Studied Music + Choir Member	Formal Diagnosis	Video Chat
Nimbus	22	Non-Binary	Indian	India	Plays Bass as a Hobby	Self-Diagnosis	Video Chat
Jared	31	Male	Hispanic	Colombia	Studied Music + Composer	Formal Diagnosis	Video Chat
Thomas	22	Male	White British	United Kingdom	Produces Music + Plays Bass as a Hobby	Formal Diagnosis	Audio Chat
Wei	27	Non-Binary	Chinese	Singapore	Uses Music to Stim	Self-Diagnosis	Text Chat
Ruby	37	Female	African American	United States of America	Multi-Instrument-alist	Self-Diagnosis	Text Chat

Figure 2: Participant Data Collected Prior to Interviews

4.3 Interviews:

Selected participants were given a brief outline of the study, along with the participant information sheet (see Appendix C). Opportunity to ask any further clarifying questions was provided to them before they signed the consent form (see Appendix

D). If they were happy to continue with their interviews being recorded and transcribed, a mutually agreed upon date was finalized for the interviews - which was hosted on Microsoft Teams. An option for video chat, audio chat or text chat on the video conferencing interface was given to account for varying communication needs and preferences (Zolyomi et al., 2022).

An interview guide (see Appendix E) was also included in the study to keep the researcher grounded on the main areas to cover during the interviews, but the order and the nature of any additional questions were determined by the interviewees' responses. The interview questions were formulated based on existing autism research literature focused on stimming. A Distress Protocol (Draucker, Martsolf and Poole, 2009) for the participants was also included along with the interview guide, in case speaking about any particular life experience such as having to suppress stims brought up any traumatic memories (see Appendix F).

All the interviews lasted between 60 to 150 minutes (median 108 minutes). The participants were informed that they could request for short breaks whenever required, or stim anytime during the interview.

4.4 Usability Tests:

After the interviews, the participants were asked to share their screen and visit Google's Tone Transfer website for the usability test. This was done to get further insights into how the participants were interacting with the tool, and whether they thought timbre transfer could be used to aid stimming. The structure of the usability test consisted of a simple task in order to ensure that each participant meaningfully interacted with the basic features of the tool, which was then followed by some post-testing questions to understand the participants' thoughts on the language and design used, along with their motivations to use this as a supplementary tool for stimming.

The task was based on synthesizing and downloading a simple audio file using Tone Transfer, first by using one of the given audio samples by Google, and then by recording/uploading their own audio.

All the usability tests lasted between 15-30 minutes (median 17 minutes). The researcher then scheduled the next round of interviews with the participants with a break of 4 weeks to give the researcher enough time to prepare for them. The participants were also given an activity wherein they either had to record a naturally occurring sound they've been stimming to, or save an audio file of their favourite song / their own produced sample that they've been stimming to for the next usability test.

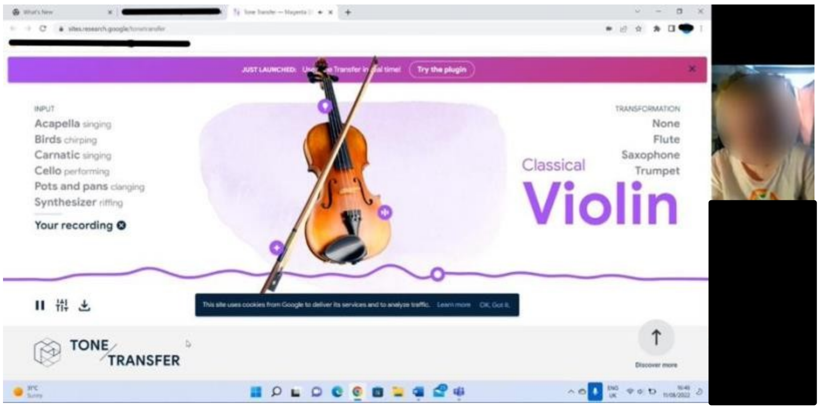


Figure 3: Testing Google Magenta's Tone Transfer with Olivia

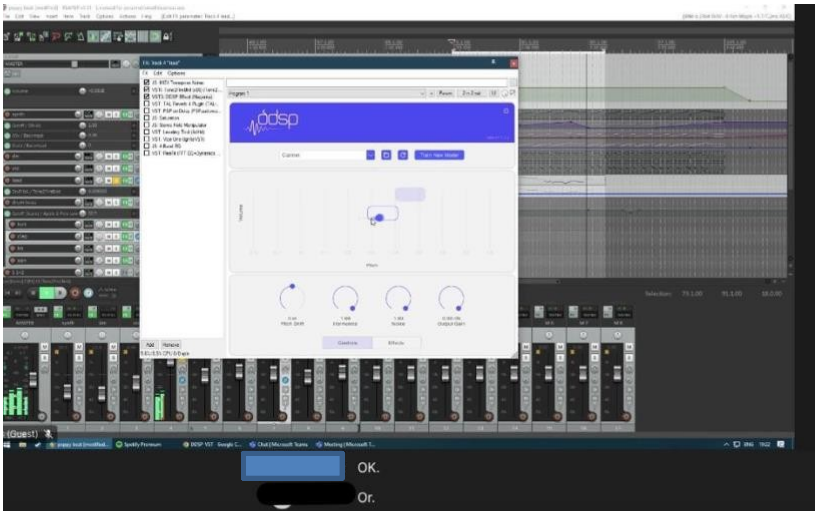


Figure 4: Testing Tone Transfer's DDSP-VST Plugin on Reaper with Thomas

4.5 Data Analysis:

After the interviews and usability tests were concluded, transcripts that were auto-generated from Microsoft Teams were lightly edited by the researcher to remove any unhelpful filler words, along with being proofread with the audio file playing to make sure all the words that were auto-transcribed were correct. The final transcripts were then emailed to the participants in order to confirm for accuracy. Upon receiving a confirmation, the researcher conducted thematic analysis on the transcripts as outlined by Braun and Clarke (2006), which consisted of six stages: (i) data familiarization, (ii) generation of initial codes, (iii) looking for themes, (iv) reviewing said themes, (v) refining and defining finalized themes, and (vi) contextualizing and representing the findings in a form of a report. A hybrid approach to analysis (Fereday and Muir-Cochrane, 2006), which combines both inductive and deductive coding, was employed in order to stay grounded on the research question, but also focus on the participants' lived experience at a semantic level without letting the researcher's personal knowledge about stimming influence the finalized themes.

Figure 5: Affinity Mapping

During the first stage, the researcher read all the transcripts repeatedly to properly immerse themselves in the data, and early impressions were written down on post-it notes.Then, affinity mapping (Pernice, 2018) was done in order to understand the relationship between all the initial impressions visually, before moving on to computer-aided thematic analysis using NVivo. The second stage involved generating initial codes using the aforementioned software for each transcript along with the pairings from affinity mapping, followed by the third stage of grouping all the codes into potential themes.

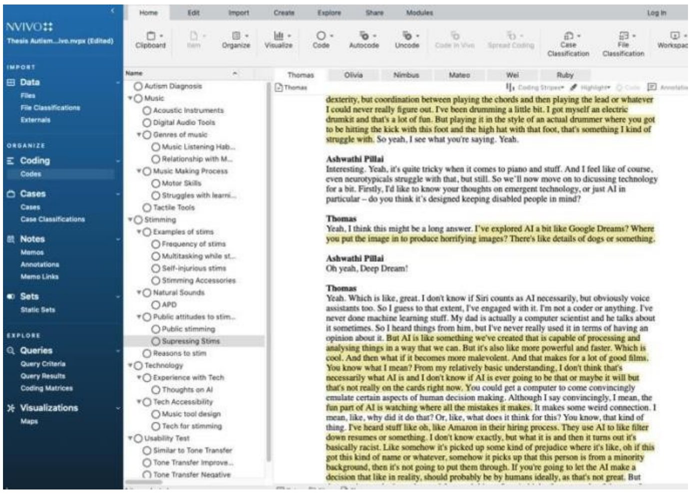


Figure 6: Looking for themes from all the generated codes using NVivo.

In the fourth stage, all the prospective themes were reviewed in order to verify if the sentiment of the responses were common amongst all the transcripts, or unique for particular participants, followed by refining and finalizing them in the fifth stage.

5. RQ1 – FINDINGS:

The data analysis led to the development of three main domains, which were further subdivided into themes. The researcher also tried sticking to exact quotes throughout the report to avoid interpretation bias.

5.1 Domain One: Stimming

Participants reflected and opened up about their personal experiences with stimming, and how it had an impact on their wellbeing. Common examples of stimming such as jumping, arm flapping, leg shaking, pacing, repeatedly listening to one song, teeth grinding, humming, grunting inscrutable words, and applying pressure on eye sockets were mentioned. Some uniquely individual ways of stimming were also discussed. Ruby said, *"Violin was one of the first instruments I learnt. If I don't have one around, I like imagining a violin in my left hand and giving a performance as a way of stimming."*, and Jared said, *"When I'm stressed at work, I sip two or three different beverages at the same time to keep myself stimulated, maybe coffee, lemon juice, and Dr. Pepper. The difference in tastes helps me stay calm somehow, not really sure why."*

All participants stated that they stimmed on a daily basis. The range calculated using reported estimations of time duration was between 1 hour to 9 hours (median 3.5 hours), although it is important to note that participants also mentioned that they often stimmed subconsciously as well, or when they were doing another activity or task as a way to stay focused or calm, so these estimated figures might not be completely accurate.

Through this domain, three main themes were identified: 1) Motivations to stim, 2) Reasons and consequences of suppressing stims, and 3) Auditory processing and stimming to natural sounds.

5.1.1. Theme One - Motivations to stim:

All participants described stimming as a calming and pleasurable activity, and that they would usually stim whenever they were in highly stressful situations such as overstimulating environments like crowded public transit, or brightly lit or very noisy places; new situations where they were not fully prepared causing anxiety; or while doing activities they didn’t enjoy such as laundry, looking at spreadsheets, or sleeping.

> Olivia said:
> *“[stimming] feels soothing and satisfying when I’m stressed, kind of like scratching an itch.”*

Some participants also mentioned that stimming helped them communicate intense and complicated emotions to loved ones when they were sensorily overloaded, as they struggled to speak or verbalize thoughts during that mental state.

5.1.2. Theme Two - Reasons and consequences of suppressing stims:

Negative public attitudes to stimming was commonly mentioned as one the reasons why the participants chose to suppress their stims, with three participants mentioning that they had to supress it because of their family possibly down to social stigma or relatively low understanding of autism. Wei said, *“I often suppressed stimming behaviour when I was younger because I was told by my parents that I shouldn't behave like that. It definitely felt horrible having to suppress it.”* When speaking about having to supress stims, Thomas said, *“I was recently having a conversation with a family member and she told me, like, ‘Stop flapping your arms. Put your arms down by your side and then say what you're going to say.’ I don't know why but I tried. And I was like, I can't. I can't do that. [...] Maybe she just doesn’t want to be associated with that. She doesn’t accept my diagnosis. She just doesn’t think it's a valid thing.”; and* Ruby said, *“I was told by [redacted] that my invisible violin stim ‘looks too methy.’. Especially when* *my wrist and fingers start contorting, doing more than just violin things. It feels sad, like I will never be completely safe to be my whole self. Not even in my own home sometimes. It hurts."*

The other three participants mentioned that they had a lot of support from family members while growing up, and that they were never told to suppress any stims or other autistic traits. However, they still reported that they would mask or camouflage stims when in public, with Olivia commenting, "Nobody's told me not to, but people have pointed out that I'm doing it. Which made me a bit self-conscious maybe. I do stim in public, but usually softly. Like, I'll tap on my phone or something instead of my preferred stims, so it doesn't necessarily look like an autistic person stimming.".

Effects of suppressing stims were reported as physically discomforting, with Thomasmentioning that "[...] it's like you really need to go for a wee or something. It's like you just can't not focus, like I need to do this", and Wei commenting, "I legitimately thought I was about to die because of how quiet it was and I wasn't allowed out of my seat.". Some participants also mentioned that they would start profusely sweating if they were suppressing or camouflaging their stims for too long.

5.1.3. Theme Three - Auditory processing and stimming to natural sounds:

Throughout the interviews, participants mentioned incidents of stimming to naturally occurring sounds in their environment. Some common examples amongst the participants were stimming to construction noises, traffic signal sounds, and the chirping of birds in the morning. Olivia mentioned: *"I work in childcare [...]sometimes I have to push a buggy and the sound of the wheels is really satisfying, especially if it's raining, that adds to the sound, and if you're on a dirty path because there's more things rubbing against the wheels. I guess because it's a cyclical sound that's quite satisfying and then sometimes something gets caught, like a leaf or a pebble, on one* *of the wheels. Every so often it goes round and you hear an extra sound which is very satisfying."*. When speaking about university course load and executive dysfunction, Jared commented, *"Do you remember the droning hum of the hypnotoad from Futurama? I was once watching an episode instead of working on my [music school] assignment. I couldn't get that sound out of my head so I added it to my project along with a chord progression I had written".* When specifically asked about stimming to natural sounds, Thomas said, *"If there's a noise like water coming off of a tap sort of thing, or construction noises with a particular rhythm, I'm like, OK, what? Whatever that noise over there, that's like the snare drum. And then I'm over here doing like a hi-hat when it comes in again. That's quite funny, you know. It's not that I necessarily hear it as that, it's more of how I conceptualise that rhythmically, I guess.";* and to the same question, Ruby answered, *"As a musical family it's normal to see us banging on the table, tapping on our glasses, turning random occurrences into songs [...] I have great difficulty with auditory processing when people are speaking. Need to watch TV with subtitles, [...] always asking my spouse to repeat himself. And multiple people talking at once is a recipe for instant panic attack for me. But with music or other sounds I seem to have the opposite problem. I can pick out the notes and voices that make up chords. I have reverse engineered songs I like, transcribing them and arranging for different instruments.*

5.2 Domain Two: Music

Given the diversity in cultural and musical backgrounds of the participants, questions around favourite genres of music and listening habits received very varied and interesting responses ranging from participants preferring popularly genres such as hip-hop, metal, and dance music, to more niche preferences such as vapourwave, Japanese folk, bluegrass, and PC music. Most participants also reported that their active music listening habits would go way beyond their neurotypical peers, responses falling under the range of 3 hours to 8 hours (median 3.5 hours). Olivia said, *"I always listen to music before choir rehearsals which is in London, and I try and remember to bring my noise cancelling headphones with me, so whenever I have a commute, travelling in London using public transport feels less stressful. I don't think I could get on the tube without headphones."*. Music seemed to have a connection with aiding emotional dysregulation too, with some participants reporting that listening to their favourite songs or making music would also help them break their repetitive negative thinking patterns. Although, Nimbus mentioned that shutdowns would make it harder for them to listen to music or help with low moods. *"I've had a depressive episode and during that time I didn't listen to the music for two and a half years, I just lost interest."*, they added. As discussed previously, the participant pool had five musicians and one non-musician. However, Wei, who described themselves as having limited musical experience, seemed to have explored quite a few instruments. *"I had to learn how to play the recorder for school, then I briefly dabbled with a Chinese instrument called erhu for a school club. I also learned piano but never made it past the beginner level. And I could play a few short ditties on the kalimba. I was learning to play the violin for a bit, but I didn't get to continue because my parent thought it would destroy my posture."*, they commented. Because of their limited engagement with all of the instruments, they said they didn't feel comfortable calling themselves a musician. With regard to music technology, all the five musicians used digital audio tools in their creative practise.

For this domain, the researcher decided to pick out one theme to get more insights for the research question.

5.2.1. Theme One: Struggles with using music tools

Difficulties with motor-coordination, information overload, not having enough creative control, and financial barriers with purchasing required software or instruments were some of the sub-themes identified, when asked about the difficulties participants faced while interacting with acoustic instruments or digital tools. Piano and drums were frequently mentioned as difficult to pick up quickly, with Olivia saying *"I think I found [playing the piano] more tricky because you have to have quite good hand eye coordination, and obviously you move your hands a lot. And I think that wasn't the most suitable instrument for me because of the dyspraxia. I think that could affect your learning rate a lot".* Nimbus commented, *"I struggle to understand most digital audio workstations. Logic and Garage Band has been straightforward, but the others seem to have too many things going on for me to grasp what's happening. I also feel like some tools do not make all the functionality of the software obvious for beginners, which makes it harder for me to grow creatively I think. I want to make new, interesting music and not just stick to the bass."*, and Ruby said, *"I'm never as skilled as I want to be. Sometimes it is difficult to focus on things I'm trying to practice on, especially if it's a complicated software. I slip into mindless stimming for hours at times, and then feel really guilty about it."*

5.3 Domain Three: Technology

To gauge the participants' experiences with technology, and openness to emerging technologies such as artificial intelligence for the study, the researcher marked this domain to be very prominent in understanding how to structure the second usability test. Most participants illustrated a fairly good understanding of technology, and had very strong opinions on the same. However, most of them also expressed that using new softwares would stress them out. Concerns about AI creating further social problems were commonly expressed. Jared replied, *"Depends on how AI's future is shaped. Like it's been proven that AI can be helpful, but there is also potential for harm. Did you hear about that AI Microsoft developed that became really racist really quickly? That's the talent of the Internet, but whatever. I don't like that."* Olivia and Thomas mentioned similar concerns about technology and corporations using it for wrong purposes, and their inability to do anything about it that made them upset. Olivia added, *"Sometimes it feels like technology is geared towards neurodivergent people in a kind of cynical way, even if it's probably not. Like with Tik Tok and just having 5 second videos and stuff, it's good for people with a short attention spans, which are usually neurodivergent people I guess. And like hearing the same audio over and over, it can be quite good if you're stimming. I've spent over 8 hours on Tik Tok some days, just wasting valuable time, and thinking about the fact that someone designed this to profit off people like me with obsessive tendencies makes me uncomfortable."* Thomas demonstrated a fairly strong understanding of a lot of technical tools, and had experimented with deep learning models for artwork before, but mentioned that having to look at code or math would take some time to process. *"I don't think anyone really understands FM synthesis apart from Brian Eno himself. I don't really do maths or code. So it's like, OK. That's very complicated. I would have to spend some time. But I'm very passionate about technology, and if given couple of days to focus on one thing, I could learn it in a detailed-manner."* he said. Ruby and Nimbus both expressed excitement about emerging technologies, and what it could mean for music and other creative applications. *"Data from Star Trek helped raise me. Whenever AI becomes sentient, I will gladly fight for their rights. In the meantime, I try to be polite to Siri."*, Ruby joked.

The researcher identified two themes relevant to the research question:

5.3.1. Theme One: Technology and Accessibility

The common consensus amongst all participants was that technology wasn't designed keeping disabled people in mind, regardless of their technical experience, with most stating that using technology regularly was stressful. However, some participants struggled to pin-point what exactly caused them anxiety or confusion about the same. *"I really don't like Microsoft products, especially Excel. I get quite anxious just using* *technology for day-to-day stuff like meetings and emails. I don't know if it's a specific autistic disadvantage or just lack of practice, but maybe it's because I'm autistic and I get overwhelmed quite easily", Olivia said.* One participant, Nimbus, mentioned that controllers with uneven buttons would overwhelm them, and that they couldn't figure out which buttons were important and which were not. "I just don't get why some tools have so many advanced features despite being advertised for beginners." they stated. Jared spoke in quite a lot of detail about how technology has the potential to support disabled people. *"Google Assistant helps blind people a lot, I imagine. I also use voice assistants to help me remember stuff, like I have an Amazon Echo in my room. I can easily turn off my computer, my TV, my lights, and music whenever I am feeling overwhelmed. I have a command for Echo that immediately turns off everything. And sometimes, if I forget to turn off the lights and I'm 100 meters away from home, it turns off the lights for me. So it's very useful for neurodivergent people too I think.*", he said.

5.3.2. Theme Two: Thoughts on Timbre Transfer:

The usability test on Google's Tone Transfer generated a lot of meaningful insights. After being told how the model works, most participants mentioned that timbre transfer could serve as a tool to support their auditory stims when they hear natural, repetitive sounds (as described in Section 5.1.2), along with being a fun tool to create off-beat, avant-garde music. Ruby said, *"Since I'm a multi-instrumentalist, the way I make music is that sometimes if I hear a certain instrument playing a certain melody, I'll just learn how to play that specific melody using that instrument. And it takes time and it can be really frustrating because you're having to learn something entirely new from the ground up, and then my imposter syndrome kicks in if I don't know how to perfectly play that particular instrument later. But this is cool because I can just sing into it, and it gives me at least an approximation of the instrument.".* Shealso found the design of the waveform *"cute and playful"* as it reminded her of her favourite storybook from childhood called Harold and the Purple Crayon. Olivia suggested that this tool could help someone that wants to make music, but struggles to grasp all the music theory concepts. Although, she also mentioned that she wouldn't use it herself as she found the generated sounds to be quite harsh for her liking, but really liked the concept and wished a mobile application like that could exist which generated "softer sounds like the synths from Logic." Participants also mentioned concerns about auditory distress from unpredictable inputs, and sounds not quite replicating the exact timbre of instruments as they had anticipated which led to disappointment.

The researcher made note of all the common pain points highlighted in this report, before working on both the timbre transfer models. However, the detailed insights of participants mentioning how they stimmed to natural sounds, along with their love and curiosity for music, helped the researcher conclude that the first research question might be true.

6. RQ2 – METHODOLOGY:

6.1. Usability Tests:

6.1.1. Model 1 - DDSP-based Approach:

For the first model, the researcher utilized a neural network building block called neural waveshaping synthesis (Hayes, 2021), or NEWT, which is a lightweight approach to neural audio synthesis. This work is part of Google's differentiable digital signal processing methods (Google, 2020) as the NEWT is paired with a differentiable harmonic-plus-noise synthesizer, which is able to produce realistic musical performances for timbre transfer applications. The researcher initially decided to train multiple autoencoders, released by Google Magenta (Engel et al., 2020), on random field recordings for abstract target instruments, but the audio generated from those models were very harsh and jarring, so she decided to make use of NEWT instead, as the sounds generated via this DDSP approach were a lot more milder and experimental, and gave a lot of creative control to the user.

The code accompanying the NEWT paper was edited and hosted on Google Colab for user testing due to its widespread use for running machine learning code. Given the participants' lack of prior experience with machine learning, the researcher decided to heavily comment out the notebook to make it more explainable, whilst also hiding all code so that the participants do not get overwhelmed with syntax they are not familiar with. Also, since autistic people typically learn new information faster when it is presented visually (Rudy, 2021), the researcher made use of two emojis throughout the notebook; 1) star emoji which indicated that the participant just had to run the cell, and need not worry about what it means too much, and 2) paintbrush emoji which indicated that the participant had to provide some creative input while running the cell, either through uploading an audio file of their choice, or adjusting the control signals before synthesizing the audio. The DDSP-based model gave participants a lot of flexibility when it came to sourcing the input audio too. They could either use a

YouTube link, upload a .wav file, or record directly from their microphone. The researcher added an optional section for music source separation via Facebook's Demucs (Facebook Research, 2021) in case they wanted to explore their current stim song in further detail through the timbre transfer model, or if they didn't have a .wav file handy.

All the participants were able to finish the activity in under 10 minutes.

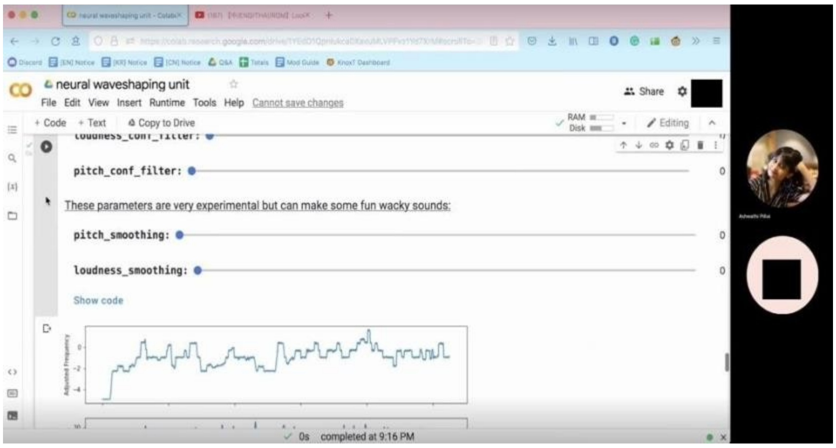


Figure 7: Tool testing the DDSP-based model with Wei

6.1.2. Model 2 - Fourier-based Approach:

For the second model, the researcher decided to make use of MIMIC Project (McCallum, 2022), short for "Musically Intelligent Machines Interacting Creatively", which is a collaborative website that allows users to explore various creative technologies such as machine learning for music and signal processing with ease. Due to the website's intuitive design, and having an existing Fast Fourier Transform (Grierson, 2022) example in the explore tab, the researcher chose this platform to host the second model. A Fast Fourier Transform is an algorithm that determines the Discrete Fourier Transformation (Max Documentation, 2022) of an input with greater speed. The document in the example made use of an Inverse Fourier Transform to convert a signal in the frequency domain to the time domain to be able to perform FFT. The code in the document also included NexusUI (Nexus, 2022), a basic interface library, and Maximillian (Grierson, 2022), one of the few audio synthesis and digital signal processing libraries that gives the user an Inverse Fourier Transform on the web. The interface had a play button and slider for the user to interact with the model. After forking the document, the researcher added some extra HTML code to the project in order to make the model more self-explanatory to use, as aspects like having to change the audio file name in the JavaScript code, or only uploading a total of 100 MB files, were to be kept in mind while using the model. A note to decrease the volume before pressing play was also added as per the findings from the previous usability test.

For the usability test, participants were asked to visit the MIMIC website and log in using the researcher's details, in order to save time on giving extra instructions about having to create an account and fork a project. They were then asked to open the Timbre Transfer project that the researcher had set up to be able to interact with it. After uploading audio files and changing the file names in the code, the participants used the slider and the play button to mix the timbres.

All the participants were able to finish the activity in under 10 minutes.

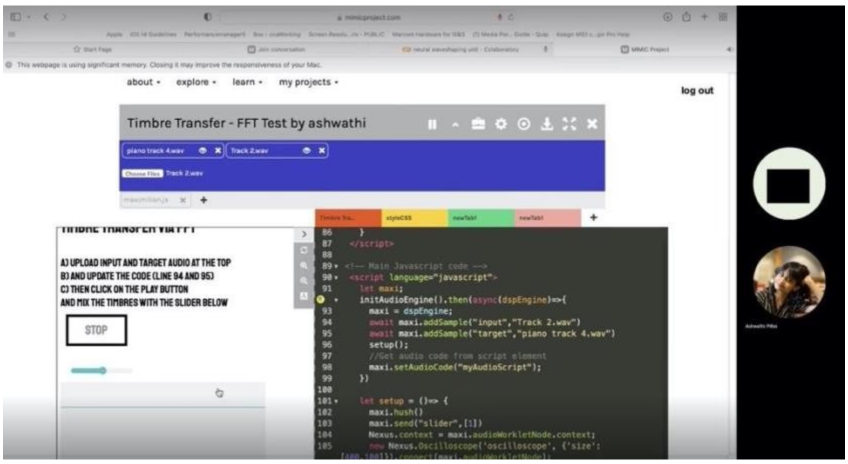


Figure 8: Tool Testing the Fourier-based model with Ruby

6.2. User Experience Evaluation:

In order to collect feedback from the usability tests, a user experience evaluation strategy was designed which comprised of two surveys: a) System Usability Scale Survey, and b) Participant Demographic + Preferences Survey.

6.2.1. Survey Structure:

a) System Usability Scale:

All the participants were asked to report their experience interacting with the models' user interface by filling out a System Usability Scale (Brooke, 1995) questionnaire (see Appendix G). The researcher chose this method due to it being a frequently used tool in the field of user experience, with over 1300 references in articles and publications (Sauro, 2022). The System Usability Scale serves as a quick tool to use on small sample sizes for reliable results on how usable a system is, without focusing on diagnostic issues. Google Forms, because of its accessibility and widespread use, was selected as the platform to host the questionnaire, which consisted of 10 statements along with a Likert-scale, a commonly used bipolar symmetric scaling method for quantitative research. This enabled the participants to comfortably express their agreement or disagreement with each statement using five response options, ordered in an ascending manner from 'Strongly Disagree' to 'Strongly Agree' to avoid left-side selection bias (Yonnie Chyung, Kennedy and Campbell, 2018). The form also had an optional field for the participants to type out any particular comments they had about each model which might not have been covered in the 10-item questionnaire.

6.2.1. Data Analysis:

a) System Usability Scale

All participants ranked each statement from 1 to 5 based on how much they agreed with it, with 'Strongly Disagree' being 1, and 'Strongly Agree' being 5 (see Step 1).

Step 1:				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5
Step 2:		**Step 3:**		**Step 4:**
X = Sum of the points for all odd numbered questions	Y = Sum of the points for all even numbered questions	X0 = X – 5 Y0 = 25 - Y	→	SUS Score = (X0 + Y0) x 2.5
Step 5:				
> 80.3 = A	68 – 80.3 = B	68 = C	51 – 68 = D	< 51 = F

Figure 9: System Usability Scale – Score Calculation

The way the SUS questionnaire is designed is that every odd numbered item is positively framed, and every even numbered item is negatively framed. So, after calculating the sum of all odd numbered items (stored in 'X') and the sum of all even numbered items (stored in 'Y'), 5 is subtracted from 'X' resulting in X0 and 'Y' is subtracted from 25 resulting in Y0 – which then can be added together to give us the SUS Score (as illustrated in Figure 9). The maximum possible score a design system could get is 100 points, with each item having a weightage of 10 points. Since a 'strongly agree' response for a positive question (odd-numbered item) would get 10 points straight away, and vice versa for 'strongly disagree' for a positive question i.e. 0 point, subtracting 1 from every odd numbered question (Step 3) ensures that the minimum is zero, and maximum being 10 for each question after it is multiplied by 2.5. The formula illustrated in the figure above makes the whole calculation very straightforward and quick.

There are many ways to interpret SUS scores, the researcher decided to go with the grading system due to its simplicity. As illustrated in Step 5, a score above 80.3 or higher would receive an A - which means it's a decently designed system. A score around 68 (which tends to be the average) would receive a C - meaning it needs some improvements. Anything below 51 is an F - which means usability needs to be a priority (Brooke, 1995).

7. RQ2 – FINDINGS:

7.1. System Usability Scale Scores:

The scores from each usability test is illustrated below:

SUS SCORES	
DDSP-based Model	**Fourier-based Model**
60	47.5
60	42.5
95	52.5
65	55
85	65
47.5	37.5

Figure 10: SUS Scores

The DDSP-based model got higher scores compared to the Fourier-based model, with the former's median being 62.5 (Final Score -> D), and the latter's median being 50 (Final Score -> F) indicating that both models needed further improvement and refinement for it to be usable.

7.2. Qualitative Feedback:

The form also had a field for the participants to point out any particular comments about the models. Some suggestions for improvement from them were:

> Nimbus wrote:
>
> *"If there is an option to normalize the volume down to lower volumes for the second model, maybe the blending will be easier, because I was noticing intense clipping in the final blended audio. Sliders to control the volumes of the two input audios could help."*

A common sentiment amongst most participants was that they liked the creative freedom that the Fourier-based model gave them of choosing their own input and target audio, but struggled to recognise the synthesised audio's timbre shift. Olivia also found the generated sound from the model really stressful, and couldn't figure out how to make the stop button work.

> Thomas wrote:
>
> *"This is really cool. But if you think about Google's Deep Dream, the whole fun of it is that it's going to come up with a totally strange, psychedelic image that you could argue would be hard to recreate by a person. Or wouldn't necessarily think to create something like that. It's just very strange, but that's what's interesting about it for me. Something similar with sound would be cool, like you give it a sound sample and tell it to make it sound like a garbled up, distorted sound, or a weird, dreamy, psychedelic sound. Basically something abstract. I think that might be way more interesting or way more fun, as with this I'm expecting it to sound like a flute but then when it doesn't it's like what the f*** was this thing?"*

The DDSP-based model's common complaint was that the generated audio sounded nothing like the actual instrument sometimes, and hence the participants were disappointed because their expectations weren't met. Ruby suggested that having synth sounds for the target audio, instead of only focusing on real instruments, might be a better approach to manage expectations.

> Wei wrote:
> *"It was really fun to be able to run an earworm song through the model and get little clips of it that sound exactly like how I hum it in my head, so I don't have to actually manually hum it. It'll be more convenient to use if it's an app instead of a webpage, but I really like the functionality of it so far!"*

Participants expressed that they would prefer using a mobile application instead so that they could stim freely with their headphones while mixing timbres, instead of being forced to sit in front of their computers. Most participants also mentioned that they disliked having to play around with code in the second model, and that they didn't like Google Colab's interface. "I can't tell when the cell is done executing, the green tick on the left isn't distinct enough I think. It needs to be more clear, or bigger.", Nimbus commented while user-testing the first model.

8. DISCUSSION:

The findings from the first research question's study illustrated a robust pattern of repetitive, sensory-seeking stimming to naturally occurring sounds, with some participants enjoying breaking down elements of music into smaller components to hyperfocus on. The investigation heavily relied on the research participants lived experiences of their childhood and adulthood, with a focus on stimming, music, and technology, and indicated that timbre transfer has a lot of potential to support such stims and improve autistic wellbeing. Findings from the second research question’s study pointed towards the fact that despite being a powerfully sensory and fun technique, neither of the timbre transfer approaches that were tested met the participants expectations due to falling short on certain design and accessibility factors. Timbre transfer also presents itself as a use case for musicians that are not autistic too, and might want to try out new ways of making music, edit their field recordings for production work, or use it as an alternative to autotune methods if required. Overall, the findings of this study underscores the importance of understanding potential benefits and limitations of timbre transfer techniques in relation to supporting the wellbeing of individuals on the autism spectrum.

9. LIMITATIONS:

The findings from this study should be considered along with a number of caveats. To begin, the participant pool was really small, limiting the breadth of findings that could be drawn from the collected data, which inadvertently also reduced the number of statistical methods that could be used to quantitatively analyse the same. All research participants were also cognitively capable of filling out surveys and questionnaires independently. Therefore, despite the sample size being very culturally and musically diverse, it was not fully representative of the autistic community with varying needs i.e., non-verbal autistic individuals, autistic people with differing cognitive or motor skills, etc. Moreover, having just one participant from each ethnic backgroundwas not ideal too as they cannot fully represent the views and experiences of their community, especially since the research participants were residing in different places geographically as well. The participant pool also had an unbalanced mix of musicians and non-musicians. Despite it making more sense to recruit more autistic musicians, one cannot conclude that such a tool would be beneficial for autistic people in general for stimming if non-musicians are not equally involved in the user-testing phase. Lastly, given that only one researcher was coding all the data to uncover findings, the chances of bias in the finalised themes is something worth considering too.

The two models that were used for testing also had some confounding variables. Since the Fourier-based model had a continuous mode of interaction, and the DDSP-based model had a discrete mode of interaction, testing both of them together wasn't ideal. The researcher initially decided to test the VST-plugin from Google (Magenta DDSP, 2022), along with the along with the DDSP autoencoder code via Google (Engel et al., 2020) to avoid confounding variables, but decided to avoid that as not all the research participants were well-versed with installing plugins into a digital audio workstation. Also, the Fourier-based model had no manual parameters to tweak in order to refine the output, whereas the DDSP-based model gave enough control to the user before the output was generated, which could possibly have led to a bias during the user experience evaluation. Due to time and resource constraints, along with limited prior experience with programming and machine learning, the researcher couldn’t attend to these issues adequately before the second usability tests for both the models were scheduled.

10. RECOMMENDATIONS FOR FUTURE WORK:

The next iteration of this study could involve organizing focus groups with autistic people of a particular ethnic group from the same country, along with including autistic people with varying needs, to get more detailed and representative insights from the analysis in order to understand if timbre transfer could be used to aid autistic stimming. Also, including more researchers during coding would be beneficial to avoid any bias. In order to ascertain how useful the model is, making use of diary studies instead of usability tests might be a better approach, as diary studies would give researchers more accurate results on how effective it is for stimming in real time, which are detailed insights that won't be possible to generate via simple 20 minute usability tests.

Few research participants also mentioned that one of the reasons they disliked most digital tools was that there was almost never any tactile input, and would love to play a digital music tool that would respond to pressure i.e., the harder the user presses a button, the louder the sound emitted, etc. Some recommendations for motion-detecting music tools were also made. Jared said, *"Don't know if this can be an Alexa skill, but it would be awesome if Alexa could play a particular song when I'm stimming in a happy way or when I'm angry. I tend to pace quite fast when I'm stressed or angry, and only certain kinds of music helps me around that time. If an AI can notice how I'm feeling based on how I'm stimming and play songs responding to that, I would want to buy that. That would really help me keep track of my emotions too."* An emotion-based tool like that would clearly have a lot of ethical concerns with regard to sensitive data leakages to third party organisations, but further exploring motion-based applications,possibly using Wekinator (Fiebrink, 2022), and giving the user the freedom to move around freely while causing the model to emit sounds they like, is something worth considering in order for the technology to truly serve as a tool to aid and encourage autistic stimming.

11. ETHICAL CONCERNS:

Making use of deep learning models such as timbre transfer to create art brings forth a slew of ethical concerns, one of them being the validity of algorithmic art being seen as honest due to its lack of intent. While this statement for a model that completely auto-generates art might hold true, most artificially intelligent models do give their users some creative control over the outputs that are generated. Moreover, the assumption that producing creative and honest art being difficult to algorithmically automate, as opposed to physical labour jobs, may be based on classist notions we may subconsciously hold about certain types of work, rather than a comprehensive understanding of what is truly hard or sincere.

The art world is known to be highbrow, and has a long history of classism and discrimination, with the wealthy and socially powerful seldom having any barriers to entry, whereas artists from particular marginalised demographics or colonized countries statistically underperforming as compared to their western, more influential counterparts (Boucher, 2019). Music schools also charge astronomically high tuition fees, and investing in certain acoustic instruments, such as the piano, might be financially difficult for some individuals from middle class backgrounds. Even disabled people, who might have restricted motor or cognitive capabilities, might be discouraged from making music using traditional approaches.

Since people from marginalized populations have equally rich and meaningful inner experiences that could be communicated through art, one could argue that advances in artificial intelligence, with regard to computational creativity, could potentially democratize the playing field - giving everyone access to appropriate tools, anda fair chance to artistically express their thoughts and life experiences. Honest art does not need to be a by-product of traditional approaches only. If an artwork produced using a deep learning model can make us feel things, or help us understand a new perspective - we are in the midst of honesty and humanness. One could also argue that these tools are similar to how smartphone photography made taking photos very accessible to people who might not possess the technical skills to use an analog camera, or have the money to invest in such equipment. The advances in the smartphone photography did not replace professional photographers, but just made it easier for regular people to preserve memories without spending a lot of money, or to practice photography before investing a lot of money in it. Art produced through deep learning models also gives us an opportunity to experience weird yet interesting art that might not have been possible to make before, or for humans to even think about creating.

12. CONCLUSION:

Over the last two decades, research into autism has grown into a business spawning a plethora of books, workshops, and interventions to cure or prevent the same. However, a diagnosis of autism is seldom accompanied by personalized support plans to aid sensory and emotional dysregulation. There is also a history of misunderstanding in the field, with autism literature referring to autistic people as "non-persons" or "behaviour problems", and popular media portrayals of autistic characters often being either antagonistic or helplessly naïve, leading to further negative misconceptions about the neurotype amongst the general public. Since behaviour is highly interpretable, it is imperative to gather primary data about stimming directly from autistic people, and to rethink assumptions or preconceived notions about restricted repetitive behaviours to truly support autistic wellbeing.

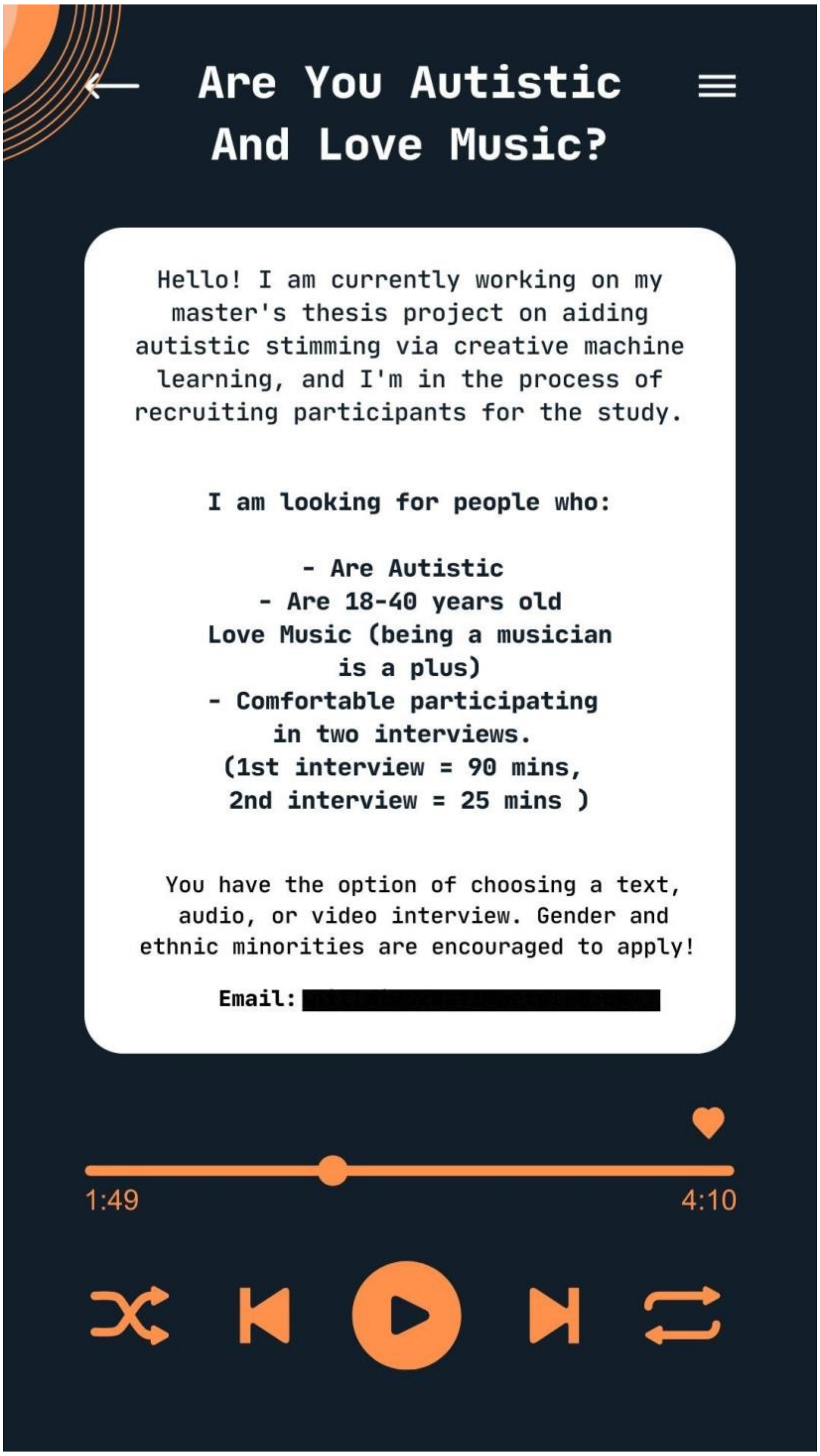
In this thesis, the researcher explored two different types of timbre transfer models as an accompanying tool to aid and encourage self-stimulatory behaviours, or stimming, in autistic adults through in-depth interviews and usability tests. Findings from this research study illustrated that participants stimmed in a wide variety of methods, including stimming to music and natural sounds that they heard in their immediate environment. The participants found timbre transfer to be an interesting approach to aid autistic stimming, and worth exploring further. However, there is still much work to be done in order to confidently conclude that such a tool could be utilised to support stimming, as discussed in Section 9 and 10.

As a closing statement, the researcher hopes that this study highlights the importance of involving and including autistic people in research that directly affects them, and the value of further investigating the benefits of combining knowledge from both music and assistive technology domains to improve autistic wellbeing.

13. APPENDICES:

A) Recruitment Material

Infographic:



Writeup:

Hello!

I’m a masters student studying Data Science and AI for the Creative Industries at UAL, and I’m currently working on my thesis project.

I am looking for six volunteers for my research study on understanding the experiences of stimming directly from autistic people, and to explore whether music could support and encourage that. The results from this study would be used to develop a music tool using machine learning that autistic people can make use of to produce music as a tool for auditory stimming, or just as a hobby to support their interest in music.

The criteria for the participants I’m looking for is as follows:

- Autistic adults
- Ages between 18-40
- Interest in music (being a musician is a plus!)
- Comfortable participating in two interviews (1st interview = 90 mins, 2nd interview = 25 mins). Participants will have the following options: Text chat, audio chat, or video chat on Microsoft Teams.

Gender and ethnic minorities are encouraged to apply! This project aims to empower autistic people by centering their experiences during research, therefore the recruitment stage of this study is crucial to adequately represent their views.

Thank you, and please feel free to reach out to me via email if you have any questions or would like further information about this study!

- A. Pillai

B) Ritvo Autism & Asperger Diagnostic Scale:

Link: https://psychology-tools.com/test/raads-14

Ritvo Autism & Asperger Diagnostic Scale (RAADS-14)

Tests | About Us | Get Help

The purpose of the study behind this instrument was to develop a screening instrument to aid in the identification of patients who may have undiagnosed ASD.

Please respond with the answer that most accurately describes how each of the statements below applies to you. For the purposes of this test, "When I was Young" refers to the age of 17 or younger.

	True Now & When Young	True Only Now	True When I Was Young	Never True
1. It is difficult for me to understand how other people are feeling when we are talking.	○	○	○	○
2. Some ordinary textures that do not bother others feel very offensive when they touch my skin.	○	○	○	○
3. It is very difficult for me to work and function in groups.	○	○	○	○
4. It is difficult to figure out what other people expect of me.	○	○	○	○
5. I often don't know how to act in social situations.	○	○	○	○
6. I can chat and make small talk with people.	○	○	○	○
7. When I feel overwhelmed by my senses, I have to isolate myself to shut them down.	○	○	○	○
8. How to make friends and socialize is a mystery to me.	○	○	○	○
9. When talking to someone, I have a hard time telling when it is my turn to talk or to listen.	○	○	○	○
10. Sometimes I have to cover my ears to block out painful noises (like vacuum cleaners or people talking too much or too loudly).	○	○	○	○
11. It can be very hard to read someone's face, hand, and body movements when we are talking.	○	○	○	○
12. I focus on details rather than the overall idea.	○	○	○	○
13. I take things too literally, so I often miss what people are trying to say.	○	○	○	○
14. I get extremely upset when the way I like to do things is suddenly changed.	○	○	○	○

Score my Answers

Sources

C) Participant Information Sheet

PARTICIPANT INFORMATION SHEET

Aiding Autistic Stimming via Creative Machine Learning Technology

Researcher: A. Pillai

Supervisor: Prof. Mick Grierson

Creative Computing Institute, University of the Arts London

Thank you for considering to take part in this study. This information sheet outlines the purpose of the study, and provides a brief description of your involvement and rights as a participant, if you do agree to take part. Please take time to read the following information carefully, and feel free to ask me questions if anything you read is not clear, or if you would like more information.

1. Who is conducting this study?

My name is A. Pillai and I shall be conducting this study as part of my thesis. I am a Masters student studying Data Science and AI for the Creative Industries at UAL's Creative Computing Institute.

2. What is the research about?

This research is about understanding the experiences of stimming directly from autistic people, and to explore whether music could be used to support that. The results from this study will be used to develop a musical interface using machine learning that autistic people can use to produce music as a tool for stimming, or just as a hobby to support their interest in music.

3. Do I have to take part?

Participation is completely voluntary. You do not have to take part if you do not wish to. If you do decide to take part, I shall ask you to sign a consent form prior to the interview. Also, at any given point during the interview, you wish to stop and withdraw from this study, please do let me know. I shall conclude the interview and discard your data. Withdrawing from the study will have no effect on you.

4. What will my involvement be?

Taking part in this study would involve two interviews. The first interview would be focused on understanding your personal experiences as an autistic adult with stimming to self-regulate emotions, along with your relationship with technology, your music preferences/listening habits, and your previous experiences with music production (if any). This information shall help me work on a prototype revolving around music to support stimming behaviours. I shall also ask you to test out an existing solution to better understand which features/aspects of the tool works well or doesn't work well for you. The second interview would just be a user testing interview wherein I would ask you to play around with the prototype I've worked on in order to collect feedback.

5. How are these interviews structured?

The first interview would last for about 90 minutes, and the second interview would last for about 25 minutes. Participants will have the following options: text, audio, or video interviews on Microsoft Teams.

6. Do I have to answer every question?

If any question during the interview makes you feel uncomfortable, you do not have to answer it. Refusing to disclose any kind of information will have no effect on you.

7. What will my information be used for?

Your information will be used for research purposes for my masters thesis, and might be used for future research or publications.

8. Will my taking part and my data be kept confidential? Will it be anonymised?

Your participation will be completely anonymous - your name will not be used in any reports or video presentation resulting from this study. The only socio-demographic data that will be collected from you is your age, gender, and ethnic background, and I shall be using a pseudonym to safeguard your identity. Lastly, none of this data shall be shared with any third party.

D) Consent Form:

CONSENT FORM

By participating in this study, you agree to the following statements:

- You confirm that you have read and understood all the information outlined in the participant information sheet prior to participating.

- You acknowledge that you have had the opportunity to ask questions about the study, either verbally or via writing.

- You understand that participation in this study is voluntary and that you are free to withdraw at any time without giving the researcher any reason.

- You understand that the researcher named above will have access to the research data collected during this project.

- You agree to your interview being audio-recorded and retained by the researcher until November 2022.

- You understand that any mention of your identity in the findings of the thesis report or video presentation shall remain anonymous.

- You agree to take part in the research project.

Signed:______________________________ Date:______________

E) Interview Guide:

IN-DEPTH INTERVIEW + USABILITY TEST

INTRODUCTION AND ASSENT:

Hello {participant name}! My name is A. Pillai, and I'm a masters student at UAL's Creative Computing Institute, studying Data Science and AI for the Creative Industries. In today's session, we're going to be talking about music, technology, and your experiences with stimming. I shall then ask you to play around with Google Magenta's Tone Transfer to better understand how you interact with it, and ask if there's anything in particular that you like or don't like. This session will take around 70-90 minutes, and I shall be recording this interview and transcribing it to remember all your comments. To protect your identity, I shall also be using a pseudonym for you in my study. Do you have a preferred pseudonym that you'd like me to use? *{Response}*

As participation in this study is completely voluntary, you have the choice to ask me to stop if you change your mind at any point during this interview. All you have to do is tell me to stop, and it will have no effect on you. Most importantly, if you do not want to answer any particular question, or are feeling any negative emotions during the interview, please do let me know.

Lastly, feel free to ask me any question if something isn't clear. You can ask me questions at any point during this interview.

Do you agree to participate in this study? *{Response}*

Do you consent to your interview being recorded and transcribed? *{Response}*

BASIC QUESTIONS:

1) What is your age?

2) What gender do you identify with?

3) What sex were you assigned at birth?

4) What is your ethnicity?

STIMMING:

1) (If formally diagnosed) At what age did you receive your diagnosis? (if self-diagnosed) At what age did you suspect that you might be autistic?

2) Do you stim? And if so, how? (If many forms): Which ones do you do the most?

3) How long or often do you stim in a day?

4) Have you ever tried stimming in public?

5) Do you use something while stimming? Like a chew toy or a lava lamp?

6) How much time do you spend stimming in a day, if you had to give a rough estimate?

7) Are there any particular time-periods or situations that causes you to stim?

8) How do you feel while stimming?

9) Have you ever tried stimming to manage overwhelming or negative emotions? Did it help?

10) Has anyone ever told you to not stim? If so, could you tell me more about why you think they might have told you so?

11) Have you ever tried supressing your stims? If so, how long was it for and how did that make you feel?

12) Lastly, has stimming ever caused you, or someone you know, any harm?

MUSIC:

1) What genres of music do you usually listen to?

2) How often do you listen to music in a day, if you had to give a rough estimate?

3) What is your relationship with music? Are you a musician? Or an avid music listener?

4) (If musician): Could you talk me through your process of making music? Do you use any particular tools / instruments?

5) (If musician): Have you explored any AI tools in your creative practise? Could you talk more about them? (if no): Would that be something you'd be interested to try?

6) (If musician): What would you say were some of the main obstacles while learning how to make music?

8) (Musician/non-musician): Have you ever used music or natural sounds to stim?

TECHNOLOGY:

1) What are your views on emerging technologies? Especially AI?

2) Do you feel like modern technological interfaces are designed keeping disabled people in mind?

3) Has there been a time where you found a particular software or application hard to use? Could you explain why?

4) If someone were to build a musical interface for creating music for you, either to stim or to just make music - how would you want it to be like? Is there something they should be mindful of?

--

Okay great! Thank you for your comments, we shall now move on to the second part of the interview.

[Brief Introduction to Google Magenta's Tone Transfer]

USABILITY TEST:

Task: To create two synthesized sounds using Tone Transfer, first sample could be using one of Google's provided audio inputs, the second could be an uploaded/recorded audio file. The participant would have 15 minutes to do so.

Testing Questions:

1) (After explaining tone transfer) What are your thoughts on using this feature for music creation?

2) Do you feel that this model could be used to aid vocal / auditory stims?

3) What do you think of the language used throughout this interface?

4) What are your thoughts on the design and layout?

Prompts if stuck during task:

1) If you were to finetune the parameters of the audio output, where would you expect to find it?

2) Do you have a favourite song or a current auditory stim recorded that you'd like to explore here?

3) Do you want to try recording yourself playing your favourite musical instrument as the input?

4) Where would you expect to find the download button for the output file?

Post Testing Questions:

1) What did you like the most about this tool?

2) What did you like the least about this tool?

3) How frequently would you use a similar product?

[INTERNAL] Did they finish the task on time? Which areas were they stuck at?

-

DEBRIFING SCRIPT:

That's all for today! Thank you so much for participating. The goal of this research is to better understand the experiences of stimming directly from autistic people, and to explore whether music could be used to support that. Your comments are extremely valuable to this study, and I'm very grateful for your time and participation.

For the next interview, it would be great if you could bring some recorded audio samples of naturally occurring sounds that you've been stimming to, or a .wav file of your current stim song / composition. We'll then explore how they sound using different instruments similar to how we did today using Tone Transfer.

Before we close, do you have any questions for me? Or anything you'd like to add that you feel wasn't covered during the interview? *{Response}*

SECOND USABILITY TEST

INTRODUCTION:

Hello! Hope you've been well since we last spoke. I'm really excited about today's session as we'll get to explore two models that I've been working on, but this time specifically using sound samples you've been stimming to.

[Brief Introduction to both the models]

USABILITY TEST:

Task: To create two synthesized sounds using both the models. The participant would have 10 minutes for each model, and they are meant to do this task independently by reading all the instructions mentioned alongside each model, but can ask questions to the researcher if they are stuck anywhere.

USER EXPERIENCE EVALUATION:

Alright! Could you please fill out simple questionnaire for both the models? It should take you around 3 minutes to complete. Do ask me questions if you're stuck anywhere!

DEBRIEFING SCRIPT:

Thank you once again for your participation in this study. This concludes your involvement in the study, is there anything else you'd like to add before we close? *{Response}*

Would you like to receive the final report of all the findings once I'm done? *{Response}*

F) Distress Protocol:

DISTRESS PROTOCOL

Written after reading best practises for designing a distress protocol for sensitive research studies (Source: Paper)

i. Distress:- If a participant verbally mentions that they're feeling uncomfortable, or is exhibiting behaviours that would suggest that they are in distress i.e. crying, fidgeting, shaking, etc. go to next step
ii. Stage 1 Response:- Temporarily pause the interview, and assess mental health status: "Could you tell me what thoughts you are having? Would you like any external support? Do you feel okay/safe?"
iii. Review:- If the participant mentions that they are okay and wishes to carry on, resume the interview. If not, go to next step.
iv. Stage 2 Response:- Conclude the interview and take the participant to a quiet area (if this is an in-person interview) and encourage the participant to contact their GP/mental health practitioner. In case the participant does not have any access to medical support, provide them with contact details of a mental health charity like the Samaritans (with their consent). Also, discard any collected data on them.
v. Follow up: After a week, follow up with the participant with a courtesy call to check up on them (if the participant consents).

G) User Experience Evaluation Survey:

Link: https://forms.gle/B87JKoA3A9b7bv1t5

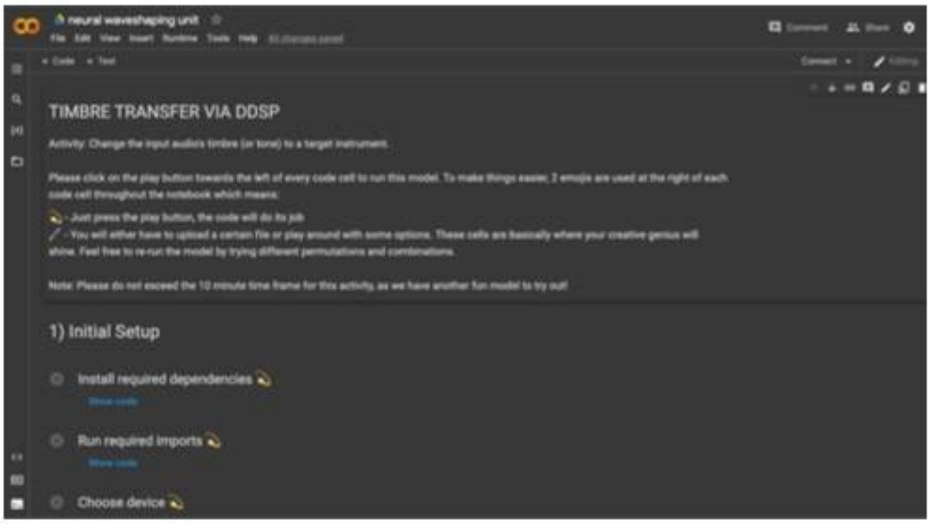
User Experience Evaluation Survey

Hello! Thank you once again for your interest in this study. Please fill out this form to help me evaluate the usability of both the timbre transfer models that we just tested. This form should take about 5-8 minutes to fill out. Please feel free to ask me questions if you need an explanation or clarification while filling this out.

* Required

1) Timbre Transfer via DDSP Approach (NEWT)

Please share your experience of using this model through the scale below.
There are 10 questions, with 5 response options ranging from Strongly Disagree to Strongly Agree (with an option to choose Neutral too but try to avoid this option if you can!)



1. Rate your level of agreement with each statement *

Mark only one oval per row.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. I think that I would like to use this system frequently.	⬭	⬭	⬭	⬭	⬭
2. I found the system unnecessarily complex	⬭	⬭	⬭	⬭	⬭
3. I thought the system was easy to use	⬭	⬭	⬭	⬭	⬭
4. I think that I would need the support of a technical person to be able to use this system	⬭	⬭	⬭	⬭	⬭
5. I found that the various functions and design in this system were well integrated	⬭	⬭	⬭	⬭	⬭
6. I thought there was too much inconsistency in this system	⬭	⬭	⬭	⬭	⬭
7. I would imagine that most people would learn to use this system very quickly	⬭	⬭	⬭	⬭	⬭
8. I found the system very cumbersome to use.	⬭	⬭	⬭	⬭	⬭
9. I felt very confident using the	⬭	⬭	⬭	⬭	⬭

system

10. I needed to learn a lot of things before I could get going with this system	◯	◯	◯	◯	◯

2. Are there any other comments you would like to add about this model?

2) Timbre Transfer via Fourier Based Approach (MIMIC)

Please share your experience of using this model through the scale below. There are 10 questions, with 5 response options ranging from Strongly Disagree to Strongly Agree (with an option to choose Neutral too but try to avoid this option if you can!)

3. Rate your level of agreement with each statement

Mark only one oval per row.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. I think that I would like to use this system frequently.	◯	◯	◯	◯	◯
2. I found the system unnecessarily complex	◯	◯	◯	◯	◯
3. I thought the system was easy to use	◯	◯	◯	◯	◯
4. I think that I would need the support of a technical person to be able to use this system	◯	◯	◯	◯	◯
5. I found that the various functions and design in this system were well integrated	◯	◯	◯	◯	◯
6. I thought there was too much inconsistency in this system	◯	◯	◯	◯	◯
7. I would imagine that most people would learn to use this system very quickly	◯	◯	◯	◯	◯
8. I found the system very cumbersome to use.	◯	◯	◯	◯	◯
9. I felt very confident using the	◯	◯	◯	◯	◯

system					
10. I needed to learn a lot of things before I could get going with this system	○	○	○	○	○

4. Are there any other comments you would like to add about this model?

That's it! Thank you so much for filling this out, and for voluntarily giving me your time for this study. Your inputs are extremely valuable to me :)

14. BIBLIOGRAPHY:

1. Adkins, S., 2022. *We Move Session: Sara Adkins - Machine Learning and Composition*, Music Hackspace [online] Youtube.com. Available at: <https://www.youtube.com/watch?v=dQfN4hL_YIk&t=2478s> [Accessed 1 September 2022].
2. Alpers, B., 1941. The Use of Physical Therapy Measures in Neurological Conditions. *Physical Therapy*, [online] 21(6), pp.299-304. Available at: <https://academic.oup.com/ptj/article-abstract/21/6/299/4747811?redirectedFrom=fulltext>.
3. American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. [online] 28(3), pp.36-37. Available at: <https://cdn.website-editor.net/30f11123991548a0af708722d458e476/files/uploaded/DSM%2520V.pdf>.
4. Arik, S., Jun, H. and Diamos, G., 2019. Fast Spectrogram Inversion Using Multi-Head Convolutional Neural Networks. *IEEE Signal Processing Letters*, [online] 26(1), pp.94-98. Available at: <https://arxiv.org/pdf/1808.06719.pdf>.
5. Askham, A. and Dattaro, L., 2022. *Backlash from autistic community pauses research, exposes communication gaps | Spectrum | Autism Research News*. [online] Spectrum | Autism Research News. Available at: <https://www.spectrumnews.org/news/backlash-from-autistic-community-pauses-research-exposes-communication-gaps/> [Accessed 22 August 2022].
6. Baron-Cohen, S., Ashwin, E., Ashwin, C., Tavassoli, T. and Chakrabarti, B., 2009. Talent in autism: hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, [online] 364(1522), pp.1377-1383. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2677592/>.
7. Basu, A., 2022. *Explained: What is Google's Tone Transfer tool - Times of India*. [online] The Times of India. Available at: <https://timesofindia.indiatimes.com/gadgets-news/explained-what-is-googles-tone-transfer-tool/articleshow/91545662.cms> [Accessed 31 August 2022].
8. Baumeister, R., Bratslavsky, E., Muraven, M. and Tice, D., 1998. Ego depletion: Is the active self a limited resource?. *Journal of Personality and Social Psychology*, [online] 74(5), pp.1252-1265. Available at: <https://faculty.washington.edu/jdb/345/345%20Articles/Baumeister%20et%20al.%20%281998%29.pdf>.
9. Bennie, M., 2019. *Understanding Stimming and Autism: The Good and Bad Side of Anxious Behaviours*. [online] Autism Awareness Centre. Available at: <https://autismawarenesscentre.com/stimming-the-good-and-bad-of-anxious-behaviours/> [Accessed 29 August 2022].

10. Boucher, B., 2019. It's Not All in Your Head—the Art World Really Is Unfair. Here Are 9 Reasons Why | Artnet News. [online] Artnet News. Available at: <https://news.artnet.com/art-world/9-reasons-art-world-is-unfair-1726653> [Accessed 21 September 2022].
11. Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, [online] 3(2), pp.77-101. Available at: https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa?needAccess=true .
12. Brooke, J., 1995. SUS: A quick and dirty usability scale. [online] Available at: https://www.researchgate.net/publication/228593520_SUS_A_quick_and_dirty_usability_scale [Accessed 21 September 2022].
13. Buijsman, R., Begeer, S. and Scheeren, A., 2022. 'Autistic person' or 'person with autism'? Person-first language preference in Dutch adults with autism and parents. *Autism*, [online] p.136236132211179. Available at: <https://journals.sagepub.com/doi/full/10.1177/13623613221117914>.
14. Cage, E. and Troxell-Whitman, Z., 2019. Understanding the Reasons, Contexts and Costs of Camouflaging for Autistic Adults. *Journal of Autism and Developmental Disorders*, [online] 49(5), pp.1899-1911. Available at: <https://link.springer.com/article/10.1007/s10803-018-03878-x>.
15. Cassidy, S., Hannant, P., Tavassoli, T., Allison, C., Smith, P. and Baron-Cohen, S., 2016. Dyspraxia and autistic traits in adults with and without autism spectrum conditions. *Molecular Autism*, [online] 7(1). Available at: <https://molecularautism.biomedcentral.com/articles/10.1186/s13229-016-0112-x#:~:text=Dyspraxia%20is%20significantly%20more%20prevalent,autistic%20traits%20and%20lower%20empathy.>.
16. Chamak, B., Bonniau, B., Jaunay, E. and Cohen, D., 2008. What Can We Learn about Autism from Autistic Persons?. *Psychotherapy and Psychosomatics*, [online] 77(5), pp.271-279. Available at: <https://www.jstor.org/stable/48511100>.
17. Chapman, R., 2021. *Why Autistic People Are Worried by Spectrum 10K*. [online] Psychology Today. Available at: <https://www.psychologytoday.com/gb/blog/neurodiverse-age/202108/why-autistic-people-are-worried-spectrum-10k> [Accessed 23 August 2022].
18. Charlton, R., Entecott, T., Belova, E. and Nwaordu, G., 2021. "It feels like holding back something you need to say": Autistic and Non-Autistic Adults accounts of sensory experiences and stimming. *Research in Autism Spectrum Disorders*, [online] 89, p.101864. Available at: <https://doi.org/10.1016/j.rasd.2021.101864>.
19. Chase, S., 2022. *What Is Timbre In Music? A Complete Guide with Videos*. [online] Hellomusictheory.com. Available at:

<https://hellomusictheory.com/learn/timbre/> [Accessed 1 September 2022].

20. Cibrian, F., Peña, O., Ortega, D. and Tentori, M., 2017. BendableSound: An elastic multisensory surface using touch-based interactions to assist children with severe autism during music therapy. *International Journal of Human-Computer Studies*, [online] 107, pp.22-37. Available at: <https://reader.elsevier.com/reader/sd/pii/S1071581917300757?token=7318F0ADD478181CE4D97F144142298D8459D6553B44A95461900DDE7E9C7A65F7CBD053E210C8B85B92B01CE2CB91B1&originRegion=eu-west-1&originCreation=20220830193243>.
21. Cooper, K., Russell, A., Calley, S., Chen, H., Kramer, J. and Verplanken, B., 2021. Cognitive processes in autism: Repetitive thinking in autistic versus non-autistic adults. *Autism*, [online] 26(4), pp.849-858. Available at: <https://journals.sagepub.com/doi/epub/10.1177/13623613211034380>.
22. Crosman, C., 2022. *Why We Should Stop Stim Suppression*. [online] In the Loop About Neurodiversity. Available at: <https://intheloopaboutneurodiversity.wordpress.com/2019/09/17/why-we-should-stop-stim-suppression/> [Accessed 29 August 2022].
23. Davidson, J., 2010. 'It cuts both ways': A relational approach to access and accommodation for autism. *Social Science & Medicine*, [online] 70(2), pp.305-312. Available at: <https://doi.org/10.1016/j.socscimed.2009.10.017>.
24. Défossez, A., Zeghidour, N., Usunier, N., Bottou, L. and Bach, F., 2018. SING: Symbol-to-Instrument Neural Generator. *arXiv preprint*, [online] Available at: <https://arxiv.org/pdf/1810.09785.pdf> [Accessed 1 September 2022].
25. DeJonckheere, M. and Vaughn, L., 2019. Semistructured interviewing in primary care research: a balance of relationship and rigour. *Family Medicine and Community Health*, [online] 7(2), p.e000057. Available at: <https://fmch.bmj.com/content/7/2/e000057>.
26. Demetriou, E., Lampit, A. and Quintana, D., 2022. Autism spectrum disorders: a meta-analysis of executive function. *Molecular Psychiatry*, [online] Available at: <https://doi.org/10.1038/mp.2017.75> [Accessed 29 August 2022].
27. DeVita-Raeburn, E., 2022. *Is the Most Common Therapy for Autism Cruel?*. [online] The Atlantic. Available at: <https://www.theatlantic.com/health/archive/2016/08/aba-autism-controversy/495272/> [Accessed 29 August 2022].
28. Donahue, C., McAuley, J. and Puckette, M., 2019. Adversarial Audio Synthesis. *arXic preprint*, [online] Available at: <https://arxiv.org/pdf/1802.04208.pdf> [Accessed 1 September 2022].
29. Draucker, C., Martsolf, D. and Poole, C., 2009. Developing Distress Protocols for Research on Sensitive Topics. *Archives of Psychiatric Nursing*, [online] 23(5), pp.343-350. Available at: <https://svri.org/sites/default/files/attachments/2016-01-13/pdf_0.pdf>.

30. Dunn, D. and Andrews, E., 2015. Person-first and identity-first language: Developing psychologists’ cultural competence using disability language. *American Psychologist*, [online] 70(3), pp.255-264. Available at: <https://docksci.com/no-title_5a678782d64ab23767a5752d.html>.
31. Edgar, J., Fisk IV, C., Berman, J., Chudnovskaya, D., Liu, S., Pandey, J., Herrington, J., Port, R., Schultz, R. and Roberts, T., 2015. Auditory encoding abnormalities in children with autism spectrum disorder suggest delayed development of auditory cortex. *Molecular Autism*, [online] 6(1). Available at: <https://molecularautism.biomedcentral.com/articles/10.1186/s13229-015-0065-5>.
32. Engel, J., Agrawal, K., Chen, S., Gulrajani, I., Donahue, C. and Roberts, A., 2019. GANSynth: Adversarial Neural Audio Synthesis. *arXiv preprint*, [online] Available at: <https://arxiv.org/abs/1902.08710?> [Accessed 31 August 2022].
33. Engel, J., Smith, I., Choi, K., Hawthorne, C. and Dinculescu, M., 2020. Encoding Musical Style with Transformer Autoencoders. [online] Magenta. Available at: https://magenta.tensorflow.org/transformer-autoencoder [Accessed 25 September 2022].
34. Eriksson, J., 2013. *Ritvo Autism & Asperger Diagnostic Scale (RAADS-14)*. [online] Psychology-tools.com. Available at: <https://psychology-tools.com/test/raads-14> [Accessed 28 August 2022].
35. Euronews, 2015. *Pink Floyd's David Gilmour inspired by SNCF jingle*. [online] Available at: <https://www.euronews.com/culture/2015/07/20/pink-floyd-s-david-gilmour-inspired-by-sncf-jingle> [Accessed 1 September 2022].
36. Experia USA, 2022. Understanding the Different Types of Stimming. [online] Available at: <https://www.experia-usa.com/blog/understanding-different-types-stimming/> [Accessed 27 August 2022].
37. Fancourt, D. and Finn, S., 2019. World Health Organization, Regional Office for Europe, & Health Evidence Network. What is the evidence on the role of the arts in improving health and well-being? A scoping review. *Nordic Journal of Arts, Culture and Health*, [online] 2(1), pp.77-83. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK553773/>.
38. Faras, H., Al Ateeqi, N. and Tidmarsh, L., 2010. Autism spectrum disorders. *Annals of Saudi Medicine*, [online] 30(4), pp.295-300. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2931781/>.
39. Fereday, J. and Muir-Cochrane, E., 2006. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. International Journal of Qualitative Methods, [online] 5(1), pp.80-92. Available at: https://journals.sagepub.com/doi/full/10.1177/160940690600500107 .
40. Fiebrink, R., 2022. Wekinator | Software for real-time, interactive machine learning. [online] Wekinator.org. Available at: http://www.wekinator.org/ [Accessed 26 September 2022].

41. Fischbach, G., 2022. *Leo Kanner's 1943 paper on autism*. [online] Spectrum |Autism Research News. Available at: https://www.spectrumnews.org/opinion/viewpoint/leo-kanners-1943-paper-on-autism/ [Accessed 23 June 2022].
42. Fletcher-Watson, S., Adams, J., Brook, K., Charman, T., Crane, L., Cusack, J., Leekam, S., Milton, D., Parr, J. and Pellicano, E., 2018. Making the future together: Shaping autism research through meaningful participation. *Autism*, [online] 23(4), pp.943-953. Available at: <https://journals.sagepub.com/doi/epub/10.1177/1362361318786721>.
43. Furfaro, H., 2022. *Race, class contribute to disparities in autism diagnoses*. [online] Spectrum | Autism Research News. Available at: <https://www.spectrumnews.org/news/race-class-contribute-disparities-autism-diagnoses/> [Accessed 27 August 2022].
44. Geretsegger, M., Elefant, C., Mössler, K. and Gold, C., 2014. Music therapy for people with autism spectrum disorder. *Cochrane Database of Systematic Reviews*, [online] 2016(3). Available at: <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD004381.pub3/epdf/full>.
45. Gernsbacher, M., 2017. Editorial Perspective: The use of person-first language in scholarly writing may accentuate stigma. *Journal of Child Psychology and Psychiatry*, [online] 58(7), pp.859-861. Available at: <https://acamh.onlinelibrary.wiley.com/doi/full/10.1111/jcpp.12706>
46. Google, 2020. DDSP: DIFFERENTIABLE DIGITAL SIGNAL PROCESSING. [online] Available at: <https://openreview.net/attachment?id=B1x1ma4tDr&name=original_pdf> [Accessed 31 August 2022].
47. Google Magenta, 2020. Sites.research.google. 2022. *Tone Transfer - Magenta DDSP*. [online] Available at: <https://sites.research.google/tonetransfer> [Accessed 31 August 2022].
48. Gould, J., 2017. Towards understanding the under-recognition of girls and women on the autism spectrum. *Autism*, [online] 21(6), pp.703-705. Available at: <https://journals.sagepub.com/doi/10.1177/1362361317706174>.
49. Grierson, M., 2022. Maximilian.js Timbre Transfer Example. [online] Mimicproject.com. Available at: https://mimicproject.com/code/7ddcb284-72a1-275f-d3d5-0ec26aece5f8 [Accessed 26 September 2022].
50. Grierson, M., 2022. GitHub - micknoise/Maximilian: C++ Audio and Music DSP Library. [online] GitHub. Available at: https://github.com/micknoise/Maximilian [Accessed 26 September 2022].
51. Hagger, M., Wood, C., Stiff, C. and Chatzisarantis, N., 2010. Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin*, [online] 136(4), pp.495-525. Available at: <https://www.researchgate.net/publication/44689726_Ego_Depletion_and_the_Strength_Model_of_Self-Control_A_Meta-Analysis>.

52. Haigh, S., Walsh, J., Mazefsky, C., Minshew, N. and Eack, S., 2018. Processing Speed is Impaired in Adults with Autism Spectrum Disorder, and Relates to Social Communication Abilities. *Journal of Autism and Developmental Disorders*, [online] 48(8), pp.2653-2662. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6041130/>.
53. Harvey, L., 2022. *Researching the Real World - Semi Structured in-depth interviewing*. [online] Qualityresearchinternational.com. Available at: <https://www.qualityresearchinternational.com/methodology/RRW4pt2pt2Semistructured.php> [Accessed 29 August 2022].
54. Hashemian, P., Mashoogh, N. and Jarahi, L., 2015. Effectiveness of Music Therapy on Aggressive Behavior of Visually Impaired Adolescents. *Journal of Behavioral and Brain Science*, 05(03), pp.96-100. Available at: < https://www.researchgate.net/publication/273906271_Effectiveness_of_Music_Therapy_on_Aggressive_Behavior_of_Visually_Impaired_Adolescents>
55. Hayes, B., 2021. Neural Waveshaping Synthesis. [online] GitHub. Available at: https://github.com/ben-hayes/neural-waveshaping-synthesis/ [Accessed 25 September 2022].
56. Hill, E. and Frith, U., 2003. Understanding autism: insights from mind and brain. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, [online] 358(1430), pp.281-289. Available at: <https://www.jstor.org/stable/3558141>.
57. Huang, A., 2020. *This new music tool is weird, incredible, and free*. [online] Youtube.com. Available at: <https://www.youtube.com/watch?v=f8CdZxWj--A> [Accessed 31 August 2022].
58. IACC, 2022. *IACC Strategic Plan For Autism Spectrum Disorder Research*. [online] Available at: <https://iacc.hhs.gov/publications/strategic-plan/2010/> [Accessed 22 August 2022].
59. James, R., Sigafoos, J., Green, V., Lancioni, G., O'Reilly, M., Lang, R., Davis, T., Carnett, A., Achmadi, D., Gevarter, C. and Marschik, P., 2014. Music Therapy for Individuals with Autism Spectrum Disorder: a Systematic Review. *Review Journal of Autism and Developmental Disorders*, [online] 2(1), pp.39-54. Available at: <https://link.springer.com/article/10.1007/s40489-014-0035-4>.
60. Janzen, T. and Thaut, M., 2018. Rethinking the role of music in the neurodevelopment of autism spectrum disorder. *Music & Science*, [online] 1, p.205920431876963. Available at: <https://journals.sagepub.com/doi/full/10.1177/2059204318769639>.
61. Johnson, C., 2022. *Types of Stimming Infographic*. [online] Just Keep Stimming!. Available at: <https://justkeepstimming.com/2019/05/21/types-of-stimming-infographic/> [Accessed 27 August 2022].
62. Judd, T., 2021. *Scriabin's Piano Sonata No. 10: 'The Sun's Kisses'*. [online] The Listeners' Club. Available at: <https://thelistenersclub.com/2021/08/23/scriabins-piano-sonata-no-10-the-suns-kisses/> [Accessed 1 September 2022].

63. Kapp, S., Steward, R., Crane, L., Elliott, D., Elphick, C., Pellicano, E. and Russell, G., 2019. ‘People should be allowed to do what they like’: Autistic adults’ views and experiences of stimming. *Autism*, [online] 23(7), pp.1782-1792. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6728747/>.
64. Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C. and Pellicano, E., 2015. Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, [online] 20(4), pp.442-462. Available at: <https://ados2training.co.uk/uploads/9/4/8/0/94806962/autism_terminology_2015_paper.pdf>.
65. Kim, C., 2013. *A Cognitive Defense of Stimming (or Why “Quiet Hands” Makes Math Harder)*. [online] Musings of an Aspie. Available at: <https://musingsofanaspie.com/2013/06/18/a-cognitive-defense-of-stimming-or-why-quiet-hands-makes-math-harder/> [Accessed 29 August 2022].
66. Kim, S. and Bottema-Beutel, K., 2019. Negotiation of Individual and Collective Identities in the Online Discourse of Autistic Adults. *Autism in Adulthood*, [online] 1(1), pp.69-78. Available at: <https://www.liebertpub.com/doi/10.1089/aut.2018.0029>.
67. Kossyvaki, L. and Curran, S., 2018. The role of technology-mediated music-making in enhancing engagement and social communication in children with autism and intellectual disabilities. *Journal of Intellectual Disabilities*, 24(1), pp.118-138. Available at: < https://journals.sagepub.com/doi/full/10.1177/1744629518772648>
68. Kupferstein, H., 2018. Evidence of increased PTSD symptoms in autistics exposed to applied behavior analysis. *Advances in Autism*, [online] 4(1), pp.19-29. Available at: <https://www.emerald.com/insight/content/doi/10.1108/AIA-08-2017-0016/full/html>.
69. Lloyd, K. and White, J., 2011. Democratizing clinical research. *Nature*, [online] 474(7351), pp.277-278. Available at: <https://www.nature.com/articles/474277a>.
70. Lynch, C., 2019. *Invisible Abuse: ABA and the things only autistic people can see*. [online] NeuroClastic. Available at: <https://neuroclastic.com/invisible-abuse-aba-and-the-things-only-autistic-people-can-see/> [Accessed 29 August 2022].
71. Magenta DDSP, 2022. *DDSP-VST Neural Synthesis in your DAW*. [online] Available at: <https://magenta.tensorflow.org/ddsp-vst> [Accessed 1 September 2022].
72. Mandy, W., 2019. Social camouflaging in autism: Is it time to lose the mask?. *Autism*, [online] 23(8), pp.1879-1881. Available at: <https://journals.sagepub.com/doi/10.1177/1362361319878559>.
73. Matson, J., Turygin, N., Beighley, J., Rieske, R., Tureck, K. and Matson, M., 2012. Applied behavior analysis in Autism Spectrum Disorders: Recent

developments, strengths, and pitfalls. *Research in Autism Spectrum Disorders*, [online] 6(1), pp.144-150. Available at: <https://www.sciencedirect.com/science/article/pii/S1750946711000778>.

74. Max Documentation,2022. MSP Analysis Tutorial 3: Using the FFT - Max 7 Documentation. [online] Available at: https://docs.cycling74.com/max7/tutorials/14_analysischapter03 [Accessed 26 September 2022].
75. McCallum, L., 2022. MIMIC Creative AI Platform. [online] Mimicproject.com. Available at: https://mimicproject.com/about [Accessed 26 September 2022].
76. McCormack, L., Wong, S. and Campbell, L., 2022. 'If I don't Do It, I'm Out of Rhythm and I Can't Focus As Well': Positive and Negative Adult Interpretations of Therapies Aimed at 'Fixing' Their Restricted and Repetitive Behaviors in Childhood. *Journal of Autism and Developmental Disorders*, [online] Available at: <https://link.springer.com/content/pdf/10.1007/s10803-022-05644-6.pdf>.
77. Mendability, 2022. *Good Music Helps Speech Development in Autism Therapy, Studies say.*. [online] Available at: <https://www.mendability.com/autism-therapy/sensory-processing/good-music-improves-speech-therapy-for-people-with-autism-and-aspergers/> [Accessed 30 August 2022].
78. Molnar-Szakacs, I. and Heaton, P., 2012. Music: a unique window into the world of autism. *Annals of the New York Academy of Sciences*, [online] 1252(1), pp.318-324. Available at: <https://nyaspubs.onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.2012.06465.x>.
79. Moore, D. and Ferguson, M., 2014. Auditory Processing Performance and Nonsensory Factors in Children with Specific Language Impairment or Auditory Processing Disorder. *Seminars in Hearing*, [online] 35(01), pp.001-014. Available at: <https://www.researchgate.net/publication/263246253_Auditory_Processing_Performance_and_Nonsensory_Factors_in_Children_with_Specific_Language_Impairment_or_Auditory_Processing_Disorder>.
80. Moses, G., 2013. *Gabe Moses - "Stimming"*. [online] Youtube.com. Available at: <https://www.youtube.com/watch?v=MZ8TTogOtCA> [Accessed 1 September 2022].
81. Neumärker, K., 2003. Leo Kanner: His Years in Berlin, 1906-24. The Roots of Autistic Disorder. *History of Psychiatry*, [online] 14(2), pp.205-218. Available at: <https://journals.sagepub.com/doi/10.1177/0957154X030142005>.
82. Nexus, 2022. Nexus Web Audio Interfaces. [online] Available at: https://nexus-js.github.io/ui/ [Accessed 26 September 2022].
83. Oliver, A., Mossialos, E. and Robinson, R., 2004. Health technology assessment and its influence on health-care priority setting. *International Journal of Technology Assessment in Health Care*, [online] 20(1), pp.1-10. Available at: <https://www.cambridge.org/core/journals/international-journal-of-

technology-assessment-in-health-care/article/abs/health-technology-assessment-and-its-influence-on-healthcare-priority-setting/BC0FC06674538B340ADC58D2132F2520>.
84. Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A. and Kavukcuoglu, K., 2016. WaveNet: A Generative Model for Raw Audio. *arXiv preprint*, [online] Available at: <https://arxiv.org/abs/1609.03499> [Accessed 31 August 2022].
85. Orsini, M. and Smith, M., 2010. Social movements, knowledge and public policy: the case of autism activism in Canada and the US. *Critical Policy Studies*, [online] 4(1), pp.38-57. Available at: <https://www.tandfonline.com/doi/abs/10.1080/19460171003714989>.
86. Panayotov, V., Chen, G., Povey, D. and Khudanpur, S., 2015. Librispeech: an ASR corpus based on public domain audio books. *IEEE*, [online] pp.pp. 5206-5210. Available at: <https://www.danielpovey.com/files/2015_icassp_librispeech.pdf> [Accessed 1 September 2022].
87. Pellicano, E., Dinsmore, A. and Charman, T., 2014. What should autism research focus upon? Community views and priorities from the United Kingdom. *Autism*, [online] 18(7), pp.756-770. Available at: <https://journals.sagepub.com/doi/epub/10.1177/1362361314529627>.
88. Pernice, K., 2018. Affinity Diagramming: Collaboratively Sort UX Findings & Design Ideas. [online] Nielsen Norman Group. Available at: https://www.nngroup.com/articles/affinity-diagram/ [Accessed 22 September 2022].
89. Quinones, J. and Lajka, A., 2017. *"What kind of society do you want to live in?": Inside the country where Down syndrome is disappearing*. [online] Cbsnews.com. Available at: <https://www.cbsnews.com/news/down-syndrome-iceland/> [Accessed 23 August 2022].
90. Quintin, E., 2019. Music-Evoked Reward and Emotion: Relative Strengths and Response to Intervention of People With ASD. *Frontiers in Neural Circuits*, [online] 13. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6759939/>.
91. Radford, A., Met, L. and Chintala, S., 2022. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv preprint*, [online] Available at: <https://arxiv.org/pdf/1511.06434.pdf> [Accessed 1 September 2022].
92. Ragone, G., Good, J. and Howland, K., 2021. How Technology Applied to Music-Therapy and Sound-Based Activities Addresses Motor and Social Skills in Autistic Children. *Multimodal Technologies and Interaction*, [online] 5(3), p.11. Available at < https://www.mdpi.com/2414-4088/5/3/11/htm>
93. Ritvo, R., Ritvo, E., Guthrie, D., Ritvo, M., Hufnagel, D., McMahon, W., Tonge, B., Mataix-Cols, D., Jassi, A., Attwood, T. and Eloff, J., 2010. The Ritvo Autism Asperger Diagnostic Scale-Revised (RAADS-R): A Scale to Assist the Diagnosis

of Autism Spectrum Disorder in Adults: An International Validation Study. *Journal of Autism and Developmental Disorders*, [online] 41(8), pp.1076-1089. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3134766/#:~:text=Abstract-,The%20Ritvo%20Autism%20Asperger%20Diagnostic%20Scale%2DRevised%20(RAADS%2DR,201%20ASD%20and%20578%20comparisons).>.

94. Robertson, S., 2022. *What does the Thalamus do?*. [online] News-Medical.net. Available at: <https://www.news-medical.net/health/What-does-the-Thalamus-do.aspx> [Accessed 30 August 2022].
95. Robson, D., 2022. *Stimming Behaviors in Autism: A Look at Common Types & Examples*. [online] Andnextcomesl.com. Available at: <https://www.andnextcomesl.com/2021/09/stimming-behaviors.html> [Accessed 27 August 2022].
96. Roestorf, A., Bowler, D., Deserno, M., Howlin, P., Klinger, L., McConachie, H., Parr, J., Powell, P., Van Heijst, B. and Geurts, H., 2019. "Older Adults with ASD: The Consequences of Aging." Insights from a series of special interest group meetings held at the International Society for Autism Research 2016-2017. *Research in Autism Spectrum Disorders*, [online] 63, pp.3-12. Available at: <https://www.sciencedirect.com/science/article/pii/S1750946718301247?via%3Dihub>.
97. Rudy, L., 2021. Visual Thinking and Autism. [online] Verywell Health. Available at: https://www.verywellhealth.com/visual-thinking-and-autism-5119992 [Accessed 25 September 2022].
98. Sandoval-Norton, A. and Shkedy, G., 2019. How much compliance is too much compliance: Is long-term ABA therapy abuse?. *Cogent Psychology*, [online] 6(1). Available at: <https://www.tandfonline.com/doi/full/10.1080/23311908.2019.1641258>.
99. Sauro, J., 2022. Measuring Usability with the System Usability Scale (SUS) - MeasuringU. [online] Measuringu.com. Available at: https://measuringu.com/sus/ [Accessed 21 September 2022].
100. Schneider, M., 2022. *Prenatal tests for autism are coming—and so are more abortions*. [online] America Magazine. Available at: <https://www.americamagazine.org/politics-society/2022/04/29/autism-abortion-prenatal-testing-242895> [Accessed 23 August 2022].
101. Sciforce, 2020. Text-to-Speech Synthesis: an Overview. [online] Medium. Available at: https://medium.com/sciforce/text-to-speech-synthesis-an-overview-641c18fcd35f [Accessed 22 September 2022].
102. Shen, Y., Lin, Y., Liu, S., Fang, L. and Liu, G., 2019. Sustained Effect of Music Training on the Enhancement of Executive Function in Preschool Children. *Frontiers in Psychology*, [online] 10. Available at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01910/full>.
103. Shiledarbaxi, N., 2022. *Guide To Differentiable Digital Signal Processing (DDSP) Library with Python Code*. [online] Analytics India Magazine. Available

at: <https://analyticsindiamag.com/guide-to-differentiable-digital-signal-processing-ddsp-library-with-python-code/> [Accessed 31 August 2022].
104. Sibbald, S., Singer, P., Upshur, R. and Martin, D., 2009. Priority setting: what constitutes success? A conceptual framework for successful priority setting. *BMC Health Services Research*, [online] 9(1). Available at: <https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-9-43>.
105. Singh, J., Illes, J., Lazzeroni, L. and Hallmayer, J., 2009. Trends in US Autism Research Funding. *Journal of Autism and Developmental Disorders*, [online] 39(5), pp.788-795. Available at: <https://med-fom-neuroethics.sites.olt.ubc.ca/files/2016/06/Trends-in-US-autism-research-funding.pdf>.
106. Smetanin, S., 2018. Google News and Leo Tolstoy: Visualizing Word2Vec Word Embeddings with t-SNE. [online] Medium. Available at: https://towardsdatascience.com/google-news-and-leo-tolstoy-visualizing-word2vec-word-embeddings-with-t-sne-11558d8bd4d [Accessed 22 September 2022].
107. Smith, L., 2022. *Stimming: Understanding this symptom of autism*. [online] Medicalnewstoday.com. Available at: <https://www.medicalnewstoday.com/articles/319714> [Accessed 29 August 2022].
108. Stanutz, S., Wapnick, J. and Burack, J., 2012. Pitch discrimination and melodic memory in children with autism spectrum disorders. *Autism*, [online] 18(2), pp.137-147. Available at: <https://pubmed.ncbi.nlm.nih.gov/23150888/>.
109. Statistics laerd, 2022. Independent t-test in SPSS Statistics - Procedure, output and interpretation of the output using a relevant example | Laerd Statistics. [online] Available at: https://statistics.laerd.com/spss-tutorials/independent-t-test-using-spss-statistics.php [Accessed 26 September 2022].
110. TensorFlow, 2022. *Neural style transfer | TensorFlow Core*. [online] Available at: <https://www.tensorflow.org/tutorials/generative/style_transfer> [Accessed 28 August 2022].
111. Van den Oord, A., Kalchbrenner, N. and Kavukcuoglu, K., 2016. Pixel Recurrent Neural Networks. arXiv preprint, [online] Available at: https://arxiv.org/pdf/1601.06759.pdf [Accessed 22 September 2022].
112. Vines, B., Krumhansl, C., Wanderley, M. and Levitin, D., 2006. Cross-modal interactions in the perception of musical performance. *Cognition*, [online] 101(1), pp.80-113. Available at:<https://reader.elsevier.com/reader/sd/pii/S0010027705001538?token=4F7454039BEAA6D97A18988D8CDFF6A3D418E53D974BF3C5AD9769B554D7EAD6C11F4FFA0BC69F8B1969940EF0AE1461&originRegion=eu-west-1&originCreation=20220830035712>.

113. Wan, C., Bazen, L., Baars, R., Libenson, A., Zipse, L., Zuk, J., Norton, A. and Schlaug, G., 2011. Auditory-Motor Mapping Training as an Intervention to Facilitate Speech Output in Non-Verbal Children with Autism: A Proof of Concept Study. *PLoS ONE*, [online] 6(9), p.e25505. Available at: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0025505>
114. Wang, Y., 2017. Tacotron: Towards End-to-end Speech Synthesis. [online] Available at: https://arxiv.org/pdf/1703.10135.pdf [Accessed 26 September 2022].
115. Williamson, V., 2022. *Do musicians have better executive function? – Music Psychology*. [online] Musicpsychology.co.uk. Available at: <https://musicpsychology.co.uk/do-musicians-have-better-executive-function/> [Accessed 30 August 2022].
116. Yonnie Chyung, S., Kennedy, M. and Campbell, I., 2018. Evidence-Based Survey Design: The Use of Ascending or Descending Order of Likert-Type Response Options. Performance Improvement, [online] 57(9), pp.9-16. Available at: https://onlinelibrary.wiley.com/doi/epdf/10.1002/pfi.21800?saml_referrer .
117. Yuan, J., Li, M. and Lu, Z., 2021. Racial/Ethnic Disparities in the Prevalence and Trends of Autism Spectrum Disorder in US Children and Adolescents. *JAMA Network Open*, [online] 4(3), p.e210771. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7936258/>.
118. Zolyomi, A., Begel, A., Waldern, J., C Tang, J., P Barnett, M., Cutrell, E., J McDuff, D., Andrist, S. and Morris, M., 2022. Managing Stress: The Needs of Autistic Adults in Video Calling. *ACM Journals*, [online] Volume 3(Issue CSCW), pp.pp 1-29. Available at: <https://dl.acm.org/doi/abs/10.1145/3359236> [Accessed 29 August 2022].