Applications of AI-Generated Music to Music Therapy and Mental Health

Sinchana M. Nama
Mission San Jose High School
August 17, 2021

Acknowledgment
Thank you Meixing Dong from the University of Michigan for the guidance in the development of this research paper.

***Abstract* - This paper will explore how the combination of music generation and music therapy can improve one's mental health by building off the fundamentals of both fields. While music therapy has been historically based on a social science model, the field is now taking a more neuroscience-guided direction, based on music perception and brain function [1]. Different forms of biofeedback, including heart rate and skin conductivity, can measure the effect of music on our brains and mental state. Using quantifiable data representative of our emotions, machine learning models can automatically generate music designed to positively impact us. This technology can be easily applied to revolutionize the field of music therapy and aid the ongoing mental health epidemic.**

I. INTRODUCTION

Mental health is an ongoing issue that plagues nearly 13% of the world's population. These numbers are only increasing with the rise of the global pandemic, COVID-19. Common disorders include anxiety, depression, substance abuse, OCD, PTSD, bipolar disorder, schizophrenia, and eating disorders. In America itself, 20% of adults don't have access to necessary therapy, hence the need to push the possibilities of treatments. Music therapy is a common treatment method that uses music-based interventions for therapeutic purposes.

Music undoubtedly has a prominent impact on one's brain activity, behavior, and physiological responses. Different features of music, including dynamics, timbre, rhythm, register, harmony, lyrics, and tonality, affect the human brain differently. A song's rhythm affects the cerebellum, influencing one's motor skills. The lyrics affect the occipital lobe as well as the frontal lobe which is primarily responsible for emotions and cognitive skills such as problem-solving and impulse control. Lastly, tonality affects the frontal and temporal lobes which are in charge of processing memories. More importantly, different parts of the brain carry out unique functions, which are deeply influenced by listening to music. By characterizing and quantifying these musical qualities and their emotional effects on a person, machine learning can be used to predict the emotional and physiological response a person will have to a certain piece.

In this paper, I will set the foundations of both music therapy and music generation and then develop the intersection between the fields. Section II will introduce current music therapy methodologies and how they have been used to aid different mental health disorders. Section III will delve into the cognitive neuroscience of music and the benefits of receptive music therapy. In Section IV, I will provide a gross introduction to music generation and common generative AI algorithms and models. I will also elaborate on a few examples of music generation and their use cases and explain how one's emotional state can be quantified. After developing how music therapy is used to benefit the aforementioned mental disorders and what different models have contributed to music generation, I will discuss possible combinations of the two fields. In section V, Future Applications, I will reference music generation studies that have used various forms of biofeedback to bridge the gap between the fields. Lastly, I will propose new applications of these studies pertaining to improving mental health in today's world.

II. MUSIC THERAPY

Music therapy is a medical practice that utilizes music to aid the treatment of a variety of mental ailments. Music therapy capitalizes on the emotional, mental, physical, and social aspects of music to better one's health [2]. Music therapists study their clients' emotional responses to different types of music to curate what's best for them. Activities involving singing, songwriting, listening to music, and playing instruments lead to better motor skills, cognitive function, behavior, and social skills and improve emotional development. These practices are commonly used in medical institutions such as cancer centers, rehabilitation centers, psychiatric hospitals, and education centers.

Music therapy treatments fall under active and receptive modes. Active techniques include improvisation, recreation, and composition. Re-creation techniques involve singing or playing instruments and are centered around strengthening motor skills, social interaction, and self-expression. Improvisation refers to spontaneous music creation, and composition involves coming up with lyrics and melodies. These two interventions help clients communicate through music and express their thoughts and emotions. While this therapist-client interaction plays a vital role in improving one's communication and behavioral skills, this paper will focus on the direct effect of listening to music on our brains. Receptive therapy involves listening to music for mood changes and relaxation and has a more direct effect on one's mental health. Selecting music for receptive therapy depends on several factors including age, therapeutic context, and preferred music. For example, music designed for relaxation usually has a slow and consistent tempo, predictable melodic, harmonic, and rhythmic features, few dynamic changes, and more strings and woodwinds. Western classical, Celtic, new-age (nature-scapes enhanced by electronic sounds), jazz, and meditative trance music are commonly used in receptive music therapy. However, user input plays a great role in determining which genres will work the best for them [3].

Music therapy currently acts as a supplemental form of therapy and hasn't been proven to single-handedly heal mental disorders the way other forms of therapy have. However, ongoing neurobiological research, regarding brain function and music perception, is changing the future direction of this field by setting the foundation for how music changes both brain and behavioral functions [1]. This shift to a more neuroscientific basis may allow music therapy to play a more central role in mental treatments in the future.

A. MUSIC THERAPY APPLICATIONS

Music therapy is commonly used to treat several mental ailments including depression, anxiety, autism, and even Alzheimer's. Depression and anxiety are often found together and co-exist in patients. Participants of a study, diagnosed with depression, were either provided standard care plus 20 music therapy sessions or standard care on its own. Those who underwent music therapy along with standard care showed significantly better results in terms of depression (Montgomery-Asberg Depression Rating Scale - MADRS), anxiety (Hospital Anxiety and Depression Scale - HADS-A), general function (Global Assessment of Functioning - GAF), and quality of life (health-related survey RAND-36) [4].

Music therapy has also proven to be beneficial to children with autism (ASD). Music therapy showed short-term improvement in several dimensions of social interaction and communication, including non-verbal and verbal communicative skills, initiating behavior, social-emotional reciprocity, social adaptation, and relationships. Music therapy techniques used here included improvisation, vocalization, and listening to music [5].

For the elderly facing dementia, music therapy has shown cognitive, physical, and psychological benefits. While research for curing Alzheimer's, the number one cause of dementia, is most commonly pharmacological, the drugs used often worsen motor skills and result in premature death. Non-pharmacological treatments, such as music therapy, provide a safer and more effective alternative. Alzheimer's patients that underwent music therapy had increased MMSE (*Mini-Mental State Examination*), NPI (*Neuropsychiatric Inventory)*, and HADS (*Hospital Anxiety and Depression Scale*) scores. MMSE scores evaluate orientation, memory, attention, and language-motor skills. NPI scores focus on behavioral functions including delusions, hallucinations, depression, agitation, irritability, aberrant motor behavior, anxiety, aggressiveness, apathy, and disinhibition. Music therapy has proven to be a safe and inexpensive way to combat these while improving mood, cognitive function, and behavior in dementia patients [6].

III. MUSIC AND THE BRAIN

By better understanding research demonstrating the connection between music and brain function, we substantiate our viewpoint that music therapy has great potential for further mental health applications. As mentioned in the introduction, different auditory aspects of music have neural correlations to different parts of the human brain. Through the use of PET and fMRI scans, studies have identified how pitch, harmony, timbre, and rhythm affect the brain [7]. Pitch processing affects multiple regions of the brain. For example, spectral and temporal sound features are encoded in the brain stem, while a conscious perception of pitch lies in the auditory cortex. Long-time-scale pitch patterns are processed through the brain's larger networks. Right and left auditory temporal areas play a large role in processing timbre. Timbre is a musical characteristic separate from pitch and intensity, such as the difference in sound between a violin and an orchestral bass. Rhythm processing includes cognitive, sensory, and motor characteristics and occurs in a distributed cortical and subcortical network. The cerebellum is also thought to affect the temporal organization of cognitive and perceptual processes.

The Mozart Effect is a long-standing theory suggesting that listening to Mozart can improve one's IQ. While it has since been proven that people with dementia and Alzheimers show better results with music they grew up with, the elements of Mozart's music that yield positive results for some still hold value. The mode of a song and its tempo can drastically change how the music makes one feel. To test this, individuals were asked to perform the paper-folding-and-cutting task (PF&C) once after listening to music and once after being in complete silence [8]. Half of the participants were asked to listen to Mozart, which is in a relatively fast-tempo and major key, while the other half listened to Albinoni's Adagio, which has a relatively slow tempo and is in a minor key. Mood and arousal levels were also measured before and after the participants listened to the music. As predicted, those who listened to Mozart had higher arousal levels and were happier compared to when they sat in silence, while those who listened to Albinoni showed no difference between listening to music and silence. In addition, those who listened to Mozart performed better on the PF&C task.

A study conducted by the Osaka University Graduate School of Medicine in 2008 used salivary cortisol levels as an endocrinological stress marker to measure which musical keys reduced the most stress in the human subject [9]. Major modes were found to decrease stressful conditions more than minor modes. Based on the patterns of stress responses found in the upper temporal cortex areas, scientists believe there is a correlation between music-induced stress reduction and how we process emotions such as happiness and sadness. These studies show why some songs have a positive effect on listeners and why listening to such music in the context of music therapy is effective.

Another study used rats to identify the effects of listening to music on the release of brain monoamines, which include dopamine and serotonin (feel-good hormones). Nigrostriatal and mesocorticolimbic pathways are the parts of the dopaminergic system that relate to reward mechanisms, motor control, and emotion-driven behavior. Dopamine and serotonin levels were measured in the caudate-putamen (CPU) and nucleus accumbens (NAcc), areas of the brain related to reward and motor control. The experimental group of rats exposed to music showed an increase in dopamine levels and the release of serotonin in the CPU, as well as a dopamine turnover in the NAcc, confirming the direct effect music has on our monoamine activity [10].

In addition, musicians have more gray matter in the part of the frontal cortex that accommodates neural networks involved in several memory processes. Overall, music has been shown to decrease negative irregularities such as anxiety, low emotions, and pain, as well as improve memory, alertness, motivation, and mood [11].

IV. MUSIC GENERATION

Automated music generation uses machine learning to simulate a musician's creativity by analyzing existing musical compositions. Before deep learning models were available, statistics and probability were used to generate music. This was known as Stochastic music, developed by Iannis Xenakis, in the early 1950s. He used stochastic theory to generate new music elements, purely dependent on such mathematical models [12].

Today, there are several deep-learning algorithms used to generate music. Deep learning is a form of statistical learning that utilizes neural networks with large amounts of hidden layers. A basic neural network consists of nodes (or neurons) organized into several layers. These layers are connected by a set of weights, which represent the strength of the connections formed. Lastly, activation functions, such as Sigmoid, Tanh, and ReLU, are applied to transform the output of every node in a layer.

In the realm of music generation, a piece of music can be represented as a sequence of events. These events can be notes, chords, or other musical elements. Pieces within classes are assigned higher probabilities, while other pieces have lower probabilities. Classes can be defined by genre, style, or composer. Generated music samples elements or pieces with high probabilities as dictated by the model. The music generation examples explored below simply serve as an insight into ways the following models have been applied to music generation and don't directly tie back to music therapy.

A. Hidden Markov Models

Hidden Markov models serve as the building blocks for computational sequence analysis and probabilistic models and are used to solve linear sequence 'labeling' problems. This involves biological sequencing, speech recognition, and even music generation. An HMM model invokes a state for each label. Each state has its own probabilities as well as transition probabilities which indicate the chance of moving from one state to another. An HMM model generates both the underlying state path as well as the observed sequence. The state path is a hidden Markov chain; a model in which the next state depends on the current state. Since HMM is a probabilistic model, Bayesian classification frameworks and theories are applicable [13].

HMM models have been used to generate classical music based on the originality, harmonic qualities, and temporal structure of the Romantic era pieces [14].

B. Recurrent Neural Networks

Recurrent neural networks, RNNs, use sequential data to produce outputs dependent on the previous elements of the sequence. In between the input and output layers of an RNN, there lie several layers. Each layer in a recurrent neural network shares the same weight parameter, unlike feedforward networks, which have different weights throughout. RNNs use the backpropagation through time algorithm (BPTT) to calculate the error gradients from the output to the input layer. Since parameters are shared across layers, BPTT sums the error at each time step. This gradient allows the model to adjust its parameters accordingly and "learn."

1. LSTM

Long Short-Term Memory networks are a type of RNN that are skilled in encoding long-term patterns. In an LSTM, information flows through cell states, which allow LSTMs to selectively remember or disregard information depending on how important it is. Previous information is stored within an LSTM cell. LSTM models have been used to create probabilistic models of sequences of polyphonic music [15].

C. Transformer Neural Networks

Transformer Neural Networks are faster and more efficient than other neural networks because all elements of their input sequence can be passed simultaneously into the model. A default transformer model has 6 encoders and decoders. An encoder encodes the input sequence and the decoder decodes the input into the next sequence. Imposing a metrical structure onto the input music can help transformers identify the beat-bar-phrase hierarchical structures in music and compose music with better rhythmic and harmonic structures, as used by the Pop Music Transformer [16]. Transformers based on self-attention have been used to elaborate on given motifs and generate accompaniments based on the given melody [17].

D. QUANTIFYING EMOTIONAL STATE

The most common way to quantify emotions is a 2-dimensional emotional plane, with arousal and valence as the two variables, and was used by several of the studies mentioned in this paper. Valence (horizontal axis) is a measure of positivity while arousal (vertical axis) is a measure of activation strength. Emotions, such as calmness, happiness, sadness, and anger are mapped onto this space. For example, anger has a negative valence and high arousal. While this method of mapping can be problematic or ambiguous at times, this numerical data is necessary for machine-learning models to generate music based on human input and impact.

V. FUTURE APPLICATIONS

We have established the benefits of listening to music and other music-related activities, furthered by the field of music therapy. However effective, music therapy isn't easily accessible due to cost, resources, and sufficient education. The biggest limitation when it comes to listening to music for medical purposes is the lack of effective music. By using machine learning, music therapy can be made much cheaper and more accessible. The aforementioned machine learning algorithms allow us to generate music specifically designed to help an individual. This can be influenced by the genres and sounds the patient grew up listening to, as listening to one's favorite music activates different parts of the brain. Using biological markers, including heart rate, MRI scans, saliva, and the galvanic skin response (GSR), we can estimate the emotional state of a person in response to listening to a certain piece of music. These signals can be used in machine learning algorithms to generate music designed to ignite a specific emotion in the recipient.

Scientists were able to generate emotionally congruent music through Hidden Markov Models [18] by evaluating skin conductivity. Shimmer3 wireless GSR+ Unit1, credited for biomedical-oriented research, is used to analyze GSR (galvanic skin response), the change in the balance of positive and negative ions in sweat secretion [19]. These scientists found a direct correlation between the type of piece (calm/scary), the participant's GSR reading, and the emotions they described in a questionnaire. This data allowed them to estimate an emotional state and the control signal, and generate new musical structures according to the musical feature similarity model. This emotionally congruent music can be beneficial to the field of music therapy as the generated music will have a higher probability of inciting the desired emotion in the patient.

Similarly, recent studies have shown a direct correlation between heart rate and mental health, making heart rate another viable indicator [20]. Long-term depression, anxiety, and stress disorders may have physiological effects, including increased heart rate, blood pressure, and cortisol levels. Heart rate variability biofeedback (HRVB) provides real-time electronic feedback monitoring one's heart rate variability (HRV). Higher HRVs indicate healthier physical and mental states, while lower HRVs represent many illnesses, ranging from asthma to depression. HRVB has proven to be of great therapeutic benefit for a wide range of medical and mental disorders, specifically, those related to autonomic nervous system dysregulation [21]. Using a heartbeat sensor, it is possible to generate music corresponding to specific emotional states. Algorithmic composition using a generation model described in Section IV can then use a music-feature mapping method to compose the necessary music. While these experiments have been done, they haven't been applied to music therapy and mental health. Using the emotional plane and feature mapping methods, music can be generated to change accordingly to help lower the patient's heart rate. The average depressed person's heart rate is about 10-15 BPM above normal. The target heart rate may vary due to age, mental and physical health, and other environmental factors. Once the target heart rate is attained, the machine learning model should learn to maintain it at that level. Certain melodies, instruments, and beats may have a different impact depending on the person. Therefore, the machine learning model will be able to design a personalized tune, specifically designed to help the patient.

Leaning away from biofeedback, Twitter posts and chatbots can also serve as an indicator of one's mental state. Natural language processing used in chatbots has been used to diagnose and treat mental illness [22]. This data can be converted into the same emotional plane and be used to generate music the person would enjoy and benefit from. Using direct user input, musical preferences, such as favorite genre, instrument, and style, can be manually inputted into the algorithm, making the process less ambiguous and randomized. The untapped potential and possibilities at the intersection of these seemingly unrelated fields are endless.

VI. CONCLUSION

Artificial intelligence has applications in nearly every field: whether it's medical, environmental, political, or even entertainment. In this paper, I was able to direct the combination of AI and music from simple entertainment to a practical application in the realm of therapeutic treatment. Due to the specific neurological effect music has on our brains, music therapy is growing to be a more influential and credible form of treatment for a variety of mental illnesses. We can quantify music's effects on the brain with several metrics, which can then be inputted into a variety of machine-learning algorithms to generate music designed to help those in need. With mental health issues on the rise, automatically generated music has practical applications and implementations across the field of therapy.

References

[1] Thaut, M. H. (2005). The Future of Music in Therapy and Medicine. *Annals of the New York Academy of Sciences, 1060*(1), 303-308. doi:10.1196/annals.1360.023

[2] Christensen, E. (2012). *Music Listening, Music Therapy, Phenomenology and Neuroscience*.

[3] Grocke, D. E., & Wigram, T. (2007). *Receptive methods in music therapy: Techniques and clinical applications for music therapy clinicians, educators, and students*. Jessica Kingsley.

[4] Erkkilä, J., Punkanen, M., Fachner, J., Ala-Ruona, E., Pöntiö, I., Tervaniemi, M., . . . Gold, C. (2011). Individual music therapy for depression: Randomised controlled trial. *British Journal of Psychiatry, 199*(2), 132-139. doi:10.1192/bjp.bp.110.085431

[5] Geretsegger, M., Elefant, C., Mössler, K. A., & Gold, C. (2014). Music therapy for people with autism spectrum disorder. *Cochrane Database of Systematic Reviews*. doi:10.1002/14651858.cd004381.pub3

[6] Gallego, M. G., & García, J. G. (2017). Music therapy and Alzheimer's disease: Cognitive, psychological, and behavioural effects. *Neurología (English Edition), 32*(5), 300-308. doi:10.1016/j.nrleng.2015.12.001

[7] Levitin, D. J., & Tirovolas, A. K. (2009). Current Advances in the Cognitive Neuroscience of Music. *Annals of the New York Academy of Sciences, 1156*(1), 211-231. doi:10.1111/j.1749-6632.2009.04417.x

[8] Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, Mood, and the Mozart Effect. *Psychological Science*, *12*(3), 248–251. http://www.jstor.org/stable/40063588

[9] Suda, M., Morimoto, K., Obata, A., Koizumi, H., & Maki, A. (2008). Emotional responses to music: Towards scientific perspectives on music therapy. *NeuroReport, 19*(1), 75-78. doi:10.1097/wnr.0b013e3282f3476f

[10] Moraes, M. M., Rabelo, P. C., Pinto, V. A., Pires, W., Wanner, S. P., Szawka, R. E., & Soares, D. D. (2018). Auditory stimulation by exposure to melodic music increases dopamine and serotonin activities in rat forebrain areas linked to reward and motor control. *Neuroscience Letters, 673*, 73-78. doi:10.1016/j.neulet.2018.02.058

[11] Jäncke, L. (2008). Music, memory and emotion. *Journal of Biology, 7*(6), 21. doi:10.1186/jbiol82

[12] Whorley, R. P., & Conklin, D. (2016). Music Generation from Statistical Models of Harmony. *Journal of New Music Research, 45*(2), 160-183. doi:10.1080/09298215.2016.1173708

[13] Eddy, S. R. (2004). What is a hidden Markov model? *Nature Biotechnology, 22*(10), 1315-1316. doi:10.1038/nbt1004-1315

[14] Yanchenko, A. (2017). *Classical music composition using hidden Markov models* (Doctoral dissertation, Duke University)

[15] Nayebi, A., & Vitelli, M. (2015). Gruv: Algorithmic music generation using recurrent neural networks. *Course CS224D: Deep Learning for Natural Language Processing (Stanford)*.

[16] Huang, Y., & Yang, Y. (2020). Pop Music Transformer. *Proceedings of the 28th ACM International Conference on Multimedia*. doi:10.1145/3394171.3413671

[17] Huang, C. Z. A., Vaswani, A., Uszkoreit, J., Shazeer, N., Simon, I., Hawthorne, C., ... & Eck, D. (2018). Music transformer. *arXiv preprint arXiv:1809.04281*.

[18] Monteith, K., Martinez, T. R., & Ventura, D. (2010, January). Automatic Generation of Music for Inducing Emotive Response. In *ICCC* (pp. 140-149).

[19] Williams, D., Hodge, V., Gega, L., Murphy, D., Cowling, P., & Drachen, A. (2019, March). AI and automatic music generation for mindfulness. In *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio*. Audio Engineering Society.

[20] Lehrer, P. M., & Gevirtz, R. (2014). Heart rate variability biofeedback: How and why does it work? *Frontiers in Psychology, 5*. doi:10.3389/fpsyg.2014.00756

[21] Moss, D., & Shaffer, F. (2017). The Application of Heart Rate Variability Biofeedback to Medical and Mental Health Disorders. *Biofeedback, 45*(1), 2-8. doi:10.5298/1081-5937-45.1.03

[22] Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. *The Canadian Journal of Psychiatry, 64*(7), 456-464. doi:10.1177/0706743719828977