Personality's Influence on Emotional Experience of Music

Daniel First, Aditya Garg, Lakshya Garg, Zac Robertson
{df2450,ag3741,lg2906,zlr2101}@columbia.edu

Abstract

In this paper, we take a Natural Language Processing approach to analyze reports of Strong Musical Experiences - that is, text reports of subjects' recollections describing a particularly intense emotional experience in response to music. In this paper, we focus on the impact of personality and gender on these experiences and reports. We hypothesize that people with different personality traits will have different emotional reactions to music and report them differently. After collecting a novel dataset of Strong Musical Experiences together with personality and demographic data, we used the Linguistic Inquiry and Word Count (LIWC) lexicon to analyze how, for example, extroverts and introverts differ in their reports of peak musical experiences. We found a number of significant differences, such as a greater presence of Social words among individuals who are Compassionate. We also found that the genders differed significantly in their reports of peak musical experiences. Women tended to use more words related to other people and family. We also found that men tend not to use words related to negative emotion, especially sadness.

1 Introduction

From time immemorial, humans have turned to music to induce powerful emotional experiences. In ancient Greece, the Athenians would perform dithyrambs at festivals, where flute players and choruses would join to induce an emotional experience in worship of the Greek god Dionysus. Later thinkers placed music at the center of the human search for happiness and meaning. Schopenhauer said that through music we are able to forget the misery of existence, and Nietzsche opined that music has the potential to inspire one to see the value of life. Today, people continue to turn to music for emotional experiences and meaning, at churches, funerals, dance clubs, and religious rituals around the world.

The central role that music plays in facilitating emotional experiences accentuates the importance of studying the psychology of music. Music psychology is thus an active field of research. Because people experience music as being formational to their identity, their emotional lives, and often even the meaning they derive from life (Lamont 2012), psychologists have begun to explore this area scientifically.

However there is only limited research into how people with different personalities might experience music differently. Do introverts experience jazz differently from the way extroverts do? A sub-inquiry within this question focuses on particularly intense emotional experiences in the context of music. Do people with different personality types tend to have different sorts of these intense experiences? In this study, we focus on this last sub-inquiry. For example, is it possible that people who are extroverts have different musical experiences than introverts? Are people who score high on neuroticism more likely to have intense sad experiences than those low on neuroticism? We are also interested in the impact of gender. We hypothesize, for example, that women will be more emotionally expressive than men.

Much prior work has been done on the field of music psychology in general. Neuropsychologists have studied the neuroscience of music to decipher how music affects our mental and physical health. Extensive research has been conducting into the nature of strong experiences of music listening, identifying a number of individual components from physiological through to psychological (Gabrielsson Lindström Wik, 2003) and their relationship with mainstream theories of happiness and sadness.

Of course, as onlookers from the outside, we do not have access to people's actual emotional experiences. We therefore rely on people's reports of their strong emotional experiences, as other studies in the field do (ibid.). In this study in particular, we will focus, then, on **"Strong Musical Experiences"** (SMEs)- particularly intense emotional events induced by music. Our aim is to investigate whether people with different personality types have different SMEs, and whether they describe these SMEs differently. Our approach will be to analyze the features in reports of SMEs:We will analyze people's **free text description reports** of SMEs using various lexical and semantic features. We expect our study will show that, for example, introverts mention sadness more than extroverts, when reporting their SMEs. Of course, given such a result, our analysis leaves open the question of whether the two groups are having different experiences, or if alternatively they are having the same experience yet reporting it differently.

A deeper understanding of how we experience music emotionally has broad applications. When extended, this study would have applications in music psychology and the music industry. For example, music therapists could use these results in conjunction with personality tests on their patients to identify types of music that are likely to bring about certain emotions. Sales teams in retail stores may want to play music that appeals to certain personality types or elicits emotions in them.

2 Hypothesis

The central guiding hypothesis of this research is that people with different personality traits are affected by music differently, and therefore relate their strong musical experiences with different physiological, emotional, psychological, and social linguistic features. **We have a novel dataset** of over seven hundred text reports of Strong Musical Experiences, as well as ratings for the personalities of the subjects who wrote the responses, as described below. By examining various syntactic features extracted from free text accounts of these reported peak musical experiences, we plan to examine if certain emotions, expressions and language features are frequently associated with different personality types with regards to music. For example, we hypothesize that extroverts will report more social words and that people high on neuroticism (that is, the opposite of emotional stability) will have more negative emotions reported, and that conscientiousness is correlated with religious experiences.

Aside from our central hypothesis about music and personality, we will also investigate whether our research confirms or contradicts other research into how people with different personality types use different styles of speech.

3 Related Work

Our research draws from a broad variety of existing literature, spanning the fields of personality and emotion research, music psychology, and semantic and lexical feature extraction.This section elaborates some of the related work and how our methodology differs from the same.

Music and Personality Studies: There have been some studies regarding the relationship between personality and emotion, but our study will differ in various significant ways from them. One study looked at the way in which people with different personalities experienced sad music. Additionally, Garrido and Schubert (2011)present a theoretical framework which investigates individual differences in enjoyment of "sad" music, and some of the personality factors that may be involved. While music was played, participants were asked to describe their experience; what thoughts or feeling arose while listening; how the emotions they experienced compared to what they perceive the music to be expressing; and, how the emotions experienced compared to emotions induced by "real-life situations." Followed by questions related to broader listening habits and demographics, participants chose a diverse range of pieces to play during the interview. Particular patterns of responses to sad music include the tendency that individuals with ruminative tendencies may have a lesser capacity to dissociate. Although this does try to marry personality and emotions experienced while listening to music, the experiment is restricted to only five participants and accounts only negative emotions related to music. Additionally, no standard mechanism is followed in testing personalities. Our study does not limit the scope in terms of emotional experiences and uses the popular OCEAN personality test to assign personality labels.

In their study titled "Strong Experiences Related to Music (SEM), a Descriptive System", Gabrielsson and Lindstrom source free descriptions of the strongest experience of music of 900 swedish volunteers, and present a comprehensive and detailed description of the components in strong experiences related to music. The reactions related in the participant's reports were classified into common broad psychological categories including General characteristics, Physical reactions and behaviors, Perception, Cognition, Feelings/Emotions, Existential and transcendental aspects, and Personal and social aspects. In addition to SEM reports, participants were asked to answer some more specific questions around their experience, aimed to help understand their reactions and draw a timeline of how a specific experience occurred, triggers involved and after effects. This study is closely related to our overall aim to analyze reports of peak musical experiences since both concepts (SEM and PEM) have many features in common. The above uses free text followed by questionnaire to deep dive into SEM of an individual, but our more targeted and focused approach gathers details of peak experience by mean of specific questions as described in methodology and relate the same with the personality of the respondent. Also analysis of SEM responses was performed manually by two individuals, which is replaced by automated NLP technique in our experiment.

Language Use Differs by Demographics: The LIWC lexicon has also been used extensively for studying gender and age. Many studies have focused on function words (articles, prepositions, conjunctions, and pronouns), finding females use more first-person singular pronouns, males use more articles, and that older individuals use more plural pronouns and future tense verbs(Schwartz 2013). Other works have found males use more formal, affirmation, and informational words, while females use more social interaction, and deictic language. Regarding age, the most salient findings include older individuals using more positive emotion and less negative emotion words, preferring fewer self-references (i.e. 'I', 'me'), and stylistically less use of negation. Additionally, Argamon et al. used factor analysis and identified 20 coherent components of word use to link gender and age, showing male components of language increase with age while female factors decrease.

Occasionally, studies find contradictory results. For example, multiple studies report that emoticons (i.e. ':)' ':-(') are used more often by females, but Huffaker Calvert (2005) found males use them more in a sample of 100 teenage bloggers. This particular discrepancy could be sample-related – differing demographics or having a non-representative sample (Huffaker Calvert looked at 100 bloggers, while later studies have looked at thousands of twitter users) or it could be due to differences in the domain of the text (blogs versus twitter).

Our research in extracting text is supported by the work of Jurafsky, Ranganath, and McFarland(2009), in which they examine speech from speed dates and classify interaction styles as awkward, friendly, or flirtatious based on survey assessments from the date participants. Their lexical features are drawn from the LIWC lexicon as they extract features including anger, assent, negative emotion, and the use of pronouns such as I, we, and you. They feed their features into a logistic regression-based binary classifier, finding 60-75% accuracy in classifying their results into the given interaction styles. We plan to draw from their work by also using LIWC features and classifying passages of text into categories, but we will expand on this research by applying these techniques to music psychology, emotion, and personality types. We will also build off of the work of Yarkoni, who explored LIWC categories and unigrams in connection with personality scoring from 406 bloggers. His team found that certain words correlated with certain personality traits - for example, 'hug' with agreeableness.

4 Methodology

For the purpose of our experiment, we collected a novel dataset using questionnaires. This questionnaire had two parts: One, for personality; the second, questions asking for reports about the Strong Musical Experience of the subject in question. In total, we had 920 responses. These were collected in part from online volunteers visiting http://www.musicaluniverse.org who opted to take the survey as well as M-turk users.

4.1 Dataset

Our dataset consists of a 2-part of questionnaires.

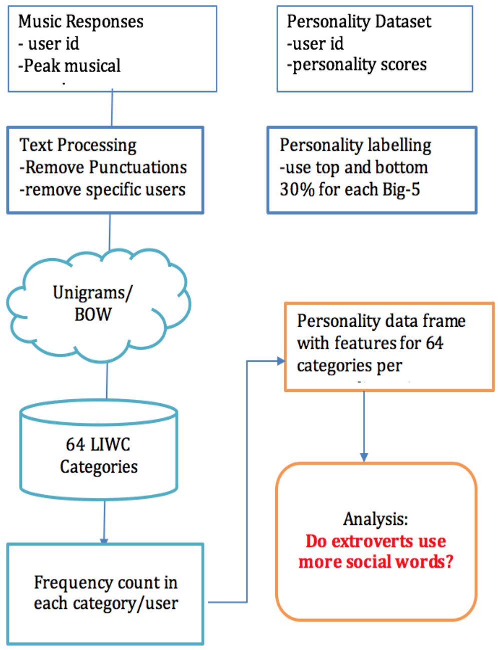
Personality Questionnaire: The first questionaire consists of the standard and widely-used

personality questions, TEN ITEM PERSONALITY MEASURE (TIPI) (Ten-Item Personality Inventory by Gosling, Rentfrow, Swann, 2003 http://gosling.psy.utexas.edu/scales-weve-developed/ten-item-personality-measure-tipi/) scale of the “Big Five” OCEAN personality traits to assign personality labels to respondents. The big five personality traits are the best accepted and most commonly used model of personality in academic psychology. The big five (or Five-Factor Model) come from the statistical study of responses to personality items. The test consists of ten items that one must rate on how true they are about oneself on a seven-point scale where 1=Disagree, 4=Neutral and 7=Agree. It takes most people 1-2 minutes to complete. Subjects are then given scores from 1-7 on five personality dimensions: Openness, Conscientiousness, Extroversion, Agreeableness, Neuroticism. It also collected various pieces of demographic information, but we only used gender from this.
Strong Musical Experience Questionnaire: The second questionnaire consists of subjective questions to capture the overall peak musical experience. Here, subjects are asked to provide free text that reports answers to questions: “Describe your peak musical experience?” “What is the meaning of music in your life?” and “What is your everyday musical experience?” There were also various follow-up questions: “How did you feel after the experience,” “What do you think was the cause of the experience?” and “How often do such strong experiences occur?”.

4.2 Test Processing

As described in Figure1,for the purpose of our experiment, we first verified each respondent answer both questionnaires. Next, we manually filtered out 47 users who gave random responses to meet the word limit-mainly consisting of responses with meaningless text, leaving 873 remaining. In the next stage, we remove certain punctuation marks from the responses that may introduce noise in the task of feature extraction.
After pre-processing of free text responses, unigrams were obtained. On average, respondents used 258 words in total across all free-response questions. Next we used The Linguistic Word Count (LIWC), a psycholinguistics lexicon that has been frequently used to incorporate semantic and psychological information into linguistic

Figure 1: Overview of free text responses and assignment of personality labels



analysis. We obtain features for each of the 64 psycholinguistic classes present in the lexicon by calculating the percentage of words in the pre-processed responses belonging to each class. We were therefore left with a 873-by-71 matrix, in which each row was a subject, and for each subject we had 65 measures of feature percentages, as well as five personality scores and gender.
Responses obtained from personality dataset are processed such that only top and bottom 30% of each category scores are labelled as extremes in that category and the middle section was ignored for that series of tests, in order to accentuate the differences between the personality archetypes. For example, in extraversion category people with the top 30% of scores are labelled as extroverts, the bottom 30% scores are labelled as introverts. We also tested another approach where we used a threshold to categorize personality on a scale of 0 or 1 based on their personality score for a particular trait is greater or less than 4. However, results were far more significant using the former approach. Thus we had features for each of the 64 categories, for each of the 5 personality traits.

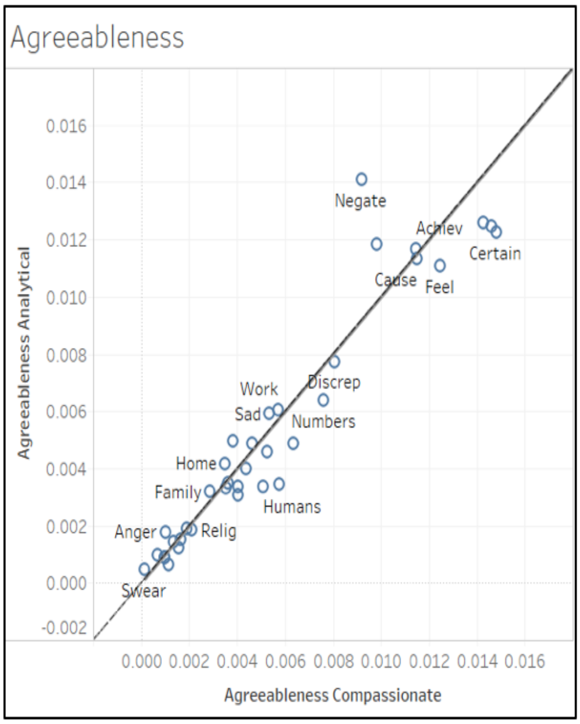
5 Results

We present results based on personality, as well as for gender.

Figure 2: p-values for 21 LIWC categories and personality labels

Personality Category	LIWC Category	P-value
Agreeableness	Funct	< 0.001
Agreeableness	Humans	< 0.001
Agreeableness	Incl	< 0.001
Agreeableness	Social	< 0.001
Agreeableness	Past	0.001
Agreeableness	They	0.002
Agreeableness	We	0.003
Agreeableness	Certain	0.007
Agreeableness	Sexual	0.016
Agreeableness	Bio	0.02
Openness	Verbs	0.006
Openness	Certain	0.01
Openness	Space	0.011
Extraversion	Hear	0.011
Extraversion	Incl	0.012
Extraversion	Posemo	0.018
Conscientiousness	AuxVb	0.002
Conscientiousness	Leisure	0.002
Conscientiousness	Past	0.006
Conscientiousness	Filler	0.008
Emotional Stability	Insight	< 0.001

Figure 3: Comparison of Agreeable analytical and compassionate personalities



Personality: To analyze our results, we performed a series of Mann-Whitney U Tests comparing each opposing personality type across the LIWC categories. This led to 43 significant results among 325 tests using a threshold of p = 0.05. The 21 results with a p-value less than or equal to 0.02 are listed in Figure 2.

These results can be visualized using a scatter plot with opposing personality types on the x and y-axes, using a 45 degree line for reference as shown in Figure 3.

In this example comparing individuals who score as Compassionate against those who score as Analytical on the Agreeableness personality trait, we can see that LIWC categories such as Humans, Feel, and Certain tend to fall below the reference line, meaning they are more associated with Compassionate people.This can be observed anecdotally from some of our highest-scoring Human responses, such as the following:

“The most profound music experience I had was at a blues festival in the middle of Wyoming. I saw Buddy Guy and Taj Mahal for the first time ever, and my life was forever changed. I felt things I’d never felt before.. I felt alive. I felt like the music spoke to me, even though I’m just an average white girl from Wyoming. That was 10 years ago and I still love the blues. I love it so much, I named my daughter Mae- a common middle name for lady blues singers and the lady subjects of many blues songs. I hope that my daughter can someday appreciate the blues as much as I do.”

One of our initial hypotheses was that Extroverts would likely use more social words than Introverts, but our analysis has not shown a significant relationship to exist (p-value = 0.537). However, individuals who rate as Compassionate on the Agreeableness personality factor do use more social, human, and inclusion words, each with a significance level of less than 0.001. That said, Extroverts do exhibit some behavior that bears an intuitive resemblance to social words, such as using significantly more Hearing, Inclusion, and Positive Emotion words than their Introverted counterparts.

Gender: We found that in a number of feature areas, women scored higher than men. We found that women used more words related to family (p=0.008), humans (.02949), insight (.003), negative emotion (.004), sadness (p=.0003), sexuality (.049), social (.0002), and verbs (.00028). Surprisingly, men did not score higher than women on any LIWC feature, beyond a threshold of statistical significance. This supports our hypothesis that women are more emotionally expressive than men, yet interestingly the differences were limited to negative emotion. This may lend support to the common observation that men are socialized to not talk about their negative emotions as much, and especially not sadness. This also confirms the findings of Argamon (2003) and Rao (2010) that women use more social words than men, in general. However, our study did not replicate the finding that men use more formal and affirmation words, nor that women use more first-personal singular pronouns. This suggests that in the future, explorations into how the languages of men and women differ can sometimes hold cross-domain, yet sometimes be limited to domain.

6 Challenges and Future Work

The dataset used for the study is very unique, given that there exists no source or open data that provides description of a person's peak musical experience, along with their personality labels and demographic data. This posed a challenge since the collection of data required incentivizing users to fill lengthy questionnaires describing their peak musical experience in detail, besides answering Ten Item Personality Measure (TIPI) and demographics-related questions. Although we were able to collect about 900 responses through extensive mailing and offering monetary incentives to people on Amazon Mechanical Turk, we would have liked to collect additional data to further improve the statistical importance of our results.

Also, given the sheer volume of feature significance tests that we ran on the dataset, it is likely that our statistical measures are prone to Type-I error of hypothesis testing. However, we obtained a large number of results that can be considered highly significant even given the volume of our tests.

Our study focused on comparing the different extremities of specific personality traits with respect to their description of peak musical experiences, for e.g. – extrovert vs introvert, curious vs. cautious etc. This study, however, does not account for composite personality traits and their implication on musical experiences of the user. For instance, it could be possible that compassionate extroverts differ in their musical experiences from detached extroverts. We could look into analyzing combinations of different personality traits in the future to gain additional insights into individuals. Additionally, we could explore how gender affects the interaction between traits. Perhaps women who score low on agreeableness differ from their high-scoring female counterparts, yet such a difference does not emerge among men.

In this study we showed how personality traits significantly affect the way individuals experience and describe music. Out of the 65 different psycholinguistic classes derived from the LIWC lexicon, we found that some of them are characteristic to different personality types. Therefore, in the future, we can use these classes as features, and combine them with other lexical and semantic features, to develop a classifier that is able to accurately predict the personality type of an individual. This approach would allow us to automatically determine the personality traits of an individual from free text responses. This study would have widespread applications in music industry and musical psychology where automated personality detection can be used to assess target customers and analyze subjects' personality.

Currently, we focused on studying the interaction of individuals with music as a whole. In the future, we can extend this study to analyze the interaction of different personalities with different genres of music. For instance, we could check whether certain personality types, say extroverts, have their peak musical experience when listening to particular genres of music, say rock. This would have useful applications in musical therapy where subjects can be recommended music based on their personality traits, and also in the musical industry.

Along with peak musical experience responses and personality data, we have also collected the demographics data such as country, age, ethnicity, education level etc. It would be interesting to analyze whether different demographic groups interact with music differently and differ in their description of their peak musical experiences. For instance, we can determine whether different age groups relate their peak musical experiences distinctly or whether people from certain countries/ethnicities have a characteristic interaction with music. This also raises the question of whether subjects are differing in their emotional experience of music, or if alternatively they are differing in their descriptions of similar emotional experiences. Future work may aim to tease out the differences here. For example, given our finding that men do not mention sadness as much as women (when unprompted to discuss sadness), perhaps this difference would not appear if we directly asked men whether music made them feel sad (prompted). If so, then this would suggest that the differences in our study are due to differences in how the genders speak, not how they experience emotion. The unique dataset we have collected with its rich feature set allows a variety of analyses and interesting opportunities to study the musical experience and its effects on different kinds of individuals.

As we have shown, an individual's personality type has a great deal of influence on how they talk about their most significant experiences with music. Via our use of a novel dataset as well as established methodologies in psychology and natural language processing, we have demonstrated a large number of relationships between personality and peak musical experience, several of which are significant even at a level of 0.001.

Additionally, we have illustrated the role of gender in peak musical experience, validating our hypotheses that women tend to use more social and emotional language.

7 Conclusions

As we have shown, an individual's personality type has a great deal of influence on how they report their most significant experiences with music. Via our use of a novel dataset as well as established methodologies in psychology and natural language processing, we have demonstrated a large number of relationships between personality and peak musical experience, several of which are significant even at a level of 0.001. Our research highlights the fact that extroverts are more likely to have positive emotion-valenced peak musical experiences, even if our findings did not confirm that extroverts are more likely to describe their peak musical experiences in social terms.

Our results also shed light on the ongoing discussion of how men and women talk about their emotional experiences differently. We hope this paper serves as a case study of how Natural Language Processing techniques can be used to weigh in on long-standing debates in gender studies. Instead of relying on stereotypes about how genders speak, we can now do actual empirical studies. And we need no longer rely on specific well-known examples - through our method, we are able to source text from non-experts. This Natural Language Processing approach can then be applied cross-culturally, to explore how gender performance might differ in different cultures. We hope these techniques will be embrace by interdisciplinary teams in years going forward.

References

David M. Blei,Andrew Y. N,Michael I. Jordan 2003. *Latent Dirichlet Allocation.* Journal of Machine Learning Research 3 (2003) 993-1022

François Mairesse, Marilyn Walker, Matthias Mehl and Roger Moore 2007. *Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text.* Journal of Artificial Intelligence Research (JAIR)

Gabrielsson, A., Wik, S. L. 2003. *Strong Experiences Related to Music: A descriptive System.* Musicae scientiae, 7(2), 157-217.

Huffaker, D. A., Calvert, S. L. 2005. *Gender, identity, and language use in teenage blogs.* Journal of ComputerMediated Communication, 10(2), 00-00.

Jurafsky, Ranganath, McFarland. 2009. *Extracting Social Meaning: Identifying Interactional Style in Spoken Conversation* HLT-NAACL: 638-646.

Lamont, A. 2012. *Emotion, engagement and meaning in strong experiences of music performance* Psychology of Music, 40(5), 574-594.

Mona Lisa Chanda and Daniel J. 2013. *The neurochemistry of music* Trends in Cognitive Sciences, April 2013, Vol. 17, No. 4.

Park, Gregory; Schwartz, H. Andrew; Eichstaedt, Johannes C.; Kern, Margaret L.; Kosinski, Michal; Stillwell, David J.; Ungar, Lyle H.; Seligman, Martin E. P. 2015. *Automatic personality assessment through social media language* Journal of Personality and Social Psychology, Vol 108(6), Jun 2015, 934-952.

Pennebaker, J. W., Booth, R. J., Francis, M. E. 2007. *Linguistic inquiry and word count: LIWC [Computer software]* Austin, TX: liwc. net.

Sandra Garrido,Emery Schubert. 2011. *Negative Emotion in Music:What is the Attraction? A Qualitative Study.* Empirical Musicology Review,2011

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... Ungar, L. H. 2013. *Personality, gender, and age in the language of social media: The open-vocabulary approach.* PloS one, 8(9), e73791.

Survey website *www.getpsychd.org "Personality and Music"*

8 Team Contributions

8.1 Daniel First:

Read and summarized 'Schwartz, H.A (2013). Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach' and 'Lamont, A. (2011). University Students Strong Experiences of Music: Pleasure, Engagement, and Meaning.' Researched the differences between open- and closed-vocabulary approaches. Wrote parts of goal and hypothesis section as well as the background section. Contributed to the writing of the overall combined draft and conclusion. Performed analyses and wrote results section for the differences between men and women.

8.2 Aditya Garg:

Read and summarized 'Lamont, A. (2012). Emotion, engagement and meaning in strong experiences of music performance' and 'Automatic personality assessment through social media language (Park, Schwartz et al, 2014). Wrote parts of Goal and Hypothesis sections. Processed raw personality data to encode into binary personality traits. Wrote code to combine personality data with lexical data and convert it into CSVs for visualizations. Performed exploratory analysis in Tableau. Wrote challenges and future work sections in the paper.

8.3 Lakshya Garg:

Read and summarized "Gabrielsson, A., Wik, S. L. (2003). Strong Experiences Related to Music: A descriptive System","Sandra Gaarrido,Emery Schubert. Negative Emotion in Music:What is the Attraction? A Qualitative Study" .Wrote parts of goal,hypothesis. Wrote the python code for text pre-processing,creating unigrams,parsing LIWC categories and deriving features from the free text musical experience responses to generate dataframes used for analysis.Wrote the methodology(dataset,text processing,block diagram) section of the paper entirely and contribute to related work section for the reading done. Formatted the paper according to the template provided.

8.4 Zac Robertson:

Read and summarized "Extracting Social Meaning: Identifying Interactional Style in Spoken Conversation" and "Linguistic inquiry and word count: LIWC." Wrote parts of Goal, Hypothesis, Methodology, and Results sections, and combined other team members' works. Performed exploratory analysis of results in Tableau as well as statistical analysis of language tendencies by personality type.