Predicting sensorimotor responses to rhythmic events : an EEG study of neuronal entrainment & beat perception (cognizant through harmonics)

Darmalingam, Kumaresh

2021

Darmalingam, K. (2021). Predicting sensorimotor responses to rhythmic events : an EEG study of neuronal entrainment & beat perception (cognizant through harmonics). Master's thesis, Nanyang Technological University, Singapore. <https://hdl.handle.net/10356/151269>

<https://hdl.handle.net/10356/151269>

<https://doi.org/10.32657/10356/151269>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 18:33:28 SGT

OHM

**PREDICTING SENSORIMOTOR RESPONSES TO
RHYTHMIC EVENTS: AN EEG STUDY OF NEURONAL
ENTRAINMENT & BEAT PERCEPTION
(COGNIZANT THROUGH HARMONICS)**

KUMARESH DARMALINGAM

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

2021

**PREDICTING SENSORIMOTOR RESPONSES TO
RHYTHMIC EVENTS: AN EEG STUDY OF NEURONAL
ENTRAINMENT & BEAT PERCEPTION
(COGNIZANT THROUGH HARMONICS)**

KUMARESH DARMALINGAM

School of Computer Science and Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Master of Engineering

2021

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

20/05/2021

.....

Date

.....

Kumaresh

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

24 May 2021

.....

Date

.....

Prof. Guan Cuntai

Authorship Attribution Statement

Please select one of the following; *delete as appropriate:

*(A) This thesis **does not** contain any materials from papers published in peer-reviewed journals or from papers accepted at conferences in which I am listed as an author.

*(B) This thesis contains material from 0 paper(s) published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.

20/05/2021

.....

Date

.....

Kumaresh

Acknowledgement

First and foremost, I wish to express my gratitude to my supervisor Prof. Guan Cuntai. Whose friendly, supportive, and nurturing nature made it possible to traverse the complexity of my candidature and scientific work with the challenges of life. It was a great opportunity to be supervised by Prof. Guan, one which I wish I could continue and fully appreciate moving forward.

I would also like to thank my co-supervisor Prof. Kat Agres, for first giving me the opportunity to be part of her Music Cognition team. Whose encouragement and mentorship guided me into the beginnings of my experience into the world of Music Cognition. For which I am eternally thankful.

I would also like to thank my co-supervisor Dr. Steffen Herff, who was there at every step of this journey, in my learning and development. Whose guidance is greatly appreciated. Whose immense efforts in making true of possibilities will always be treasured.

Table of Contents

Table of Contents	vi
List of Figures	ix
List of Tables	xii
Summary	xiii
Chapter 1. Background	1
Introduction.....	1
Research Gap	2
Research Objective	3
Research Question	3
Research Significance	3
Chapter 2. Literature Review	5
Evolutionary Traits	5
Neuroscientific Functioning.....	9
Temporal Perception.....	11
Rhythm Cognition.....	11
Neural Resonance Theory.....	13
Neural Entrainment.....	14
Beat Perception	15
Dynamic Attending Theory	18
Sensorimotor Synchronization.....	19
Frequency Domain Analysis.....	21
In Summary.....	24
Chapter 3. Dataset & Experiment	25
Dataset.....	25
Protocol.....	27
Analysis Review	28

Chapter 4. Study Methodology	31
Tools	31
Methodology	31
Analysing in the Frequency Domain	32
Tapping Data.....	33
Song Data.....	34
EEG Data	34
Predictive Modelling.....	35
Multinomial Logistic Regression.....	35
Range of Observance – On Tapping.....	35
Range of Observance – On EEG	36
Beat Frequency and Harmonics	36
Normalising on Harmonics	37
Leave-n-Out Cross-Validation (LnO CV)	38
Chapter 5. Results & Analysis	40
Tapping Data.....	40
Time Series Tapping Signal.....	40
Tapping Frequency Data.....	41
Tapping Frequency: FFT vs Median.....	42
Tapping Distribution.....	46
Music Information Retrieval - Beat Frequency	48
Song vs Stimuli Influences on Beat vs Metre	49
Beat Extraction from Songs	54
Harmonizing Frequencies	55
EEG Analysis.....	60
Predictive Modelling.....	61
Predictive Modelling Profile.....	62

Best Performing Models	64
Chapter 6. Validation	66
Validity	66
Structural Assumption	66
Data Assumption.....	67
Tapping Responses	67
EEG Responses	70
Chapter 7. Conclusion & Future Work	74
Conclusion	74
Limitations	75
Future Research	76
Epilogue	77
References	79

List of Figures

Figure 1: Summary diagram of tracts in the human dorsal auditory stream (Patel, 2014)	6
Figure 2: Cognitive circuitry involved in sequential and temporal processing (Merchant & Honing, 2014)	7
Figure 3: Animals displaying their capability of audiomotor entrainment, by head bobbing to musical stimuli (Kotz, Ravignani, & Fitch, 2018)	8
Figure 4: Coupling between auditory and premotor cortices (Zatorre, Chen, & Penhune, 2007)	10
Figure 5: The interconnectivity addressing the breakdown from acoustic signals to cyclic patterns to beats and onto metrical representation (Kotz, Ravignani, & Fitch, 2018)	13
Figure 6: Activations in highlighted areas as observed by fMRI studies (Merchant, Grahn, Trainor, Rohrmeier, & Fitch, 2015)	16
Figure 7: Spectrogram of input stream (top) used to observe tapping responses of listeners (bottom) (Patel & Iversen, 2014)	17
Figure 8: A 1-Hz sinusoid wave viewed in the temporal and frequency spectrum (Zhou, Melloni, Poeppel, & Ding, 2016)	22
Figure 9: A 5-Hz sinusoid wave viewed in the temporal and frequency spectrum. The frequency spectrum shows multiple signals at both the fundamental frequency and its harmonics (Zhou, Melloni, Poeppel, & Ding, 2016)	23
Figure 10: NMED-T Research Paradigm	28
Figure 11: Amplitude envelopes and beat hierarchies found in the stimuli (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017)	29
Figure 12: Analysis of tapping responses and stimuli tempo (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017)	30
Figure 13: Overall Strategy Adopted	31
Figure 14. (a) Left: Tapping timings. (b) Conversion of tapping timings into a time series binary signal. (c) FFT of tapping data	34
Figure 15. Conversion of EEG data into the frequency domain for participant 1, song 1, and electrode 100	35
Figure 16: Extracting Features from EEG Frequency Spectrum Data.	37
Figure 17: Centering of EEG features based on harmonics	38
Figure 18: Leave One Out Cross Validation process flow	39
Figure 19: Tapping data stream plotted for a subset of the tapping dataset	41

Figure 20: Tapping frequency signal plotted for a subset of the tapping dataset	42
Figure 21: Tapping frequency signal with median frequency (red dotted line) plotted for a subset of the tapping dataset	43
Figure 22. Identifying tapping frequency for all participants and song 1. (Left) Identifying using FFT peak extraction. (Right) Identifying using median tapping frequency.	44
Figure 23. Comparing the extraction methodologies of beat frequency among participants with stark difference. The frequency computed using the median methodology is highlighted with a red dashed line. The strength of the FFT process, can be observed.	45
Figure 24. Peak extraction for participant 20 song 1. Highlighting the shortcoming of the FFT process.....	46
Figure 25. Participant tapping distribution among songs	48
Figure 26. Similarities and differences in beat frequency extractions from songs & stimuli..	50
Figure 27. Metre frequency of songs & stimuli	51
Figure 28. Beat and metre frequency of songs	52
Figure 29. Beat and metre frequency of stimuli	53
Figure 30. Comparing tapping distribution against extracted frequency using beat and metre	54
Figure 31: Flow of computing the fundamental frequency and its harmonics	56
Figure 32: Frequency distribution of tapping responses of all participants across all songs...	57
Figure 33: Frequency distribution of tapping responses of all participants across all songs...	58
Figure 34: Proportion of tapping responses aligning with a harmonic	59
Figure 35: Extracting features based on the beat frequency and its harmonics.....	60
Figure 36: EEG responses at beat frequency for Song 1	61
Figure 37: Process flow depicting the modelling effort.....	62
Figure 38: Mean accuracy profile by participant - left view	63
Figure 39: Mean accuracy profile by participant - right view	63
Figure 40: Mean accuracy profile by participant - top view.....	64
Figure 41: 3D topography of electrodes - side view	65
Figure 42: 3D topography of electrodes - top view	65
Figure 43: Mean accuracy profile by participant (Randomized tapping) - left view.....	68
Figure 44: Mean accuracy profile by participant (Randomized tapping) - right view	69
Figure 45: Mean accuracy profile by participant (Randomized tapping) - top view.....	69
Figure 46: Mean accuracy profile by participant (Randomized EEG) - left view.....	71
Figure 47: Mean accuracy profile by participant (Randomized EEG) - right view.....	72

Figure 48: Mean accuracy profile by participant (Randomized EEG) - top view	72
Figure 49: Mean accuracy profile by participant (Randomized EEG) – bottom left view	72
Figure 50: Schematic depiction of audio stimuli. Left: The Auditory Soundwave. Right: The music notation (Grahn J. A., 2009)	77

List of Tables

Table 1: The 10 songs used for stimuli (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017).	26
Table 2: Listing of files in the NMED-T dataset (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017).....	27
Table 3: Accuracy performance of the best models.....	64
Table 4: Performance of actual best model (Randomized tapping).....	70
Table 5: Performance of actual best model (Randomized EEG).....	73

Summary

Motivation: To embark on implementable methodologies which can inspire practical solutions. The phenomenon of musical beat perception and neural entrainment has been studied substantially in the field with great understanding in a multitude of scenarios. Building models which rely on an individual's perception to musical beat and its influence on neural entrainment demonstrates the applicability of a dynamic real-time feedback system. Such systems could lead the way in assistive technologies used to facilitate recovery and cognitive processing in psychotherapy, physiotherapy, physical training, or even speech therapy in children. However, most studies only consider beat frequency, rather than neural harmonics, which could carry valuable predictive information.

Objectives: This research endeavours to explore the phenomenon of neural entrainment which is a key neurological function that plays a critical role in beat perception among listeners of music and other rhythmic stimuli. Specifically, this project aims to build predictive models of sensorimotor responses to music stimuli, such as induced rhythmic tapping behaviour, using Electroencephalogram (EEG) measurements at beat frequency as well as its harmonics as predictors.

Methods: A publicly available dataset, called NMED-T of EEG recordings and tapping timings of 20 listeners over 10 songs is used to develop a predictive model. The predictive focus of the model is to identify the frequency at which a listener will tap to a given music. The model is driven by the theory of neural entrainment, beat perception, and Music Information Retrieval (MIR) methodologies for extracting beats from audio sources.

Findings: A multinomial logistic regression model was able to predict the frequency at which a listener will tap to using EEG measurements, with a precision of 0.74 ($s=0.19$). This suggests that tapping behaviour can be reliably predicted using neural responses. Importantly, findings also highlight that the most informative features were extracted from EEG signals at the beat frequency and its harmonics, rather than the perceived beat frequency itself. This suggests that research investigating neural entrainment or Brain-Computer-Interfaces that aims to use entrainment should consider neural harmonics too, rather than beat frequency alone. EEG surface maps further support this by showing that neural entrainment consistently occurs within the sub-harmonics range across participants and songs.

Chapter 1. Background

Introduction

Music is an artform which is appreciated through our sensory processes and relied upon in our communicative endeavours. Intonations within communicative speech possesses musical undertones, which are both universally distinguishable and uniquely represented in diverse forms across cultures. Despite these cultural variations, perception of musical beats, metres and rhythms are still universal across cultures and play an important role in social coherence. For example, it has been shown that rhythmic synchronization or entrainment among humans, either in small pairs or larger groups, have a significantly positive impact at a societal level. Such as the joy and harmony experienced in societal celebrations like births, weddings, and successes where costumes, dancing or chanting are involved. Where individuals through these activities merge within a communal group in a cohesive manner. This phenomenon is described as “collective effervescence”. Synchronized movements like drumming and dancing is considered a common and influential method to achieve “communitas”. These cultural innovations are speculated to be rooted in evolutionary needs. Societies with heritable tendencies who partake in such communal activities benefit from stronger social bonds and higher offspring compared to those who do not practise such activities (McNeill, 1995; Haidt, Seder, & Kesebir, 2008; Durkheim, 1965; Turner, 1995; Eliade, 1964; Ehrenreich, 2006).

When listening to music, tempo is spontaneously and uniquely processed by the cognitive system. Listeners can mentally represent a rhythmic cycle of a stimuli with a steady isochronous pulse and tease apart the various accents in a metric representation (Herff, et al., 2020). For example, strong accenting beats have shown greater discrimination in cognitive processing generating greater attention around cyclic pitches and notes, conforming to regularities (Fujioka, Trainor, Large, & Ross, 2009; Jones M. R., 1987). When analysing the music’s acoustic signals along with the EEG signals, there is an alignment along the frequency domain where the tempo of the musical stimuli is observed (Nozaradan, 2014).

This shows that the cognitive system is highly sensitive to tempo signals, considering that it is embedded within complex acoustic signals. Listeners could theoretically adopt various temporal interpretation based of these auditory signals, yet listener’s generally focus on a constant isochronous pulse at frequencies around 0.5Hz to 5Hz, preferred tapping frequencies

tend to be between 0.8Hz to 2.5Hz (Fujioka, Trainor, Large, & Ross, 2009; Repp B. H., 2005; Patel & Iversen, 2014; London J. , 2012).

This thesis details the efforts in answering questions relating to the cognitive processing of beat perceptions and neural entrainment, by building a predictive model that utilises both neural entrainment at beat frequency, as well as beat frequency harmonics as predictors.

This study is based on the understanding that stimuli, either visual or auditory, presented constantly at a frequency will evoke brain responses mirroring those exact periodicities of the given stimuli. These constant periodic responses are termed Steady State Evoked Potential (SS-EPs), which showcases the phenomenon of neural entrainment (Prado-Gutiérrez, et al., 2019).

In studying these periodicities, the most direct form would be to study them in the frequency domain. Though it has been clearly illustrated that neural responses cycling at a frequency do not always contribute as a response on the Fourier spectrum (Zhou, Melloni, Poeppel, & Ding, 2016). Critically however, neural entrainment results in strong neural responses on the frequency domain that correlates with the stimulus rhythm on the temporal scale (Zhou, Melloni, Poeppel, & Ding, 2016). This is the foundation of the approach taken here.

The NMED-T dataset from Stanford provided an extremely useful dataset to pursue evaluations of questions relating to beat perception supported by EEG data analytics. Pertinently, it offered behavioural data in the form of tapping responses to music stimuli that can be referenced and benchmarked against the EEG measurements (Stanford University, 2017).

Research Gap

Until now, research literature has not thoroughly explored building predictive models of the phenomenon of neural entrainment to complex music stimuli. Current explorations also primarily focus on fundamental frequencies like the beat or meter. This study extends explorations not only to those fundamental frequencies but also its resulting harmonics. Therefore, translating beat and harmonic entrainment of neurons from complex sources, into understanding or predicting actual behavioural responses is unexplored.

NMED-T dataset provides the opportunity to observe and study the naturalistic cognitive processing of commercially available music against instinctive sensorimotor responses from listening to those complex auditory signals.

Research Objective

This research takes the opportunity to further understand beat perception and behavioural responses, and to extend the understanding of the role neural entrainment plays in beat perception and resulting sensorimotor responses. Opportunities are also taken to expand on how these interactions with MIR methodologies to explore fundamental frequencies and their relationship with harmonics furthers our understanding. The objective is to build a predictive model that identifies the frequency at which listeners replicate the beat frequency of complex auditory signals in terms of behavioural responses such as tapping, using neural responses as predictors.

Research Question

The hypothesis is that, as listeners neurally entrain at different frequencies to the same musical stimuli, EEG responses from listening to these stimuli will provide the fundamental data measurements found within the frequencies of the beat and its harmonics for predicting resulting tapping behaviours.

What is the role beat harmonics plays in informing predictions of tapping frequency, and can predictive models be built based on EEG measurements informed by beat frequencies and its harmonics?

Research Significance

The significance of this research will be the evidence that even with complex music stimuli, the cognitive system is able to tease apart beat information from all other acoustic details and these perceptions play a significant role in physiological manifestations induced from motor cortex processing.

If these complex processing can be manipulated to affect behavioural responses, it can be a significant gateway to multiple applications such as in physiological movement like the training of athletics relaying on synchronised or cyclic movements, or in physiotherapy training among geriatrics. It can also be used as an aid in assistive technologies that are aimed towards psychotherapy by regulating emotional states.

Critically, recent developments in EEG technology allows mobile data collection using single electrodes, enabling both researchers to collect data in ecologically valid scenarios and non-researchers to develop user-focused tools, such as BCIs for music composition. However, for non-experts the initial question of electrode placement and frequency selection provides a substantial entry barrier. The present project aimed at providing a ‘rule of thumb’ location and frequency bin, that can inform the placement of the single electrode for such mobile devices. Consequently, the methodology was chosen to address this problem, and simulates the scenario envisioned by our research motivation, by testing each electrode at an informed frequency band by literature across participants.

Chapter 2. Literature Review

The topics covered in the literature review will not only focus on elements directly applied and implemented in this work, but also on the fundamental knowledge gained in current literature that speaks to the reasoning and development of approaches adopted. Understanding how both the physiological and psychological components are linked and influence the approaches or methodologies adopted to tackle the questions being explored, helped to both innovate and validate the effort.

Evolutionary Traits

“The perception, if not the enjoyment, of musical cadences and of rhythm is probably common to all animals, and no doubt depends on the common physiological nature of their nervous systems” - (Darwin, 1871).

Recent modern works advocating the theory of “neural resonance”, builds upon Darwin’s observations. These studies draw the relationship between beat perception, neural entrainment and its underlying oscillatory signals holding temporal patterns. Signals and patterns described as pulses or metric (Patel & Iversen, 2014; Hulse, Bernard, & Braaten, 1995; Large & Jones, 1999; Large & Snyder, 2009; Jones & Boltz, 1989; Schroeder & Lakatos, 2009; Buzsáki, 2009). (Buzsáki, 2009) discusses on how neuronal activities are organised by oscillatory signals throughout the cognitive system, which are driven by the co-evolution of the brain structure and its cognitive functions.

Neural signals that plays a part in the auditory system to process and predict beats, rely on the dorsal auditory pathway connecting to the motor control system via the parietal cortex. This pathway connection is suggested to be stronger in humans because of the evolution of vocal training (Patel & Iversen, 2014). Figure 1, shows the connections between the auditory regions and regions within the parietal cortex, and the connections between the parietal cortex and the dorsal premotor cortex. The tracts highlighted showcase the regions and areas that play a part in the ability for vocal learning (Patel, 2014). Though sea lions of the species *Zalophus Californianus*, who do not exhibit naturalistic movement to rhythmic stimuli in its natural habitat, have shown the capacity to be flexible to tempo stimuli. They can be trained to synchronise to auditory beats and are able to transfer those learnings to new tempi. Though these species of sea lions have not shown evidences of underlying vocal learning, they belong

to the family Otariidae which are descendants of the same ancestry as that of the true seals of family Phocidae and to walruses Odobenidae, both of whom have shown to be vocal learners (Janik & Slater, 1997; Schusterman & Reichmuth, 2008; Oller & Griebel, 2008; Cook, Rouse, Wilson, & Reichmuth, 2013; Patel, 2014).

Figure 1: Summary diagram of tracts in the human dorsal auditory stream (Patel, 2014)

(Patel & Iversen, 2014) argue in their work based on studies conducted on monkeys and humans that though both mammals share neural mechanisms for interval-based timing, they differ in beat-based timing. They hypothesize the basis for the ability for beat perception in humans relies on simulating cyclic intentions within the motor control regions of the brain, and on the connectivity between the motor regions and auditory regions (Merchant & Honing, 2014; Zarco, Merchant, Prado, & Mendez, 2009).

(Merchant & Honing, 2014) proposes the gradual audiomotor evolution hypothesis. The basis of the hypothesis is that rhythmic neural entrainment produced from beat-based timing is the result of gradual evolutionary development in primates. While humans share the capability to process interval-based timing with other related primates, the ability to process beat-based

timing is much more honed and developed in humans comparatively. Figure 2, shows the finding from functional image studies revealing the similarity in the brain's functional capacity between monkeys and humans. It highlights via the blue arrow the involvement of the motor cortico-basal ganglia-thalamo-cortical circuit (mCBGT) in both sequential and temporal processing. This circuitry involves the supplementary motor cortex (SMA) and putamen (Merchant & Honing, 2014; Grafton, Hazeltine, & Ivry, 2007; Harrington, Zimbelman, Hinton, & Rao, 2010; Wiener, Turkeltaub, & Coslett, 2010; Coull, Cheng, & Meck, 2011).

Figure 2: Cognitive circuitry involved in sequential and temporal processing (Merchant & Honing, 2014)

It is largely theorized that there is a need for species to entrain to each other. In nature we observe the synchrony in crickets chirping or the synchronous flashing of firefly species (Patel & Iversen, 2014). Studies have provided evidences of the requirement to be synchronous in a social or naturalistic environment among non-human species (Nagasaka, Chao, Hasegawa, Notoya, & Fujii, 2013). While other studies have shown interactions such as tapping or behavioural actions based on audio stimuli (Hasegawa, Okanoya, Hasegawa, & Seki, 2011; Seki & Tomyta, 2019; Cook, Rouse, Wilson, & Reichmuth, 2013; Fitch, 2013). Figure 3, provides an illustration of how animals display the phenomenon of audio-motor entrainment while listening to music stimuli. The illustrations are based on very popular studies of a cockatoo and sea lion (Kotz, Ravignani, & Fitch, 2018).

Trends in Cognitive Sciences

Figure 3: Animals displaying their capability of audiomotor entrainment, by head bobbing to musical stimuli (Kotz, Ravignani, & Fitch, 2018)

The ability to cognitively perceive beats require the ability to predict timing within sequences which involves the motor system. This is discussed further in this work. There are efforts by others that demonstrate the involvement of the premotor context and the dorsal auditory stream having evolved for the functionality to decipher timings within structured audio sequences (Patel & Iversen, 2014; Rauschecker & Scott, 2009; Leaver, Lare, Zielinski, Halpern, & Rauschecker, 2009).

Developmental studies have shown that infants can indeed perceive beats and metrical representations and anticipate periodic pulses from external rhythmic streams (Larsson, Richter, & Ravignani, 2019 ; Hannona & Johnson, 2005; Phillips-Silver & Trainor, 2005; Winkler, Háden, Ladinig, Sziller, & Honing, 2009). Evidences of evolutionary traits of such cognitive functions being responsible for learning other skills as suggested by Jean Piaget's theory of cognitive development.

Literature explorations on evolutionary roots has laid the foundations for understanding the facets of and the significance of neural entrainment, beat perception and sensorimotor behaviours. It has provided an introduction for its functions, needs and influences on and for

cognitive processing. These can now be extended to understanding how neurological structures and processes play a role in these phenomena.

Neuroscientific Functioning

A magnetoencephalography (MEG) study was conducted by (Iversen, Repp, & Patel, 2009) on metrical interpretation and on imagined placement of beats. Beats perceived within metrical interpretation evoke neural responses in the beta range between 20Hz-30Hz, while responses measured were stronger by 35% on average when the beat was imagined. The study was able to draw a conclusion of the influence of the motor system in interpreting metrical information, even without any actual action taken.

Likewise, similar conclusions have been drawn from studies of the cognitive link between auditory perception and to motor functions even when no overt movement is present or expected. A functional magnetic resonance imaging (fMRI) study showed the activation of the supplementary motor area, mid-premotor cortex (PMC), and cerebellum while only listening to the auditory input. The ventral mid-premotor cortex and dorsal part was activated only when tapping was required (Chen, Penhune, & Zatorre, 2008). Another fMRI study conducted by (Grahn & Brett, 2007), found that the basal ganglia and supplementary motor area (SMA) are recruited for beat perception. In fact, it was discovered when musicians were tested, they showed higher activation independent of rhythm in the premotor cortex, cerebellum, pre-SMA and SMA. In the fMRI study conducted by (Lahav, Saltzman, & Schlaug, 2007), music listeners were not required to perform any actions, saw activations in the frontoparietal motor-related network, the Broca's area, the premotor region, the intraparietal sulcus, and the inferior parietal region.

fMRI studies have shown the involvement of the basal ganglia when beat based stimuli are processed. Where, activations tended to be less prominent when involving stimuli with a missing beat sequence. While activations were even greater when an internal beat representation is required. When functional connectivity was examined it showed the putamen being linked to cortical motor and auditory areas, especially when comparing rhythmic vs non-rhythmic sequences. These suggest the crucial role basal ganglia plays in the processing of beat perceptions (Grahn J. A., 2009). More evidences of beat perception involving the putamen, have been reiterated in other studies (Geiser, Nottter, & Gabrieli, 2012; Bartolo & Merchant, 2015).

The natural cognitive interactions with music, evokes processes linking auditory processing, perception and motor functions. The motor control system is responsible for processes functions relating to timing, sequencing and spatial reasoning. These processes are managed in the premotor cortices, supplementary motor cortices, cerebellum and basal ganglia. While within the auditory processing, the pathway from the dorsal and ventral streams to the parietal and premotor cortices plays a critical part in organising motor functions to audio processing. Neuroimaging shows evidences of co-activation links between the auditory and motor system when listening to music. Deeper analysis shows the activation of the ventral premotor regions during sensorimotor activities while activations are detected in the dorsal premotor regions during rhythm processing. Figure 4, illustrates the regions and their spatial relevance discussed (Zatorre, Chen, & Penhune, 2007).

Figure 4: Coupling between auditory and premotor cortices (Zatorre, Chen, & Penhune, 2007)

Attention, motor and sensory functions generally operate in a rhythmic cycle, as discussed under *Evolutionary Traits*, where neuronal oscillations are the result on the focus of meaningful information within the input stream while filtering out the noise. In the absence of decipherable cyclic rhythm to adhere to, functions like attention operates in a heightened mode, filtering to oscillation within the noise (Schroeder & Lakatos, 2008; Pineda, 2005).

(Rajendran, Harper, Garcia-Lazaro, Lesica, & Schnupp, 2017) identified midbrain pre-processing of auditory information and its involvement to low-level auditory processing, which may influence the cognitive processing of beat perception. Neural entrainment pertaining to

beat perception has shown to manifest within the cortical and subcortical regions (Fujioka, Trainor, Large, & Ross, 2009).

The acquired understanding of neurological functioning lays the foundations for the observations discovered and explored in this study. In particular, to focus on understanding the role of harmonics in identifying entrainment and how this contributes to psychological behaviours, and to build a predictive model of this phenomena.

Temporal Perception

A key aspect in discussing the cognitive capability of processing beats within a complex audio stream is the cognitive mechanism of predicting events. Studies have shown that the motor system is a critical component in providing the functional capability for predicting auditory beats.

Evidences has been provided of the engagement of the cortico-subcortical network, putamen, supplementary motor area (SMA) and premotor cortex (PMC) in temporal sequence analysis promoting cognitive functionalities such as prediction and generation of rhythmic sequences. Where the auditory-premotor interactions are found to be key, furthering evidence of the involvement of the motor system playing a pertinent role in predicting events through audiomotor or visuomotor representation (Grahn & Rowe, 2009; Zatorre, Chen, & Penhune, 2007; Patel & Iversen, 2014; Schubotz, 2007).

Temporal perception of complex acoustic streams is based on the predictive ability of temporal expectancies of the rhythmic regularities within that acoustic sequence, even when significant deviations exist (Large & Palmer, 2002).

Understanding the neurological link to the cognitive processing of temporal perception influences anticipations for observing a model greatly influenced by features which captures these in the EEG based model.

Rhythm Cognition

Rhythm is an acoustic dance of temporal events, which can be complex in character, structured by cyclic patterns of both regular and irregular acoustic events consisting of onsets, offsets, and durations. Beat is the steady heartbeat of an acoustic stream, providing the regular periodic

pulse which can be both explicitly and implicitly supplied. Metre is the temporal signature of an acoustic source measured by the group of beats alternating in saliency of accents in periodic fashion or syncopated rhythms.

The cognitive ability to process rhythmic information and translating those to behavioural outcomes, is a culmination of multiple separate modular cognitive processes functioning in unison. Modular functions such as periodic motor pattern generation, beat perception sourcing from visual or audio modalities and neuronal entrainment. The perception of beats and metres, is also because of the ability to decipher cyclic patterns presented in a hierarchical structure (Kotz, Ravignani, & Fitch, 2018). Potentially each of these cognitive processes is a resultant of its own biologically driven evolutionary development (Kotz, Ravignani, & Fitch, 2018; Fitch, 2013; Patel, 2006).

Periodicity is the cyclic or phasal recurrence of events, whose repetition rate, also known as frequency, is principally uniformed. Motor periodicity is essentially any repetitive motion which follows a periodic pattern. Beat perception is the ability to identify the fundamental pulse in a complex stream, which results from the cognitive ability of the motor system to make predictive conclusions stemming from neuronal entrainment. This cognitive phenomenon is essentially stemming from internal oscillation of brain activities being influenced by external oscillatory information, which characteristically has a fundamentally endogenous pulse. Meter perception is the understanding of metrical phonology and on how multiple cyclic events with its own endogenous pulse characteristics are grouped together. And where perception can home in on the strongest pulse while being flexible enough to still latch onto other key durational patterns in unison. Figure 5, illustrates the connection between acoustic signals, its cyclic patterns and how these leads to the formation of metrical details from beats (Kotz, Ravignani, & Fitch, 2018).

Figure 5: The interconnectivity addressing the breakdown from acoustic signals to cyclic patterns to beats and onto metrical representation (Kotz, Ravignani, & Fitch, 2018).

Neural Resonance Theory

Neural Resonance Theory suggests that neuronal rhythms or oscillations, synchronise or entrain to auditory rhythms at the fundamental frequency of the stream (Nozaradan, Peretz, Missal, & Mouraux, 2011; Okawa, Suefusa, & Tanaka, 2017; Large & Snyder, 2009). Entrainment to cortical rhythms activates cognitive processes such as temporal prediction, error correction and attention for acoustic streams, in a distributed fashion in the cortical and subcortical areas (Large & Jones, 1999; Nozaradan, Peretz, Missal, & Mouraux, 2011; Will & Berg, 2007; Large E. W., 2014).

The neural resonance theory was tested via MEG recordings by (Tal, et al., 2017), and showed consistent results adhering to the neural resonance theory. Findings supported the observation of neural oscillations at the fundamental pulse frequency, signifying the relationship between neural entrainment to pulse perception.

Neural resonance theory proposed by Large, and further established in his following works, discusses on how a stimulus even with a non-periodic cyclic structure interacts with the oscillatory brain rhythms can produce oscillations of a periodic pulse representing the stimuli.

Beyond the fundamental frequency within the stimuli but also its subharmonics (Tal, et al., 2017; Large & Snyder, 2009; Snyder & Large, 2005).

Studies since the efforts from Large et al, has shown that stronger than expected steady state EEG responses are measured for syncopated acoustic stimuli based on its acoustic structure, in line with the expectations of what can deduced from neural response theory (Tal, et al., 2017; Nozaradan, Peretz, Missal, & Mouraux, 2011; Nozaradan, Peretz, & Mouraux, 2012; Nozaradan, Peretz, & Keller, 2016).

Neural Resonance Theory and the learning accumulated in current literature advocates at a neuroscientific level the significance of examining EEG signals at specific frequencies informed by the stimuli and its harmonics.

Neural Entrainment

Neural entrainment is the process resulting from the interaction of between the internal neural oscillation which varies as per its neurofunction and region, and the external stimuli ultimately leading the rhythm perception (Velasco & Large, 2011; Large, Almonte, & Velasco, 2010).

When an audio stimulus having rhythmic cycle at a frequency of f Hz is provided, a dominant neural response is also observed at the corresponding frequency. Neural tracking shows that neural responses do not return to baseline frequency cycles but slowly drift to reflect the stimulus in terms of frequency evocation. When there is a rhythmic input stream, the cognitive system complies to its implicit oscillatory enforcement reflected by the cyclic oscillation via neuronal activity. This phenomenon is described as Neural Entrainment, and occurs when any form of rhythmic stream is introduced to the sensory processing system (Zhou, Melloni, Poeppel, & Ding, 2016; Schroeder & Lakatos, 2008).

(Nozaradan, Peretz, Missal, & Mouraux, 2011) showed that the dynamic cognitive processing of neural entrainment to beats can be observed via EEG. They were not only able to identify the EEG responses relating to the beat of the stimuli, but also imaginary periodicities corresponding to the perceived beat. These EEG responses links the similarity between both the sub-conscious and conscious cognitive processing of musical beats. And on how the responses from these two separate conscious and unconscious cognitive processing, manifests by analysing measurements from EEG using the same analytical methodology. Beat is

perceived in the frequency range between 0.5-8Hz (London J. M., 2004; Lerdahl & Jackendoff, 1983; Large & Snyder, 2009).

Further works from (Nozaradan, Peretz, & Mouraux, 2012) showed that SS-EPs were even exaggerated when the underlying acoustic energy were not prominent. They were also able to identify that these SS-EPs enhancements were selective within a resonance frequency range. These were even observed for internal representations, providing insight that neural entrainment can occur both spontaneously and intentionally.

Neural entrainment to pulses within complex rhythms have shown to occur in the auditory and motor cortices linking to temporal expectancy and, attentional and movement coordination (Large, Herrera, & Velasco, 2015; Large & Snyder, 2009). This continues to be true even when rhythms are only imaginarily anticipated rather than auditorily sensed (Herff, et al., 2020).

EEG and MEG studies have shown that delta, beta and gamma band responses to auditory rhythms comply with Dynamic Attending Theory and its predictions. The link between beta band responses in motor processing and long-range intra-cortical interaction, supports the theory that the motor system influences the perception of sound, even in the absence of overt movement (Large, Herrera, & Velasco, 2015). This conclusion supports the claim that voluntary actions such as tapping to the beats of a music directly correlates to the sub-conscious processing of musical tempos.

Beat Perception

“A beat: A perceived periodic pulse that structures the perception of musical rhythm and which serves as a framework for synchronized movement to music.” (Patel & Iversen, 2014). It is alignment to this periodic structure that listeners are naturalistically drawn to, often expressed in physiological behaviours such as tapping along to (London J. , 2012).

There are multiple studies that validate the findings, activations and functional couplings among the following areas in the brain when it comes to processing beat perception, such as the premotor cortex, basal ganglia (putamen), supplementary motor regions, motor planning regions, auditory cortex and, including the dorsal auditory pathway and cortico-subcortical network (putamen, supplementary motor area, and premotor cortex), (Patel & Iversen, 2014; Chen, Penhune, & Zatorre, 2008; Rowe, 2009; Kung, Chen, Zatorre, & Penhune, 2013; Grahn & Brett, 2007; Grahn & Rowe, 2009; Geiser, Notter, & Gabrieli, 2012) (Nozaradan, Schwartze,

Obermeier, & Kotz, 2017; Bengtsson, et al., 2009; Chapin Heather, 2010). Figure 6, highlights these areas in the brain which are activated as observed by fMRI studies (Merchant, Grahn, Trainor, Rohrmeier, & Fitch, 2015).

Figure 6: Activations in highlighted areas as observed by fMRI studies (Merchant, Grahn, Trainor, Rohrmeier, & Fitch, 2015).

Music are spectotemporally complex sound information. Listeners often respond to music by moving rhythmically synchronising to its periodic pulse. This synchronicity to the musical beat as perceived is one stemming from a predictive process, highlighting a highly accurate temporal predictive mechanism (Iversen, Repp, & Patel, 2009; Rankin, Large, & Fink, 2009; Noorden & Moelants, 1999). The predictive ability to auditory beats has a broader significance on the cognitive ability to learn other equally rhythmically beat potent materials (Povel & Essens, 1985). In addition, the role of basal ganglia is shown to be closely linked to the process of beat prediction vs identifying the beat itself (Grahn & Rowe, 2013).

Beat perception as discussed can vary among listeners. It is very possible for listeners to emphasize on different phases such as the upbeat or the downbeat. Even switching between different beat phases during the listening experience. This flexibility also leads to another observation where listeners maintain a beat outside of the strong signals found within the audio streams, leading to the hypothesis of the constructed nature that do exist in the cognitive process of beat perception. Figure 7, shows the multiplicity present naturally in beat perception (Patel & Iversen, 2014). The spectrogram with its inverted arrows points to places of strong beat acoustics. The plot at the bottom shows the corresponding tapping timing to those strong beat acoustics and shifts in between (Repp B. H., 2007; Patel & Iversen, 2014; Repp, Iversen, & Patel, 2008; Iversen, Repp, & Patel, 2009).

Figure 7: Spectrogram of input stream (top) used to observe tapping responses of listeners (bottom) (Patel & Iversen, 2014).

Beat perception also plays a clearer constructive role in rhythm perception. The perception of where the beat falls within the same temporal pattern of the input stream among listeners greatly impacts on how the rhythm is perceived. The influence of how the beat is perceived is strong enough to alter the perception of the rhythmic structure of the input stream,

overpowering the endogenous characteristics of the beat within its temporal patterns (Repp, Iversen, & Patel, 2008; Repp B. H., 2007; Iversen, Repp, & Patel, 2009).

These functional characteristics do emphasize that beat perception is not only the result of neural entrainment but also something that can be consciously imposed, the intrinsic interpretation of extrinsic stimuli. This is critical in understanding how examining tapping responses in the frequency domain can overcome this flexibility in phase variation and multiplicity in beat perception among listeners.

An external physical measure of beat perception is to record the tapping frequency of listeners to the music stimuli. Accuracy in the ability to measure tapping duration is critical as it will be a key component in validating against the features being examined for its predictive capability. It has been observed that accuracy can be obtained from listeners when the beat intervals are perceived between 200ms-1800ms, highest accuracy being at 600ms (Fujioka, Trainor, Large, & Ross, 2009; Repp B. H., 2005). While another study found beat perception to be between 250ms-2s, where the strongest are achieved between 400-1200ms, but preferred at around 600ms (Patel & Iversen, 2014; London J. , 2012). Studies from (Fraisse, 1982; Moelants, 2002) place the ability of listeners to perceive beats best between 100-140 beats per minute (BPM) or where it was possible to extend the range between 80-160 BPM for selective distinctively interpretable stimuli.

Beat perception is also inextricably tied to other cognitive phenomena such as Dynamic Attending Theory and Sensorimotor Synchronization. All these modular cognitive components come together to play a critical role in our study and efforts to develop a predictive model. To be able to monitor the flexibility of the cognitive system to perceive beat frequency and the listeners' physiological manifestations, in our study being tapping responses.

Dynamic Attending Theory

Dynamic Attending Theory is the ability when listening to complex auditory stimulus, to subconsciously focus on the isochronous beat which is both rhythmic and the focal event being attended to. While simultaneously exhibiting the flexibility to shift phase-wise to a faster or slower rhythmic pattern decipherable within the same auditory stimulus (Drakea, RiessJones, & Barucha, 2000; Toiviainen & Snyder, 2003; McKinney & Moelants, 2006; Martens, 2011; Patel & Iversen, 2014).

These phenomena are important to comprehend, as we will observe this flexible behaviour being displayed in tapping responses among the listeners, as we extract features to build a predictive model on. Adding justification to the approach of not only monitoring the computed beat frequency of each stimulus but also to extract features from its subharmonics and harmonics as well.

Sensorimotor Synchronization

Sensorimotor synchronization (SMS) is the rhythmic coordination of beat perception and physical action (Repp B. H., 2005; Repp & Su, 2013).

(Toiviainen & Snyder, 2003) work tested the feasibility of using behavioural responses to demonstrate the processing of beat perception, to model the influences of rhythmic stimuli on sensorimotor synchronization. This has been pragmatically exercised by multiple studies using behavioural responses such as tapping to beat, metric or rhythm. And this methodology has been used in varying forms on humans and non-human species, to study and answer questions relating to evolution, anthropology, neuroscience, cognitive science, developmental psychology and social cognition. Example of such studies are (Drakea, RiessJones, & Barucha, 2000; Toiviainen & Snyder, 2003; McKinney & Moelants, 2006; Martens, 2011; Patel & Iversen, 2014; Rankin, Large, & Fink, 2009; Provasi, Anderson, & Barbu-Roth, 2014).

The ability to both consciously and unconsciously synchronise to a beat in a SMS tasks is critically reliant on the cognitive ability to process error correction. Cognitively this can be distinguished in two functional mechanism, the phase-correction mechanism and the period-correction mechanism. The phase correcting focuses on the unconscious relative phase alignment which are too inconspicuous, while the period correcting mechanism is responsible for the conscious corrections of the internal timekeeping of internal rhythm (Repp B. H., 2001; Repp & Keller, 2004; Comstock, Hove, & Balasubramaniam, 2018). This ultimately explains why some listeners can maintain a clearer synchronised tapping replicating the beat of the stimuli being listened to compared to others. Also, explains the differences in ability of individual listeners to maintain a steady beat in tapping or to maintain a cyclic tapping expression for stimuli without a salient beat information.

Understanding the neural processes involved in auditory timing is to acknowledge the multiple timing related cognitive subprocesses, such as SMS, interval timing, rhythm perception, timing

recall, time perception and motor timing. Auditory rhythm perception and temporal processing relies on the motor network, supplemental motor area (SMA), primary motor cortex, lateral premotor cortex, anterior cingulate, basal ganglia, and cerebellum. While the SMA and pre-SMA are linked to motor timing, subcortical regions (networks recruiting the basal ganglia and cerebellum) are linked to sub-second time perception, sub-second interval timing and rhythm timing. Cortical structures, such as the SMA and prefrontal cortex are linked to supra-second interval timing. Rhythm perception is linked to the dorsal auditory stream which connects the auditory cortex, posterior parietal cortex and the motor cortex, (Repp & Su, 2013; Janata, Tomic, & Haberman, 2012; Iversen & Balasubramaniam, 2016; Ross, Iversen, & Balasubramaniam, 2016; Coull, Vidal, & Burle, 2016; Merchant & Yarrow, 2016; Schwartz, Rothermich, & A.Kotz, 2012) (Wiener, Turkeltaub, & Coslett, 2010; Grahn & Rowe, 2009; Coull, Cheng, & Meck, 2011; Teki, Grube, Kumar, & Griffiths, 2011; Hove, Fairhurst, Kotz, & Keller, 2013; Patel & Iversen, 2014; Ross, Iversen, & Balasubramaniam, 2018) (Comstock, Hove, & Balasubramaniam, 2018).

It is worth noting the developmental aspect of sensorimotor synchronization, from the perspective of cognitive functionality. The ability to synchronize with little variability and robustness is greatly tied to the development of motor functionalities. Cognitive development plays a greater role in rhythmic reproduction and the ability to maintain synchronicity (Monier & Droit-Volet, 2019). The performance on tasks relating to sensorimotor synchronization has shown to develop significantly during childhood right up to late teens and thereafter remaining consistent until old age (Drewing, Aschersleben, & Li). Numerous studies on atypically-developing children and adults have shown that cerebellar dysfunction plays a critical impairment role on the ability for Sensorimotor synchronization (Repp & Su, 2013).

There are also efforts furthering such research by providing new scientific methodological tools for the purpose of designing, conducting and analysing experiments focused on sensorimotor responses and synchronisation to cues with high temporal accuracy. Tools like MatTap from (Elliott, Welchman, & Wing, 2009), FTAP from (Finney, 2001) and Tap-It from (Kim, Kaneshiro, & Berger, 2012). Tap-it was used in the NMED-T experiment, which is the basis for this work (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017).

Frequency Domain Analysis

The cognitive system processing of dynamic sensory stimuli can result in rhythmic signals captured by electroencephalography (EEG), magnetoencephalography (MEG), or local field potential recordings (LFP), (Zhou, Melloni, Poeppel, & Ding, 2016; Wang, et al., 2012; Robert Galambos, 1981; Nozaradan, Peretz, Missal, & Mouraux, 2011). Signal processing methodologies such as Fourier Transformation where signals captured on the temporal scale are analysed on the frequency spectrum. Performing a Fourier analysis not only highlights the fundamental frequency of a periodic signal but also its related harmonics (Zhou, Melloni, Poeppel, & Ding, 2016).

The key representations of EEG and tapping signals represented in this thesis can be explored theoretically. Tempos are essentially isochronous pulses and the resulting mimicry of tapping responses also produces a sinusoidal signal marked at the fundamental frequency. Given a signal of period T , it cycles at a rate of $1/T$, which represents the fundamental frequency the sinusoidal signal. Figure 8, illustrates a clear isochronous pulse being represented in the frequency domain. Signal of such form are what will be observed when a listener can synchronise his tapping to the tempo perfectly. However, such signals are theoretical and behavioural variations will produce subtle to drastic noise. When signals are noisy, the noise too is bound to be captured and represented within a range of higher and lower powered signals around the fundamental frequency. In less erratic tapping synchronicity, a higher measure of power will be observed on harmonically related frequencies based of the fundamental frequency (Zhou, Melloni, Poeppel, & Ding, 2016).

Figure 8: A 1-Hz sinusoid wave viewed in the temporal and frequency spectrum (Zhou, Melloni, Poeppel, & Ding, 2016)

Fourier Transformation essentially analyses a signal sampled at multiple frequencies. There will only be recorded signals in the frequency domain when a non-zero signal is observed in the periodic signal at the sampled frequency. A high oscillating periodic signal will produce a proportionally high-powered spectrum signal, at the fundamental frequency. As the periodic signal is sampled at multiple frequencies, proportionately powered signals are translated in the spectrum frequency at multiples of the fundamental frequency. This is what results in harmonic information in the frequency spectrum. Figure 9, illustrates this phenomenon (Zhou, Melloni, Poeppel, & Ding, 2016).

Figure 9: A 5-Hz sinusoid wave viewed in the temporal and frequency spectrum. The frequency spectrum shows multiple signals at both the fundamental frequency and its harmonics (Zhou, Melloni, Poeppel, & Ding, 2016).

Another critical phenomenon observed in Figure 9 is the shift in the representation of the strongest frequency signal from the fundamental frequency at 5Hz to the 4th harmonics at the 4Hz. An ambiguity exists in the analysis of Fourier Transformed signals (Zhou, Melloni, Poeppel, & Ding, 2016).

This motivates the exploration of not only the signal at the fundamental frequencies discovered during neural entrainment, but also the signals at its calculated harmonics. In this study, analysing harmonically decomposed signals over fundamental signals leads to improved understanding of how best to predict tapping responses of the primary frequency captured via EEG.

In Summary

This study focuses on building a predictive model by leveraging on EEG measurements validated against tapping responses. EEG measurements based on the phenomena of neural entrainment and beat perception, are extracted at specific frequencies pertaining to the beat frequency and its harmonics. Harmonics are explored for the first time through this effort to explore their ability to provide further knowledge on the cognitive processing of beats within acoustic signals based on its inherent relationship to both the perception of beats and entrainment of neurons.

In line with this effort, the current literature has been reviewed and key fundamentals, theories, data, and findings have been explored. Beginning with an evolutionary understanding, comprehending the neuroscientific backing and all the various modular cognitive processing and sub-processing involved. Extending these explorations onto how these various facets contribute to the specific elements and crucially their interactions with one another. These collectively provide the fundamentals to approach the objective of this study in building a predictive model to identify the role harmonics play in how listeners perceive acoustic signals and respond through sensorimotor actions.

Chapter 3. Dataset & Experiment

Dataset

This study and analysis were carried out on the Naturalistic Music EEG Dataset – Tempo, NMED-T dataset (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017).

The NMED-T provides dense-array EEG recordings, tapping responses, and behavioural ratings of familiarity and enjoyment, captured during and post music listening sessions. Basic demographic information of the participants is also provided. Pre-processed EEG readings which have been aggregated across participants and within each of the individual songs are also published in MATLAB format.

Recorded from 20 participants, (Age: 18-29 *Mean*: 23, Male: 14, Female: 6). Of the 20, some have formal musical training (Training: 17, *Mean*: 8.4yrs), and as a cohort generally listens to music on average 14.5 hours/week.

EEG was recorded using 128 electrode channels at 1 kHz, with vertex reference. Participants also provided their ratings based on the dimensions of familiarity and enjoyment for each of the full-length songs.

Stimuli audio used for EEG recordings were made of 10 distinct songs, (Length: 270s–300s, BPM: 56-150). All 10 songs played for a duration between 4mins - 5mins long, were of various genres, displayed wide variability in tempos and inherited electronically produced beats in duple meter. The stimuli used for tapping measurements were extracted in part from these songs. Table 1 lists the songs and basic information about the audio stimuli.

Song #	Title	Artist	min:sec
1	First Fires	Bonobo	4:38
2	Oino	LA Priest	4:31
3	Tiptoes	Daedelus	4:36
4	Careless Love	Croquet Club	4:54
5	Lebanese Blonde	Thievery Corporation	4:49
6	Canopée	Polo & Pan	4:36
7	Doing Yoga	Kazy Lambist	4:52
8	Until the Sun Needs to Rise	Rüfüs du Sol	4:52
9	Silent Shout	The Knife	4:54
10	The Last Thing You Should Do	David Bowie	4:58

Table 1: The 10 songs used for stimuli (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017).

In total the dataset contributes 55 files, totalling to 39GB worth of data, shared under “CC-BY Attribution” license. Table 2, provides a detailed listing of all the files shared under NMED-T dataset.

Dataset files

- **NMED-T_README.pdf** (this document)
Informational document describing the dataset (1 file).
- **songSS_Imputed.mat**
Cleaned and aggregated EEG recordings for song SS (10 files).
- **PP_R_raw.mat**
Single raw EEG recording for participant PP, recording R (40 files).
- **behavioralRatings.mat**
Aggregated behavioral ratings (1 file).
- **TapIt.zip**
Archive of aggregated and raw tapping responses (1 file).
- **participantInfo.mat**
Basic demographics of participants (1 file).
- **Code.zip**
Collection of Matlab scripts for performing illustrative analyses in the Losorelli et al. (2017) ISMIR paper, as well as helper functions and files (1 file).

The contents of these files are described in detail in the Losorelli et al. (2017) ISMIR paper.

Table 2: Listing of files in the NMED-T dataset (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017).

This dataset provides an opportunity to examine further details on human processing of naturalistic music, supported by a variety of contributing data.

Protocol

Participants first provided basic bio-details. Next as per the protocol for each of the 10 commercially available songs, participants listened to a song in full while an EEG reading were being recording simultaneously. At the end of each song, they provided their feedback on likeness and familiarity. Next, they tapped along to a 35sec excerpts, 60secs into the song, played in a random order (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017). Figure 10 provides a simple illustrates of the research paradigm.

Figure 10: NMED-T Research Paradigm

Part of the data provided were the raw EEG data, pre-processed EEG measurements which were down sampled to 125Hz, raw tapping responses using Tap-It, and demographic information (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017).

Participants sole tasks were to listen attentively to the music without any preconceived expectations. Participants were not aware of the beat tapping phase of the experiment, ensuring a closer to naturalistic listening experience for measurement via EEG during the listening session. These EEG recordings offer data measurements which can now be used to explore various aspects of unbiased music cognition processes.

Analysis Review

Though beat perception is the phenomenon of interest, it is important to realise that these songs are characteristically complex and exhibit significant auditory presence at various frequencies. Plotting the magnitude spectra as described by (Nozaradan, Peretz, & Mouraux, 2012), highlights spectral peaks at the harmonics and sub-harmonics of the beat frequency (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017). Figure 11 is a replica of the graphs published in

their paper. These hierarchical timescale characteristics could be informative to studying the influence on beat perception and behavioural responses.

Figure 11: Amplitude envelopes and beat hierarchies found in the stimuli (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017).

Using PCA filters on averaged responses highlights beat related frequencies and harmonics in the EEG spectra between 5-10 Hz (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017). Across listeners, perception varies peaking at the tempo and tapering off at the harmonic and sub-harmonic frequencies. This could be hypothesised as an artefact of resonating oscillator in the perceptual-motor system. The resonance curve explains this phenomenon of distribution (Noorden & Moelants, 1999).

Figure 12, is a replica of the graph analysis illustrated in the (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017) paper. User tempo tapping activity shows most can keep in tempo to a specific song. Song 1 and 2 has the distinction of a significant number of participants tapping at twice the rate. This discrepancy and the desire to understand better on how tapping responses were being influenced by cognitive processes, was the inspiration for the objective of this study.

Tapping responses. Top: Instantaneous tempo over time for individual participants (gray), with median across participants in black. Bottom: Histograms of median tempo, over time, for individual participants. Ground-truth tempos are shown with orange ($1/2x$ tempo frequency), green (tempo frequency), and red ($2x$ tempo frequency) lines.

Figure 12: Analysis of tapping responses and stimuli tempo (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017).

NMED-T by (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017), studied the music stimuli both at an acoustic level and its perception. Music perception and MIR approaches reveal cortical and behavioural evidence of entrainment to musical beat. It showed links observed through EEG recordings the cognitive processes, the brain regions activated and how the musical acoustic information infers those processing activations. These have paved the way to provide the basics and basis for extending those learning by taking it further into exploring pragmatic applications. This study extends those learning from MIR and EEG analysis, into harnessing harmonics for predictive information which in conjunction with EEG recording offers further learnings. It further proposes the functionality of pre-recorded EEGs in being predictive of future behavioural activities which were not explicitly instructed for. Like the NMED-T study, this study also links the learnings from MIR based analysis to cognitive processing and perception. It shows how harmonics play a fundamental role in the MIR analysis explored relating to neural entrainment and beat perception. And finally meshing the collective learnings of these complex facets into a predictive model of sensorimotor behaviours based off participant musical processing and perception.

Chapter 4. Study Methodology

Tools

For the analysis, MATLAB R2019a was used, along with its toolboxes MIRtoolbox v1.7.1 (Lartillot & Toivainen, 2007) and EEGLAB v14.1.2b (Delorme & Makeig, 2004).

Methodology

The objective is to develop a predictive model leveraging on harmonics, to explore the importance of harmonics in beat and neural entrainment in identifying the frequency at which a listener would instinctively tap along to for a song. This requires a multimodal approach taking into considerations the EEG signals, tapping timings and MIR details for the song. These sources of modalities need to be processed in a manner that allows these modalities to be analysed together in a common domain. Figure 13 illustrates the process flow, which addresses the individual methodologies adopted for processing these data and how these separate elements will be brought together collectively into a predictive model.

Figure 13: Overall Strategy Adopted

The tapping data needs to be first converted into a time orientated data stream before FFT can be applied to convert it into the frequency domain. Once in the frequency domain, the signal with the highest registered power will be considered the tapping frequency. This will be the dependant variable for the predictive model.

The songs will be analysed using MIR methodologies to extract the beat frequency. With the beat as the fundamental frequency, the harmonics and sub-harmonics will be computed. These computed frequencies will be markers to be explored within the EEG signals.

The processed EEG data stream is provided in the time domain and first needs to be converted into the frequency domain. Within the frequency domain, data are extracted at the beat frequency and its subharmonics and harmonics, for every single electrode. These frequencies are computed for each respective song using MIR methodologies. These extracted data features are fed into the predictive model as independent variables.

The predictive model is generated using a multinomial regressive model. Each electrode and frequency extract provide the accuracy ratings for identification.

This strategy effectively is a cross-disciplinary effort between MIR, Cognitive Science and Computational Science approaches and methodologies being incorporated for a research objective.

Analysing in the Frequency Domain

One consideration when performing analysis on tapping timings, EEG recordings and beat frequency with its harmonics, is that these data are from different sources, of different measurements, and in different domains. To analyse the relationships between these data, they must be compared under a common domain. Considering the phenomenon being explored, an obvious baseline for assessment among these data can be achieved in the frequency domain.

Fast-Fourier Transformation (FFT) was applied to all the data signals to convert them into the frequency domain. The first step in the conversion into the frequency denomination is to identify the baseline specifications. Since the pre-processed EEG recordings were of the highest resolution at 125Hz and the most resolution sensitive data. The tapping data was up sampled to match with the resolution of the EEG data to ensure that EEG information is not lost during the transformation process. Tapping data were already a binary data stream and up

sampling was the process of adding '0's to pad the data for each second. The bin sizes used in the FFT were the maximum allowable size computed based on the signal length. The size of the bin represents the region of spectrum sampled and the resolution of the sample in terms of frequency. While the sampling rate of the processed EEG data is 125Hz, the effective range extractable post an FFT process is between 0-62.5Hz. The signals were first normalised by the length of the signal followed by max-min normalization. The signal details representing inactivity were noise and subsequently removed.

Tapping Data

Tapping by participants was timed for the 35 sec stimuli, producing a data series of timings referenced from the start (*Figure 14a*). This tapping data first needs to be converted into a binary time series format, where a '1' indicates a tap. The tapping data is converted into a format representing a sampling rate of 125 Hz, where 125 entries represent one second. The time series tapping data varied in length among participants for songs. For uniformity the longest recorded duration was taken as the reference and all tapping data were extended to 37 sec (*Figure 14b*).

These processing ensure that all the converted tapping data are of the same length, representing a sampling frequency of 125Hz. FFT is performed to convert the signal from the time domain and into the frequency domain. These produces a frequency bin size of 0.0286Hz and a total of 2188 bins.

Peak extraction was used to identify the most consistent rate of tapping by the participant (*Figure 14c*). The width of the peak signifies the variance found in the tapping. The narrower the peak, the more consistent the tapping. The frequency of the peak signal and their harmonics will be the frequencies at which the EEG signals are extracted for the predictive model.

Figure 14. (a) Left: Tapping timings. (b) Conversion of tapping timings into a time series binary signal. (c) FFT of tapping data.

Song Data

“Humans possess a unique ability to process music, and many topics in Music Information Retrieval (MIR) involve computational modelling of human perception.” - (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017)

MIR methodologies provide the tools to breakdown a complex acoustic signal into specific measurable or quantifiable parameters for discussion. Of interest are specific measurements on tempo and metre.

We used MATLAB toolbox, MIRtoolbox (Lartillot & Toivainen, 2007) to extract the tempo information from the songs. The extracted tempo frequency becomes the fundamental frequency to which tapping data are aligned to and from which harmonics are calculated from. The tempo frequency from the songs and its derived harmonics will also be used as markers to extract specific data from the EEG frequency spectrum as features for the modelling.

EEG Data

The pre-processed EEG data was taken in its current form to be analysed. FFT was performed on the EEG data, for each of the 125 EEG electrodes. This converts the EEG data into the frequency domain as shown in Figure 15, where its peaks can be examined against the peaks from the tapping data.

Figure 15. Conversion of EEG data into the frequency domain for participant 1, song 1, and electrode 100.

The hypothesis driven approach is focused on finding the relationship between tapping and the power in the EEG signal at beat frequency and its harmonic. Importantly, the EEG was extracted whilst listening, but not tapping, to the song. The tapping data was collected in a second phase, after the EEG data was already collected.

Predictive Modelling

Multinomial Logistic Regression

One of the objectives of this thesis was to proof the concept that future tapping behaviours can be predicted from EEG extracted features at the beat frequency and its harmonics, from earlier music listening. A multinomial logistic regression was selected as the choice model of methodology for its simplicity and effectiveness.

Range of Observance – On Tapping

Once in the frequency domain, it becomes obvious that there is an effective range for examining beat perception. The frequency range for observance can be deduced by two factors, 1) Understanding the ceiling limit of the speed that is humanely possible to tap along to, 2) The

range within which the cognitive system is able to perceive beats, which is estimated to be between 0.5-5Hz (Nozaradan, 2015).

A range of 0-7Hz was selected for observance, to ensure that we encapsulate the ranges defined in current literature and leaving room for possible presence outside expected range of performance.

Range of Observance – On EEG

Despite trying to use the highest resolution afforded by the FFT process, there will be misalignments introduced by the bin sizes between tapping data and the EEG data. To ensure that when using the identified tapping frequency as reference for extracting EEG power ratings, corrections need to occur.

One way to overcome, this challenge was to extend the number of bins when referencing the harmonics. One data driven methodology to identify an appropriate range is to take the maximum observable distance between the tapping frequency and the nearest harmonic. Taking the maximum distance across all harmonics for a song and then across all songs and dividing this distance by the computed bin size, provides the number of bins apart these two references. Performing a ceiling rounding would provide the number of bins to extend the frequency band by for observation.

Beat Frequency and Harmonics

As a hypothesis driven approach, features were extracted from the EEG data based on the beat frequency of the songs and its harmonics, as illustrated in Figure 16. The features extracted at the specific frequencies were its power intensity. These features were compared against the tapping frequency to determine the most predictive features extractable from EEG data for building the predictive model.

Figure 16: Extracting Features from EEG Frequency Spectrum Data.

Normalising on Harmonics

The variances in EEG features produced from different songs and listeners' perception to beats needs to be addressed, to permit analysis across songs and participants. With the identification of the fundamental beat frequency from the songs itself, the harmonics at which features are to be extracted can be identified. Though these harmonic frequencies are dependent on the songs, the harmonic positionings themselves are neutral. Normalising the data to fundamental frequency aligns the harmonics across songs and participants, as shown in Figure 17.

Using the power ratings extracted from the harmonics across songs and participants can be analysed. To ensure that a wide spectrum of harmonics was examined, 20 sub-harmonics and 20 harmonics were examined. We indexed these frequencies starting with 1 at the lowest sub-harmonic, making the fundamental frequency the 21st harmonic index.

Figure 17: Centering of EEG features based on harmonics

Leave-n-Out Cross-Validation (LnO CV)

Because of the small dataset, an exhaustive cross-validation is not only possible but prudent to maximise the predictive ability of the model being built. A LnO cross-validation approach was taken to maximise parameter estimates and minimise variance in performance statistics. The partition was achieved across the participant dimension, as our focus was on predicting participant behaviour.

Figure 18, provides a simple illustration highlighting how participants are being rotated out as validation data, while the remaining participant data are used for training the model. The mean performance of the collective runs is used to report on the performance the model.

Leave-One-Out Validation

- Leave one participant out cross validation.

Figure 18: Leave One Out Cross Validation process flow

Chapter 5. Results & Analysis

This thesis aims to distinctively focus on the question of future behavioural predictability, using prior EEG data. In line with this objective the data and different modalities are examined and discussed with a broader scope in mind. This is to explore opportunities for investigations and to better understand the characteristics of the different elements and how they interact with one another. A key focus being to leverage on harmonics in addition to beat frequency, to inform EEG features to best predict future sensorimotor behaviours.

Objectively some of the methodologies used in studying the NMED-T dataset will also be reviewed and referenced against approaches taken in this study. This is to provide verifiability and add validity to the approaches and methodologies adopted taken here.

Tapping Data

The main objective of analysing the tapping data is to identify the tapping frequency. (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017) approach was to take the median tempo timing and convert that into a tapping frequency. The approach taken in this work was to instead use FFT on the tapping timing data to identify the most salient frequency. An FFT approach will be much more sensitive to tapping responses while being robust enough to withstand variations within the data which is often seen in tapping tasks. A direct comparison is performed between these approaches as discussed in the Tapping Frequency: FFT vs Median section.

The tapping data will be analysed in its various stages of processing, to not only draw the line along the rational progression of the data being prepared to be in its final form for building the model, but also to discuss the impact of the frequency extraction methodology adopted. The analysis of the tapping data will be examined in the following forms, 1) Time stream data, 2) Frequency signal data, 3) Median frequency extraction and 4) Comparing Results from Both Methodologies.

Time Series Tapping Signal

The first step was to convert the timing of taps into binary time series signal data. The signal data for Song 1 when plotted are shown below, a subset of the participant data is shown in

Figure 19. Examining the plots shows that participants are not always able to maintain a steady tapping rate from start to finish. It can be observed that, there are tapings with even graver inconsistencies as in the form of missing data from participant 9 and erratic tapping behaviour from participant 7. Empirical observation does suggest that a more robust tapping frequency method can be beneficial for analysis.

Figure 19: Tapping data stream plotted for a subset of the tapping dataset

Tapping Frequency Data

The binary timing data is now ready for FFT processing. The frequency signals for Song 1 are plotted and shown below, a subset of the participant data is shown in Figure 20. As discussed in the review, it can be observed that the methodology is robust in managing variability. Despite the variability present in the data, there always is a distinct frequency signal with the highest power rating identifiable. Participant 7 with the most erratic tapping behaviour, when observed in the frequency domain also clearly demonstrates a distinguishable peak. In the case

for participant 2, where there are two peaks of almost similar power ratings, are also distinguishable as they are not identical. Through the full tapping dataset for all the participants and songs, there always was one distinguishable peak with the highest power rating. The multiple peaks observed are an artefact of the methodology. If we marked the peak with the highest power as the fundamental frequency and it can be computed that the other discernible peaks are harmonics to the fundamental frequency.

Figure 20: Tapping frequency signal plotted for a subset of the tapping dataset

Tapping Frequency: FFT vs Median

Using the technique described in (Losorelli, Nguyen, Dmochowski, & Kaneshiro, 2017), the tapping frequency was computed using median observation. The median frequency signals for Song 1 are plotted superimposed over the frequency equivalence for a subset of the participant data is shown in Figure 21. The median frequency signals are plotted as a red dashed line and the height is representative of its own power rating. On the frequency spectrum it can be easily

observed that both techniques approximately come to the same conclusion. The greater difference between the frequencies indicated by both methodologies appears for participant 7. Though the differences in general seem small, the frequency approach we use provides a higher resolution at 0.03Hz, and is more robust against a noisy data (e.g., Participant 5 below), or when participants change between half and double time tapping (e.g., Participant 7).

Figure 21: Tapping frequency signal with median frequency (red dotted line) plotted for a subset of the tapping dataset

Examining the fundamental frequency being identified by both methodologies will provide us a better understanding on the differences between the performances of both techniques. Figure 22, shows the performance from both methodology for all participant tapping data for Song 1. It can be observed that the FFT technique comparatively produces a more salient understanding of the tapping frequencies displayed by the listeners with lesser variations. Using the FFT methodology, only three fundamental frequencies were identified across all participants, and all these frequencies are harmonically distanced from each other and were computable to a very

high precision. While comparing the median tempo approach, six fundamental frequencies were identified, and these frequencies do not all align harmonically either. Comparing the captured frequencies between both methodologies also shows that they do not align either. This is significant, as choosing the right model will greatly impact the predictability of the model.

Figure 22. Identifying tapping frequency for all participants and song 1. (Left) Identifying using FFT peak extraction. (Right) Identifying using median tapping frequency.

Picking out the participants who produce significantly different tapping frequencies due to the respective methodologies is a great place to start identifying the more reliable technique. Figure 23, clearly shows plots for participants when tapping periodicity is erratic, the FFT process is far more robust in identifying the tapping frequency.

Figure 23. Comparing the extraction methodologies of beat frequency among participants with stark difference. The frequency computed using the median methodology is highlighted with a red dashed line. The strength of the FFT process, can be observed.

An exceptional case which lends itself to discussing the phenomenon of both techniques is participant 20, illustrated in Figure 24. FFT technique can not only capture accurately both paces exhibited by the participant, but also illustrate the dominance of either tapping pace over the length of the exercise. The median tapping interval is 1.0217sec which translates to 0.9429Hz.

Figure 24. Peak extraction for participant 20 song 1. Highlighting the shortcoming of the FFT process.

Our analysis clearly shows that overall, the FFT approach provided a higher resolution for screening for the fundamental tapping frequency. It was also sensitive to tapping responses in terms of periodicity and variability. Its processing can be further improved via filtering techniques, which is beyond the immediate need for this proof of concept.

Tapping Distribution

Based on the discussed FFT approach to extracting participant tapping frequency, the tapping frequency of all participants for all the songs were extracted. The distribution of the tapping frequencies among participants within a song can be examined in Figure 25. The spread in the tapping frequency is key to identifying the viability of a predictive model, especially with the presently available limited data. In addition, a narrow spread of the tapping responses to a stimulus, backs the hypothesis behind the approach to predict these responses.

The participants adhered to a tempo for all songs with little variation. Most songs had dispersions of 2-3 distinct frequencies. Only Song 5 and Song 6 had a higher variation. It is also important to note that despite this variation within these 2 songs, most remained within the 2-3 frequency variations and only individual outliers were the cause of more variations. In general, these tapping frequency variations within a song, shows that despite having the same stimuli, it is uncommon to have the whole cohort interpret the beat frequency in a homogenous fashion. In fact, only Song 2 managed to elicit such uniformity.

The similarity in tapping responses among participants across songs highlights an underlying cognitive mechanic. However, the variation also was of interest on how this cognitive mechanic functioned and was the inspiration to identify if it was possible to monitor these individual differences resulting from the underlying cognitive process.

Figure 25. Participant tapping distribution among songs

The tapping distribution highlights a key characteristic of the observed tapping frequency. Most tapping frequencies within a song are harmonically distanced, meaning a multiple/division by 2 to its nearest neighbour, to a very high precision. Most of the variation that exists is at a tolerance of less than 0.05Hz. When a tapping frequency is observed outside of this accuracy frame, there is usually only a single observation exclusively at that frequency. Furthermore, only 4 such observations exist in the dataset. These outliers tapping outside a direct harmonic placement suggests that these participants had a hard time keeping up with the beat and tapped with greater variability while attempting to keep up with the beat. Such variations can already be observed among the tapping profiles highlighted in Figure 23 and Figure 24.

In conclusion, we see the robustness and accuracy with which the FFT methodology extracts the tapping frequency. Analysing the data also verifies the underlying cognitive processes discussed in our review and highlights a pattern in the data which could be taken advantage of in our predictive modelling endeavour.

Music Information Retrieval - Beat Frequency

The primary link between the EEG and tapping data is the frequency at which the beat is perceived. This key information can be extracted directly from the music files themselves using the MIRTtoolbox (Lartillot & Toiviainen, 2007).

The beat frequency is derived from the tempo extracted from the audio files using the *mirtempo* function. Tempo can be extracted using the classical version of detecting periodicities in a

range of BPMs, or a metre-based version where tempo is extracted from the audio's hierarchical metrical structure (Lartillot & Toivainen, 2007; Lartillot, Eliard, Trost, Rappaz, & Grandjean, 2013). Beat vs Metre are interesting parameters to explore to tease apart their influences in the entrainment phenomenon.

The source material used for tempo extraction, is another feature worth exploring. If the extraction should be from the full song or strictly from the stimuli. This is interesting to contemplate the impact on the perception of an isochronous pulse or its accuracy derived from the full song itself or the specific portion used in the stimuli. It might inform to what the listener is directly responding to within the stimuli.

Song vs Stimuli Influences on Beat vs Metre

Analysing the extraction of beat and metre frequency from songs and stimuli highlights expected similarities while providing interesting discrepancies for discussion on the impact of their influences. Where extractions from songs leveraged information from the full song used during the EEG recording phase while, only 60s from within the stimuli that were used during the tapping exercise.

Figure 26, shows the beat extraction for both the songs and stimuli are the same. Only a slight variation is found for Song 1. This uniformity in signal is expected as the isochronous pulse present is to be a constant pulse throughout the song. This ultimately ties up the concerns on variations between the EEG signal stemming from a full song and the tapping exercise stemming from a partial stimulus. This supports the approach of analysing the EEG signal for the full song as it would provide a stronger signal at the beat frequency.

Figure 26. Similarities and differences in beat frequency extractions from songs & stimuli

The metre frequency extracted from songs and stimuli is shown in Figure 27. It highlights that the frequency observed by listeners could vary depending on the source. Half the songs did not maintain consistency between the songs and stimuli. When there was consistency, they did not always match with the expected beat frequency observed via the beat extraction methodology shown in Figure 27. This becomes important as the influence of a varying metre coupled with the phenomenon of perception and attention covered in our review, could have on the listener's perception and ultimately in their sensorimotor behaviour. It also highlights the potential considerations of the impact metre frequency and the proper source material has.

Figure 27. Metre frequency of songs & stimuli

The beat and metre frequencies compared against each other for both songs and stimuli, highlights the potential differences found in the hierarchical structure of the songs playing an influencing role. However, it is the perception of listeners from processing these details which ultimately impacts their responses during the tapping exercise.

Figure 28 highlights the significant differences both in terms of frequency among the songs between both these methodologies. While observing Figure 29, shows a more conforming result among the stimuli for both methodologies. This highlights the characteristics of metre being a heretical representation of the acoustic signal, playing a greater dynamic role across the full song versus shorter span within the stimuli. However, when there are differences among the stimuli representation, they are quite large. As observed for four songs, being quite significant in a 10 songs design protocol.

These differences within the four songs is a cause for pause. This suggests that the frequency being tapped to by the participants during the tapping exercise will potentially fall close to either approach. It can be observed that in the case of using metre as a reference, sourced from

the stimuli it is more likely representative of the listener's perception when it comes to tapping behaviours.

Figure 28. Beat and metre frequency of songs

Figure 29. Beat and metre frequency of stimuli

Having identified the most likely representative and performative combinations of song/stimuli and beat/metre, being beat extraction over full songs and metre extraction over stimuli, the task of selecting one approach over the idea is afoot. The critical comparison is between the frequencies extracted from either methodology aligning with the distribution among the tapping frequencies measured. And matching it with the learning from the reviews of current literature.

Figure 30, shows where the frequencies extracted align with at least one of the tapping clusters within a song or stimulus. The size of the cluster is representative of the number of participants tapping at that frequency. An important learning from this plot is that when the frequencies are extracted, they themselves do not always align to each other. The metre frequency from the stimuli aligns with the largest cluster twice (Songs 3 and 5) vs once for the beat frequency from songs (Song 10). Else they align at the largest cluster for 70% of the songs/stimuli.

Figure 30. Comparing tapping distribution against extracted frequency using beat and metre

Beat Extraction from Songs

In conclusion, to inform the decision on which approach to take, let us take the following factors into consideration: On examining the beat frequency from the songs and stimuli, there was almost no discrepancy. As expected from the perspective of a beat being an isochronous pulse generally persistent throughout a song. However, when dealing with metre it was significantly varying between songs and stimuli. As expected from the perspective of deriving those measures by breaking down the acoustic signal's hierarchical structure.

Taking the most representative versions as in beat with songs and metre with stimuli. Comparing these computed extractions to the tapping frequencies, highlights general similarity except for variations in primarily four songs. When considering the variations and alignment to tapping responses, metre for stimuli might have a stronger compelling reason to be most appropriate. But this would not continue to hold through when considering other details in the context of our objective.

When considering the nature of precepting beat or metre along with the phenomenon of neural entrainment, it certainly shows that beat lends itself to being more influential in the context of this study. Because studies on neural entrainment to acoustic stimuli including complex acoustic arrangements like music, has shown isochronous characteristics as evidenced by the review in current literature (Nozaradan, 2015; Nozaradan, Peretz, & Keller, 2016; Nozaradan, Peretz, Missal, & Mouraux, 2011). This is also evidenced by the huge variations observed when comparing metre extractions between a song and metre. Which means that if based on metre, the listener would not be able to maintain a single steady cyclic beat throughout the song. And in consideration with the fact that perception is going to be studied adopting EEG measurements and not only through tapping, requires a more consistent feature extraction methodology.

Another key factor that these difference between beat and metre will have lesser significance is in our robust approach in extracting features from EEG frequency signals. When referencing the beat frequency and expanding our scope of observance to harmonics and sub harmonics, the metre frequencies are also being captured, as observed in Figure 30.

These analysis and interpretations add credence to beat frequency taking precedence. This strategy also supports the approach to link both the tapping frequency and extracted beat frequency in informing the features to be extracted from EEG measurements. Though variations do exist, the dominance of beat frequency regardless of its presence in either the song or stimuli is evident.

The full song is utilised to extract beat frequency using MIR methodologies, and, along with the full EEG measurements, it is possible to maximise power of our extracted features. It is still possible that metre frequency could provide information in addition to or of a wholly different detail, worthy of exploration which is beyond the scope of this study.

Harmonizing Frequencies

The three key data required to build the predictive model are all in frequencies, namely the tapping frequency, beat frequency with its harmonics and EEG frequency. Harmonics were also investigated for the first as they were never explored in previous studies within this context. These different frequency data from various sources spread across participants and

songs, need to be normalised. Normalization occurs referencing to the beat frequency and its harmonics.

A critical challenge in building the model was how these harmonic features were linked together across songs and participants. Establishing the hegemony among the features extracted based on harmonics becomes critical to be able to study the responses for each song across participants interchangeably. To achieve this the beat frequency is first taken to be the fundamental frequency and then the harmonics are aligned to the EEG frequency data based on their distance to the nearest sampled frequency and finally categorised based on their harmonic index. This allowed the h^{th} harmonic to be universally compared with one another.

Figure 31, illustrates the process through which the harmonics are computed from the beat frequency which is extracted from the songs using MIR. As 20 sub-harmonics and harmonics were examined, this makes the 21st harmonic the fundamental frequency.

Figure 31: Flow of computing the fundamental frequency and its harmonics

Figure 32, shows the distribution of all tapping responses for all participants across all songs. Based off the harmonics for the beat frequency for each song, the tapping frequencies across

all participants and songs are recategorized into harmonics. Figure 33, shows the same data seen in Figure 32, converted to be represented by harmonics.

Figure 32: Frequency distribution of tapping responses of all participants across all songs

Examining the distribution of the tapping frequencies aligned to the harmonics as plotted in Figure 33, shows that majority of participants tapped on the fundamental frequency. In fact, most tapping aligned with the fundamental frequency or its immediate neighbour. Only outliers appeared outside of this ± 1 harmonics tolerance.

Examining Figure 33, also validates two critical understandings our approach is based on. Firstly, is that the fundamental frequency extracted by the MIR methodology is indeed representative to how most participants perceived and tapped along to. Second, the relationship between the fundamental frequency and the tapping frequency, as participants predominately tapped the harmonics or its nearest neighbour.

Figure 33: Frequency distribution of tapping responses of all participants across all songs

Another validation seen in the data that supports the approach of using harmonics, is that tapping responses overwhelmingly aligned with the beat frequency or one of its harmonics directly. Figure 34, shows that 90% of the data aligned to a harmonic which includes the beat frequency. This is in line with the hypothesis that there is a link which is directly derivable between the tapping frequency and the beat frequency extracted from music.

Figure 34: Proportion of tapping responses aligning with a harmonic

The final harmonization is on how features are to be extracted from a complex EEG signal on the frequency domain. Figure 35, illustrates how a subset of beat frequency and harmonics being extracted would be plotted set against an EEG measurement. The full extraction process will result in 41 values, with 20 sub-harmonic and harmonic values, as described in Figure 31.

Figure 35: Extracting features based on the beat frequency and its harmonics

EEG Analysis

The first processing step is the conversion of the EEG signals from the time domain and into the frequency domain. Key parameters relating to the conversions of all the EEG by song are as follows. The general bin size is between 0.0034-0.0037Hz, the effective frequency range captured is between 0-62.5Hz for all songs and the length of signal data was dependent on the length of the song offering between 16946-18611 data readings.

Our hypothesis and review provide us a basis for targeting specific frequencies when examining EEG data. Namely the beat frequency and its harmonics. This simplifies the process of exploring a complex EEG data, in order to correlate to cognitive processing and sensorimotor responses, which we want to predict.

A rudimentary analysis by observing the feature extractable based on the beat frequency, shows elevated power ratings. And this can be observed even when randomly examining nodes, identifying a pattern of interest directly related to our enquiry. Figure 36 illustrates this

relationship between beat frequencies extracted from the audio signals using MIR and signals observed at the specific frequency within the EEG measurement.

Figure 36: EEG responses at beat frequency for Song 1

Predictive Modelling

At this stage the dataset has been prepared for our predictive modelling. Adopting a multinomial logistic regression to build our model, it invariably requires the dependent variable and independent variable. In our model this being the tapping frequency and electrode/harmonics combination respectively. As highlighted in Figure 33, there are five possible categories covering harmonic 18-23. The model as illustrated in Figure 37, will be parameterised to accept the specific power rating at an electrode and harmonic frequency, and to verify against the tapping frequency. Each electrode and harmonic combination will be tested across all combinations of a Leave-One-Out methodology where the average performance across all iterations is reported, as described in Figure 18. The process will yield the performance for each combination from which the most accurate model can be selected. The objective being to identify the most accurately predicting electrode and harmonic combination.

Considering all the electrode and harmonic combinations in consideration a total of 125 Electrodes X 40 Harmonics will be considered, totalling to 5000 models per participants. With

a Leave-One-Out training and validation approach as illustrated in Figure 18, each iterative run will train on as much as 5000 Electrode/Harmonic X 9 Participants, totalling to 45000 models. The range of observance was generously selected based on review and understanding of analysis conducted thus to cast a wide enough net to encompassing the variability in perceptions.

Figure 37: Process flow depicting the modelling effort

Predictive Modelling Profile

Figure 38-Figure 40, shows the profile of the accuracy of each electrode/harmonic model trained and validated cycling through all iterations of a Leave-One-Out methodology. These iterations here refer to participants being left out, where each iteration cycles through all tapping responses for 10 songs of a single participant to be used to validate the model for that cycle.

Generally, the performance is higher between harmonic 8-18, across all electrodes. A reason for the general high performance among all the models producing an on an average baseline accuracy of over 0.67 is because of the skewed nature of the tapping responses against which the model is trained and validated against. Thus, it is harder to tease apart the performance of top performing models with robust characteristics.

Figure 38: Mean accuracy profile by participant - left view

Figure 39: Mean accuracy profile by participant - right view

Figure 40: Mean accuracy profile by participant - top view

Best Performing Models

There were 2 models that performed with the highest accuracy being at 0.74, via LnO-CV. These models used features from electrode 86 and 79, at the 12th and 13th harmonics respectively, performing models are displayed in Table 3. Each with a standard deviation of 0.19.

Best Performing Models		
Electrode/Harmonic	86/12	79/13
Accuracy	0.74	0.74

Table 3: Accuracy performance of the best models

Figure 41-Figure 42 provides a topographical illustration on where the best performing electrodes were located. These were capturing signals from premotor cortex, aligning with literature on cognitive processing of beat processing.

Figure 41: 3D topography of electrodes - side view

Figure 42: 3D topography of electrodes - top view

Chapter 6. Validation

Taking a modelling approach of directly using EEG power rating to be predictive of tapping behavioural responses, validation analysis can be rigorously performed using these primary sources of the modelling process to validate its predictive capability.

Our validation approach is based on Naylor's and Finger's three step approach (Naylor & Finger, 1967). Dealing with 1) The model's face validity, 2) The model's structural assumption and 3) The model's data assumption.

Validity

The primary concept of the model is to monitor the frequency of EEG signals at specific harmonics of a song's beat frequency to educate us on the behavioural responses of listeners. This concept and approach were derived from literature reviews covering works providing evidence from the perspectives of both neuroscientific and cognitive processing. These studies covered in the literature reviews have also shown evidences of consistency in terms of expected results and behaviours when testing out hypothesis on neural entrainment, beat perception and sensorimotor responses.

Based on these findings, a hypothetical model was proposed and build with strong evidence of predictability.

Moving further into the validation process, similar models can be generated with face validity by randomising these data sources.

Structural Assumption

The model's structural assumption is verifiable by comparing the performance of our model using a single feature vs multiple feature. As these features were selected based on the theories supported by literature reviews detailed in this thesis. The features selected by this work is derived from performances at best identifying specific harmonic categories. Thus, by aiming to add these features we should see the overall performance of the models generated perform better than the single feature models. And these improvements in performance will support the hypothesis used to motivate these selections.

This exploration was not examined in depth to remain focus on actual model developed and validating its premise through data assumption.

Data Assumption

Two approaches can be taken to validate the data assumptions used to develop this model. The first to manipulating the tapping responses within a participant's own responses. The other being the actual EEG power measurements extracted across participants and songs.

The first approach is to randomise the tapping responses within each participant. Such that the user's tapping responses for a song to be interpreted as the response for another song by the same participant. It is important to note that this approach will cause the least amount of variation when compared to the real data, as most participants tapped at the beat frequency.

The second approach to maximise variability to the real data is to randomise the EEG power ratings used. This time the randomisation is expanded across both participants and songs. Such that a participant's EEG rating at a specific harmonic can be interpreted as the EEG rating for another participant for a different song and at another harmonic.

The main purpose of these two approaches is to highlight the extreme performance of the model after a best/worst case data randomisation approach.

As majority of the tapping responses fall within the beat frequency, the model's validation via these responses will produce limited variability when compared to the model using the real data. Variability in terms of the best performing models and its overall accuracy and recognition performance. Figure 33, shows that most of the tapping responses were at the beat frequency leaving little variability in the dataset overall such that even if randomising would not have a strong influence in such a small dataset. Still validation can be fruitfully derived understanding the dataset and principles behind the modelling process.

Tapping Responses

By randomising tapping responses within individual participants we can validate the significance of tapping responses in building the predictive model. It will also demonstrate based on the nature of the dataset how influential this component is.

Figure 43-Figure 45, shows the profile of the accuracy of each electrode/harmonic model trained and validated cycling through all iterations of a Leave-One-Out methodology. The surface plots clearly show a dampening in performance activity outside the key region of interest. This makes sense as most of these low performing electrode/harmonic models gets silenced when the key tapping responses data gets randomized leaving very little information available to build a predictive model.

Figure 43: Mean accuracy profile by participant (Randomized tapping) - left view

Figure 44: Mean accuracy profile by participant (Randomized tapping) - right view

Figure 45: Mean accuracy profile by participant (Randomized tapping) - top view

As expected, the variability in best performances are low, still reduction is shown in overall accuracy. Primarily due to the skewed nature and low variability of the tapping responses. Observance of the performance of the two best predictive models based on the actual data now using randomized tapping data, is shown in Table 4 below. The two best performing models from the actual dataset (Table 3) now performing significantly worse.

Actual Best Performing Models on Randomized Tapping Data		
Electrode/Harmonic	86/12	79/13
Accuracy	0.69	0.68

Table 4: Performance of actual best model (Randomized tapping)

Significant variability is observed among the general performance of all the models as shown in the differences between Figure 38-Figure 40 and Figure 43-Figure 45, and the best performing models observed as shown in the differences between Table 3 and Table 4. This signifies the importance of tapping responses play in the predictive nature of the model and validates the approach from the perspective of tapping responses.

EEG Responses

By completely randomising the EEG responses we can validate the significance of these power rating in predicting participant responses. Further demonstrating the link between these features and the tapping responses of participants. And as expected the greater influence this would have on predictability as compared to our strategy of randomizing tapping responses.

Figure 46-Figure 49, shows the profile of the accuracy of each electrode/harmonic model trained and validated cycling through all iterations of a Leave-One-Out methodology. The surface plots clearly show erratic performance behaviour and this time eliminating the key region of interest observed in the true data modelling. Specifically, on the general performance of accuracy observed between harmonic 8-18, as shown in Figure 38-Figure 40.

Figure 49, shows an interesting outcome where there are greater degenerative spikes throughout the landscape as would be expected when the links between EEG power to tapping responses has been eradicated by randomization.

Figure 46: Mean accuracy profile by participant (Randomized EEG) - left view

Figure 47: Mean accuracy profile by participant (Randomized EEG) - right view

Figure 48: Mean accuracy profile by participant (Randomized EEG) - top view

Figure 49: Mean accuracy profile by participant (Randomized EEG) – bottom left view

As expected, the variability in best performances are higher, with significant degradation in overall accuracy. Primarily due to the significant role and variability among the dataset observed among the EEG data. Observance of the performance of the two best predictive models based on the actual data now using randomized EEG data, is shown in Table 5 below. The two best performing model from the actual dataset (Table 3) now performed significantly worse and closer to the average baseline.

Actual Best Performing Models on Randomized EEG Data		
Electrode/Harmonic	86/12	79/13
Accuracy	0.67	0.69

Table 5: Performance of actual best model (Randomized EEG)

Comparatively to randomizing tapping responses, even greater variability is observed using randomised EEG data among the general performance of all the models as shown in the differences between Figure 38-Figure 40 and Figure 46-Figure 49, and the best performing models observed as shown in the differences between Table 3 and Table 5. This signifies the importance of EEG frequency power measurements play in the predictive nature of the model and validates the approach from the perspective of EEG data.

Chapter 7. Conclusion & Future Work

Conclusion

This study focused on many premises through the efforts of building a predictive model. The capability to use EEG signals to predict future sensorimotor behaviours specifically tapping responses. Utilising the phenomena of neural entrainment to identify beat frequency and the role its harmonics play in processing acoustic signals ultimately represented in sensorimotor activities like tapping along to its beat. The salient role of neural entrainment found within cognitive processing of beats and metres, and the role harmonics plays in these.

The results convincingly displayed that a predictive model can be generated with applicable accuracy that a listener's perception of beat frequency can be predicted based on EEG measurement extracted at the frequencies of the harmonics. The best predictive model had an accuracy of 0.74 with standard deviation of 0.19. In fact, the results showed that the harmonics were much more predictive of tapping responses compared to the beat frequency, even though beat frequency has been predominantly used in such investigations in current literature, please refer to Beat Perception and Sensorimotor Synchronization.

These findings are significant as this demonstrates the validity and potential gain in exploring harmonics in addition to the fundamental frequency in BCI analysis or when investigating entrainment. It also clearly demonstrates the generalizability of EEG measurements to predict future processes when it comes to neural entrainment and beat perceptions. An example would be to explore the formation of harmonic traces when imaginary beats were being explored as in (Nozaradan, 2015) study. (Nozaradan, Peretz, & Keller, 2016) effort on the other hand found direct links between cortical and rhythmic entrainment drawing the link that frequency-tagged brain activities are functionally relevant for beat perception, it could be informative to learn if new links are discovered when examined via harmonics. Another study from (Nozaradan, Peretz, & Mouraux, 2012) focused on the perception of periodicities by investigating the neuronal entrainment to beats and metres could also benefit from extending that explorations to harmonics.

Another interesting learning from this study was the possible variations observed between extracted beat and metre frequencies when examining the full songs or the stimuli excerpts. But despite these variabilities these frequencies always aligned on a fundamental frequency or

on one of its harmonics, supporting further for the consideration of harmonics when studying beat perception or neural entrainment. Generally, beats and metres were often conflated, even though MIR extractions focus on different conceptual elements within the acoustic signal. They both could be used intentionally to fine tune the focus of one's investigations or predictions, where each parameter serves its own unique purpose.

Another valuable finding is that the FFT approach adopted to identifying the tapping frequency is much more robust than the median approach that is commonly used for identifying tapping frequency. This study has provided ample evidence and justification for future studies to adopt this approach, rather than median.

The predictive models were most performative from EEG measurements taken from electrodes 86 and 79. Figure 41-Figure 42 highlights that these electrodes best represent readings from the premotor cortex. Validating the link and role the premotor cortex plays in beat perception and why the model can predict future sensorimotor responses to those cognitive processing accurately.

A deeper question and analysis need to be made into the cognitive processing of neural entrainment and beat processing in reference to the stimuli being processed. It was observed that the distribution of tapping responses for the same stimuli among participants was almost always present, though it usually only extended to be generally apart by a harmonic. In fact, only Song 2 had the distinct characteristic where all 20 participants coherently tapped along to at the same frequency. What causes this variation and how can it be measured or predicted when a variation occurs in perception. Harmonics could play a vital role in such an investigation.

Limitations

The biggest limitation in our attempt to achieve a predictive model was the small dataset, both in terms of participants and songs. This severely limited enough response data in the less frequently tapped harmonics, inhibiting a resilient model. The lack of data also exasperated the skewed nature of the data, further reducing the resiliency of the model performing well in categories outside the norm of the most frequently tapped responses. Nonetheless, the effort was successful in practically proving that a predictive model was achievable. There are strong evidences of a more resilient model being achievable with a larger and diverse dataset,

especially with a more robust modelling methodology, such as adopting machine learning techniques.

Future Research

Current stimuli flattened the stereo dynamics of the music into a mono output. By adding this additional layer of dynamics, could differences be observed in the cognitive ability of processing these complex stimuli? And how would these manifests physiologically?

A Multinomial Logistic Regression was adopted to build the model, but more sophisticated machine learning models could also be explored like Residual Neural Network. Lasso Regression could also be used to minimize to effective features.

Another approach which could have been explored to improve predictiveness and distinctiveness among tapping responses is by collating EEG data with a more distinctive feature to learn from. Literature of neuroimaging studies have shown that musicians have attenuated signals when listening to music (Grahn & Brett, Rhythm and Beat Perception in Motor Areas of the Brain, 2007). Such captured signals could have provided more predictive prowess in building a base model to be translatable to non-musician listeners.

Our approach in adopting the frequency domain for our analysis was based on previous research by (Nozaradan, Peretz, & Mouraux, 2012; Zhou, Melloni, Poeppel, & Ding, 2016). Its advantages for quick isolation of features for exploration and direct observability of strong signals at calculable frequency provided an excellent premise for exploratory modelling effort. Though a time domain analysis as conducted by (Rajendran, Harper, Garcia-Lazaro, Lesica, & Schnupp, 2017) could also be conducted to understand signal behaviour over a wider spectrum while subverting challenges presented in exploring similar broader spectrum in the frequency domain (Zhou, Melloni, Poeppel, & Ding, 2016).

The stimuli in this study were primarily chosen for its potential unfamiliarity among listeners and its unambiguity in discerning a beat. Understanding the cognitive processes involved in discerning a beat in a stimulus, it will be interesting to explore how if the stimuli varied not only in terms of their measured frequency but their inherent characteristics cues such as syncopated pulses. For example, stimuli having different accent types. Accents can be dynamic, determined by change in volume or subjectively with no discernible external cues forcing listeners to emphasise tones in the stimuli. The degree in which these vary can also be

classified in terms of regular volume accents having a prominent beat, duration accents having weaker beats or even those with no accents having no externally discernible beat having to be internally calculated. Figure 50, shows examples of such stimuli depictions from the study conducted by (Grahn J. A., 2009).

Figure 50: Schematic depiction of audio stimuli. Left: The Auditory Soundwave. Right: The music notation (Grahn J. A., 2009)

Epilogue

This work made multiple valuable findings and contributions to the current body of literature. Through the development of a predictive model it showed that 1) EEG measurements of processing acoustic signals can be generalized and used to predict future perceptions of acoustic signals. Together with MRI feature extractions from acoustic signals as the ground truth, the combinations are able to reveal critical knowledge for predictive efforts, 2) Harmonics carry a lot of predictive power and play a significant role in understanding cognitive processing of neural entrainment and beat perception, which are being always neglected in current literature, 3) A robust method for identifying tapping frequency accurately even when challenged by individual variability and ability to process, correct and maintain synchronised tapping to a stimulus, 4) The roles beat and metre play in neural entrainment and beat

perception along with the understanding that harmonics can also play a huge part in understanding or bridging these differences.

References

- Bartolo, R., & Merchant, H. (2015). β Oscillations Are Linked to the Initiation of Sensory-Cued Movement Sequences and the Internal Guidance of Regular Tapping in the Monkey. *The Journal of Neuroscience*, 35(11), 4635–4640. doi:<https://dx.doi.org/10.1523/JNEUROSCI.4570-14.2015>
- Bengtsson, S., Ullén, F., Ehrsson, H., Hashimoto, T., Kito, T., Naito, E., . . . Sadato, N. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex*, 45(1), 62–71. doi:<https://doi.org/10.1016/j.cortex.2008.07.002>
- Buzsáki, G. (2009). *Rhythms of the Brain*. doi:10.1093/acprof:oso/9780195301069.001.0001
- Chapin Heather, Z. T. (2010). Neural Responses to Complex Auditory Rhythms: The Role of Attending. *Frontiers in Psychology*, 1, 224. doi:<https://doi.org/10.3389/fpsyg.2010.00224>
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to Musical Rhythms Recruits Motor Regions of the Brain. *Cerebral Cortex*, 18(12), 2844–2854. doi:<https://doi.org/10.1093/cercor/bhn042>
- Comstock, D. C., Hove, M. J., & Balasubramaniam, R. (2018). Sensorimotor Synchronization With Auditory and Visual Modalities: Behavioral and Neural Differences. *Frontiers in Computational Neuroscience*, 12, 53. doi:<https://doi.org/10.3389/fncom.2018.00053>
- Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California Sea Lion (*Zalophus californianus*) Can Keep the Beat: Motor Entrainment to Rhythmic Auditory Stimuli in a Non Vocal Mimic. *Journal of Comparative Psychology*, 127(4), 412–427. doi:<https://doi.org/10.1037/a0032345>
- Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). California sea lion (*Zalophus californianus*) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *American Psychological Association*, 127(4), 412–427. doi:<https://content.apa.org/doi/10.1037/a0032345>
- Coull, J. T., Cheng, R.-K., & Meck, W. H. (2011). Neuroanatomical and Neurochemical Substrates of Timing. *Neuropsychopharmacology*, 36, 3–25. doi:10.1038/npp.2010.113

- Coull, J. T., Vidal, F., & Burle, B. (2016). When to act, or not to act: that's the SMA's question. *Current Opinion in Behavioral Sciences*, 8, 14-21. doi:<https://doi.org/10.1016/j.cobeha.2016.01.003>
- Darwin, C. (1871). *The descent of man, and Selection in relation to sex, Vol 1* (Vol. 1). doi:<https://content.apa.org/doi/10.1037/12293-000>
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134, 9-21.
- Drakea, C., RiessJones, M., & Barucha, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, 77(3), 251-288. doi:[https://doi.org/10.1016/S0010-0277\(00\)00106-2](https://doi.org/10.1016/S0010-0277(00)00106-2)
- Drewing, K., Aschersleben, G., & Li, S.-C. (n.d.). Sensorimotor synchronization across the life span. *International Journal of Behavioral Development*, 30(3), 280–287. doi:<https://psycnet.apa.org/doi/10.1177/0165025406066764>
- Durkheim, E. (1965). *The Elementary Forms of the Religious Life: A Study in Religious Sociology*. Macmillan.
- Ehrenreich, B. (2006). *Dancing in the streets: A history of collective joy*. Metropolitan Books.
- Eliade, M. (1964). *Shamanism: Archaic Techniques of Ecstasy*. Librarie Payot.
- Elliott, M. T., Welchman, A. E., & Wing, A. M. (2009). MatTAP: A MATLAB Toolbox for the Control and Analysis of Movement Synchronisation Experiments. *Journal of Neuroscience Methods*, 177(1), 250-257. doi:<https://doi.org/10.1016/j.jneumeth.2008.10.002>
- Finney, S. A. (2001). FTAP: A Linux-based program for tapping and music experiments. *Behavior Research Methods, Instruments, & Computers*, 33(1), 65-72. doi:<https://doi.org/10.3758/BF03195348>
- Fitch, W. T. (2013). Rhythmic Cognition in Humans and Animals: Distinguishing Meter and Pulse Perception. *Frontiers in Systems Neuroscience*, 7(68). doi:<https://doi.org/10.3389/fnsys.2013.00068>
- Fraisse, P. (1982). Rhythm and Tempo. In D. Deutsch, *Psychology of Music* (pp. 149-180).

- Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and Gamma Rhythms in Human Auditory Cortex During Musical Beat Processing. *The Neurosciences and Music III: Disorders and Plasticity*, 1169, 89-92. doi:<https://doi.org/10.1111/j.1749-6632.2009.04779.x>
- Geiser, E., Notter, M., & Gabrieli, J. D. (2012). A Corticostriatal Neural System Enhances Auditory Perception through Temporal Context Processing. *The Journal of Neuroscience*, 32(18), 6177–6182. doi:<https://dx.doi.org/10.1523/JNEUROSCI.5153-11.2012>
- Grafton, S. T., Hazeltine, E., & Ivry, R. (2007). Functional Mapping of Sequence Learning in Normal Humans. *Journal of Cognitive Neuroscience*, 7(4), 497-510. doi:<https://doi.org/10.1162/jocn.1995.7.4.497>
- Grahn, J. A. (2009). The Role of the Basal Ganglia in Beat Perception: Neuroimaging and Neuropsychological Investigations. *Annals of the New York Academy of Sciences*, 1169(1). doi:<https://doi.org/10.1111/j.1749-6632.2009.04553.x>
- Grahn, J. A., & Brett, M. (2007). Rhythm and Beat Perception in Motor Areas of the Brain. *Journal of Cognitive Neuroscience*, 19(5), 893-906. doi:<https://doi.org/10.1162/jocn.2007.19.5.893>
- Grahn, J. A., & Rowe, J. B. (2009). Feeling the Beat: Premotor and Striatal Interactions in Musicians and Nonmusicians during Beat Perception. *Journal of Neuroscience*, 29(23), 7540-7548. doi:<https://doi.org/10.1523/JNEUROSCI.2018-08.2009>
- Grahn, J. A., & Rowe, J. B. (2013). Finding and Feeling the Musical Beat: Striatal Dissociations between Detection and Prediction of Regularity. *Cerebral Cortex*, 23(4), 913–921. doi:<https://doi.org/10.1093/cercor/bhs083>
- Haidt, J., Seder, J. P., & Kesebir, S. (2008). Hive Psychology, Happiness, and Public Policy. *The Journal of Legal Studies*, 37(2), 133-156. doi:<https://doi.org/10.1086/529447>
- Hannona, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology*, 50(4), 354-377. doi:<https://doi.org/10.1016/j.cogpsych.2004.09.003>

- Harrington, D. L., Zimbelman, J. L., Hinton, S. C., & Rao, S. M. (2010). Neural Modulation of Temporal Encoding, Maintenance, and Decision Processes. *Cerebral Cortex*, 20(6), 1274–1285. doi:<https://doi.org/10.1093/cercor/bhp194>
- Hasegawa, A., Okanoya, K., Hasegawa, T., & Seki, Y. (2011). Rhythmic synchronization tapping to an audio–visual metronome in budgerigars. *Scientific Reports*, 1(1), 120. doi:<https://doi.org/10.1038/srep00120>
- Herff, S. A., Herff, C., Milne, A. J., Johnson, G. D., Shih, J. J., & Krusienski, D. J. (2020). Prefrontal high gamma in ecog tags periodicity of musical rhythms in perception and imagination. *eNeuro*, 7(4), 413-419. doi:<https://dx.doi.org/10.1523%2FENEURO.0413-19.2020>
- Hove, M. J., Fairhurst, M. T., Kotz, S. A., & Keller, P. E. (2013). ynchronizing with auditory and visual rhythms: An fMRI assessment of modality differences and modality appropriateness. *NeuroImage*, 67(15), 313-321. doi:<https://doi.org/10.1016/j.neuroimage.2012.11.032>
- Hulse, S. H., Bernard, D. J., & Braaten, R. F. (1995). Auditory discrimination of chord-based spectral structures by European starlings (*Sturnus vulgaris*). *Journal of Experimental Psychology: General*, 124(4), 409–423. doi:<https://doi.apa.org/doi/10.1037/0096-3445.124.4.409>
- Iversen, J. R., & Balasubramaniam, R. (2016). Synchronization and temporal processing. *Current Opinion in Behavioral Sciences*, 8, 175-180. doi:<https://doi.org/10.1016/j.cobeha.2016.02.027>
- Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down Control of Rhythm Perception Modulates Early Auditory Responses. *Annals of the New York Academy of Sciences*, 1169(1), 58-73. doi:<https://doi.org/10.1111/j.1749-6632.2009.04579.x>
- Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology*, 141(1), 54–75. doi:<https://doi.apa.org/doi/10.1037/a0024208>
- Janik, V. M., & Slater, P. J. (1997). Vocal Learning in Mammals. *Advances in the Study of Behaviour*, 26, 59–99. doi:[https://doi.org/10.1016/S0065-3454\(08\)60377-0](https://doi.org/10.1016/S0065-3454(08)60377-0)

- Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics*, 41, 621–634. doi:<https://doi.org/10.3758/BF03210494>
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96(3), 459–491. doi:<https://content.apa.org/doi/10.1037/0033-295X.96.3.459>
- Kim, H.-S., Kaneshiro, B., & Berger, J. (2012). Tap-It: An iOS App for Sensori-Motor Synchronization Experiments. *12th International Conference on Music Perception and Cognition*.
- Kotz, S., Ravignani, A., & Fitch, W. (2018). The Evolution of Rhythm Processing. *Trends in Cognitive Sciences*, 22(10), 896-910. doi:<https://doi.org/10.1016/j.tics.2018.08.002>
- Kung, S.-J., Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2013). Interacting Cortical and Basal Ganglia Networks Underlying Finding and Tapping to the Musical Beat. *Journal of Cognitive Neuroscience*, 25(3), 401-420. doi:https://doi.org/10.1162/jocn_a_00325
- Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action Representation of Sound: Audiomotor Recognition Network While Listening to Newly Acquired Actions. *Journal of Neuroscience*, 27(2), 308-314. doi:<https://doi.org/10.1523/JNEUROSCI.4822-06.2007>
- Large, E. W. (2014). *Rhythm Perception: Pulse and Meter*. doi:DOI 10.1007/978-1-4614-7320-6_106-2
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1), 119–159. doi:<https://doi.apa.org/doi/10.1037/0033-295X.106.1.119>
- Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, 26(1), 1–37. doi:https://doi.org/10.1207/s15516709cog2601_1
- Large, E. W., Almonte, F. V., & Velasco, M. J. (2010). A canonical model for gradient frequency neural networks. *Physica D: Nonlinear Phenomena*, 239(12), 905-911. doi:<https://doi.org/10.1016/j.physd.2009.11.015>
- Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural Networks for Beat Perception in Musical Rhythm. *Frontiers in Systems Neuroscience*, 9(159). doi:10.3389/fnsys.2015.00159

- Large, E., & Snyder, J. (2009). Pulse and meter as neural resonance. *The Neurosciences and Music III Disorders and Plasticity*, 1169(1), 46-57. doi:doi: 10.1111/j.1749-6632.2009.04550.x.
- Larsson, M., Richter, J., & Ravignani, A. (2019). Bipedal Steps in the Development of Rhythmic Behavior in Humans. *Music & Science*, 2(1), 1-14. doi:https://doi.org/10.1177%2F2059204319892617
- Lartillot, O. |., Eliard, K., Trost, W. J., Rappaz, M.-A., & Grandjean, D. (2013). Estimating Tempo and Metrical Features by Tracking the Whole Metrical Hierarchy. *The 3rd International Conference on Music & Emotion*.
- Lartillot, O., & Toiviainen, P. (2007). A Matlab Toolbox For Musical Feature Extraction From Audio. *Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07)*, (pp. 10-15).
- Leaver, A. M., Lare, J. V., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). Brain Activation during Anticipation of Sound Sequences. *Journal of Neuroscience*, 29(8), 2477-2485. doi:https://doi.org/10.1523/JNEUROSCI.4921-08.2009
- Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. New York, NY: Oxford University Press.
- London, J. (2012). *Hearing in Time: Psychological Aspects of Musical Meter*. Oxford University Press.
- London, J. M. (2004). *Hearing in Time: Psychological Aspects of Musical Meter*. New York, NY: Oxford University Press.
- Losorelli, S., Nguyen, D. T., Dmochowski, J. P., & Kaneshiro, B. (2017). Naturalistic Music EEG Dataset - Tempo (NMED-T). *Stanford Digital Repository*.
- Losorelli, S., Nguyen, D. T., Dmochowski, J. P., & Kaneshiro, B. (2017). NMED-T: A tempo-focused dataset of cortical and behavioral responses to naturalistic music. *Proceedings of the 18th ISMIR Conference*, (pp. 339-346). Suzhou, China.
- Martens, P. A. (2011). The Ambiguous Tactus: Tempo, Subdivision Benefit, And Three Listener Strategies. *Music Perception*, 28(5), 433-448. doi:https://doi.org/10.1525/mp.2011.28.5.433

- McKinney, M. F., & Moelants, D. (2006). Ambiguity in Tempo Perception: What Draws Listeners to Different Metrical Levels? *Music Perception*, 24(2), 155–166. doi:<https://doi.org/10.1525/mp.2006.24.2.155>
- McNeill, W. H. (1995). *Keeping Together in Time: Dance and Drill in Human History*. Harvard University Press.
- Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, 7, 274. doi:<https://doi.org/10.3389/fnins.2013.00274>
- Merchant, H., & Yarrow, K. (2016). How the motor system both encodes and influences our sense of time. *Current Opinion in Behavioral Sciences*, 22–27. doi:<https://doi.org/10.1016/j.cobeha.2016.01.006>
- Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. *Philosophical transactions of the Royal Society of London*, 370(1664). doi:<https://dx.doi.org/10.1098%2Frstb.2014.0093>
- Moelants, D. (2002). Preferred Tempo Reconsidered. *Proceedings of the 7th International Conference on Music Perception and Cognition*, (pp. 580-583). Sydney.
- Monier, F., & Droit-Volet, S. (2019). Development of Sensorimotor Synchronization Abilities: Motor and Cognitive Components. *Child Neuropsychology*, 25(8), 1043-1062. doi:<https://doi.org/10.1080/09297049.2019.1569607>
- Nagasaka, Y., Chao, Z. C., Hasegawa, N., Notoya, T., & Fujii, N. (2013). Spontaneous synchronization of arm motion between Japanese macaques. *Scientific Reports*, 3(1). doi:<https://doi.org/10.1038/srep01151>
- Naylor, T. H., & Finger, J. M. (1967). Verification of Computer Simulation Models. *Management Science*, 14(2), B-1-B-139. doi:<https://doi.org/10.1287/mnsc.14.2.B92>
- Noorden, L. v., & Moelants, D. (1999). Resonance in the Perception of Musical Pulse. *Journal of New Music Research*, 28(1), 43-66.
- Nozaradan, S. (2015). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1658). doi:10.1098/rstb.2013.0393

- Nozaradan, S., Peretz, I., & Keller, P. E. (2016). Individual Differences in Rhythmic Cortical Entrainment Correlate With Predictive Behavior in Sensorimotor Synchronization. *Scientific reports*, 6. doi:<https://doi.org/10.1038/srep20612>
- Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective Neuronal Entrainment to the Beat and Meter Embedded in a Musical Rhythm. *Journal of Neuroscience*, 32(49), 17572-17581. doi:<https://doi.org/10.1523/JNEUROSCI.3203-12.2012>
- Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the Neuronal Entrainment to Beat and Meter. *The Journal of Neuroscience*, 31(28), 10234–10240. doi:10.1523/JNEUROSCI.0411-11.2011
- Nozaradan, S., Schwartze, M., Obermeier, C., & Kotz, S. A. (2017). Specific Contributions of Basal Ganglia and Cerebellum to the Neural Tracking of Rhythm. *Cortex*, 95, 156-168. doi:<https://doi.org/10.1016/j.cortex.2017.08.015>
- Okawa, H., Suefusa, K., & Tanaka, T. (2017). Neural Entrainment to Auditory Imagery of Rhythms. *Frontiers in Human Neuroscience*, 11, 493. doi:<https://doi.org/10.3389/fnhum.2017.00493>
- Oller, D. K., & Griebel, U. (2008). *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication*. doi:10.7551/mitpress/9780262151214.003.0003
- Patel, A. D. (2006). Musical Rhythm, Linguistic Rhythm, and Human Evolution. *Music Perception*, 24(1), 99–104. doi:<https://doi.org/10.1525/mp.2006.24.1.99>
- Patel, A. D. (2014). The Evolutionary Biology of Musical Rhythm: Was Darwin Wrong? *PLOS BIOLOGY*. doi:<https://doi.org/10.1371/journal.pbio.1001821>
- Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, 8, 57. doi:<https://doi.org/10.3389/fnsys.2014.00057>
- Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the Beat: Movement Influences Infant Rhythm Perception. *Science*, 308(5727), 1430. doi:<https://doi.org/10.1126/science.1110922>

- Pineda, J. A. (2005). The Functional Significance of Mu Rhythms: Translating "Seeing" and "Hearing" Into "Doing". *Brain Research Reviews*, 50(1), 57-68. doi:10.1016/j.brainresrev.2005.04.005
- Povel, D.-J., & Essens, P. (1985). Perception of Temporal Patterns. *Music Perception: An Interdisciplinary Journal*, 2(4), 411-440. doi:https://psycnet.apa.org/doi/10.2307/40285311
- Prado-Gutiérrez, P., Otero, M., Martínez-Montes, E., Weinstein, A., Escobar, M.-J., El-Deredy, W., & Zañartu, M. (2019). A Method for Tracking the Time Evolution of Steady-State Evoked Potentials. *Journal of Visualized Experiments*, 147. doi:https://dx.doi.org/10.3791/59898
- Provasi, J., Anderson, D. I., & Barbu-Roth, M. (2014). Rhythm perception, production, and synchronization during the perinatal period. *Frontiers in Psychology*, 5(1048). doi:https://dx.doi.org/10.3389%2Ffpsyg.2014.01048
- Rajendran, V. G., Harper, N. S., Garcia-Lazaro, J. A., Lesica, N. A., & Schnupp, J. W. (2017). Midbrain adaptation may set the stage for the perception of musical beat. *Proc. R. Soc.* 284. Royal Society Publishing. doi:http://dx.doi.org/10.1098/rspb.2017.1455
- Rankin, S. K., Large, E. W., & Fink, P. W. (2009). Fractal Tempo Fluctuation and Pulse Prediction. *Music Perception*, 26(5), 401-413. doi:https://doi.org/10.1525/mp.2009.26.5.401
- Rankin, S. K., Large, E. W., & Fink, P. W. (2009). Fractal Tempo Fluctuation and Pulse Prediction. *Music Perception*, 26(5), 401-413. doi:https://dx.doi.org/10.1525%2Fmp.2009.26.5.401
- Rauschecker, J. P., & Scott, S. K. (2009). Maps and Streams in the Auditory Cortex: Nonhuman Primates Illuminate Human Speech Processing. *Nature Neuroscience*, 12(6), 718-24. doi:https://doi.org/10.1038/nn.2331
- Repp, B. H. (2001). Processes underlying adaptation to tempo changes in sensorimotor synchronization. *Human Movement Science*, 20(3), 277-312. doi:https://doi.org/10.1016/S0167-9457(01)00049-5

- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6), 969-992. doi:<https://doi.org/10.3758/bf03206433>
- Repp, B. H. (2007). Hearing a melody in different ways: Multistability of metrical interpretation, reflected in rate limits of sensorimotor synchronization. *102*(3), 434-454. doi:<https://doi.org/10.1016/j.cognition.2006.02.003>
- Repp, B. H., & Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: Effects of intention, attention, and awareness. *The Quarterly Journal of Experimental Psychology*, 57(3), 499-521. doi:<https://doi.org/10.1080/02724980343000369>
- Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, 20(3), 403-452. doi:<https://doi.org/10.3758/s13423-012-0371-2>
- Repp, B. H., Iversen, J. R., & Patel, A. D. (2008). Tracking an Imposed Beat within a Metrical Grid. *Music Perception*, 26(1), 1-18. doi:<https://doi.org/10.1525/mp.2008.26.1.1>
- Robert Galambos, S. M. (1981). A 40-Hz auditory potential recorded from the human scalp. *Proceedings of National Academy of Sciences of the United States of America*, 78 (4), pp. 2643-2647. doi:<https://doi.org/10.1073/pnas.78.4.2643>
- Ross, J. M., Iversen, J. R., & Balasubramaniam, R. (2016). Motor simulation theories of musical beat perception. *Neurocase*, 22(6), 558-565. doi:<https://doi.org/10.1080/13554794.2016.1242756>
- Ross, J. M., Iversen, J. R., & Balasubramaniam, R. (2018). The Role of Posterior Parietal Cortex in Beat-based Timing Perception: A Continuous Theta Burst Stimulation Study. *Journal of Cognitive Neuroscience*, 30(5), 634-643. doi:https://doi.org/10.1162/jocn_a_01237
- Rowe, J. A. (2009). Feeling the Beat: Premotor and Striatal Interactions in Musicians and Nonmusicians during Beat Perception. *Journal of Neuroscience*, 29(23), 7540-7548. doi:<https://doi.org/10.1523/JNEUROSCI.2018-08.2009>

- Schroeder, C. E., & Lakatos, P. (2008). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, 31(1), 9-18. doi:<https://doi.org/10.1016/j.tins.2008.09.012>
- Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, 32(1), 9-18. doi:<https://doi.org/10.1016/j.tins.2008.09.012>
- Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, 11(5). doi:<https://doi.org/10.1016/j.tics.2007.02.006>
- Schusterman, R. J., & Reichmuth, C. (2008). Novel sound production through contingency learning in the PaciWc walrus (*Odobenus rosmarus divergens*). *Animal Cognition*, 11, 319–327. doi:[10.1007/s10071-007-0120-5](https://doi.org/10.1007/s10071-007-0120-5)
- Schwartz, M., Rothermich, K., & A.Kotz, S. (2012). Functional dissociation of pre-SMA and SMA-proper in temporal processing. *NeuroImage*, 60(1), 290-298. doi:<https://doi.org/10.1016/j.neuroimage.2011.11.089>
- Seki, Y., & Tomyta, K. (2019). Effects of metronomic sounds on a self-paced tapping task in budgerigars and humans. *Current Zoology*, 65(1), 121–128. doi:<https://doi.org/10.1093/cz/zoy075>
- Snyder, J. S., & Large, E. W. (2005). Gamma-band Activity Reflects the Metric Structure of Rhythmic Tone Sequences. *Brain Res Cogn Brain Res*, 24(1), 117-126. doi:<https://doi.org/10.1016/j.cogbrainres.2004.12.014>
- Stanford University. (2017). *Naturalistic Music EEG Dataset - Tempo (NMED-T)*. Retrieved from Research from Stanford University: <https://purl.stanford.edu/jn859kj8079>
- Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., & Golumbic, E. Z. (2017). Neural Entrainment to the Beat: The “Missing-Pulse” Phenomenon. *Journal of Neuroscience*, 37(26). doi:<https://doi.org/10.1523/JNEUROSCI.2500-16.2017>
- Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct Neural Substrates of Duration-Based and Beat-Based Auditory Timing. *Journal of Neuroscience*, 31(10), 3805-3812. doi:<https://doi.org/10.1523/JNEUROSCI.5561-10.2011>

- Toiviainen, P., & Snyder, J. S. (2003). Tapping to Bach: Resonance-Based Modeling of Pulse. *Music Perception*, 21(1), 43–80. doi:<https://doi.org/10.1525/mp.2003.21.1.43>
- Turner, V. (1995). *The Ritual Process: Structure and Anti-Structure*. Transaction Publishers.
- Velasco, M. J., & Large, E. W. (2011). Pulse Detection in Syncopated Rhythms Using Neural Oscillators. *ISMIR*.
- Wang, Y., Ding, N., Ahmar, N., Xiang, J., Poeppel, D., & Simon, J. Z. (2012). Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence. *Journal of Neurophysiology*, 107(8), 2033-2041.
- Wiener, M., Turkeltaub, P., & Coslett, H. (2010). The image of time: A voxel-wise meta-analysis. *NeuroImage*, 49(2), 1728-1740. doi:<https://doi.org/10.1016/j.neuroimage.2009.09.064>
- Will, U., & Berg, E. (2007). Brain Wave Synchronization and Entrainment to Periodic Acoustic Stimuli. *Neuroscience Letter*, 424(1), 55-60. doi:<https://doi.org/10.1016/j.neulet.2007.07.036>
- Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America*, 106(7), 2468-2471. doi:<https://dx.doi.org/10.1073/pnas.0809035106>
- Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond Timing in Primates: Comparison of Interval Production Between Human Subjects and Rhesus Monkeys. *Journal of Neuropsychology*, 102(6), 3191-3202.
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8(7), 547–558. doi:<https://doi.org/10.1038/nrn2152>
- Zhou, H., Melloni, L., Poeppel, D., & Ding, N. (2016, 06 06). Interpretations of Frequency Domain Analyses of Neural Entrainment: Periodicity, Fundamental Frequency, and Harmonics. *Frontiers in Human Neuroscience*, 10, 274. doi:[10.3389/fnhum.2016.00274](https://doi.org/10.3389/fnhum.2016.00274)

~ The End ~