KU LEUVEN

FACULTEIT PSYCHOLOGIE EN
PEDAGOGISCHE WETENSCHAPPEN

What Does Your Spotify Wrapped Say About Your Personality? A Lens Model Investigation Into Personality Impressions Based on Music Preferences

Master's thesis submitted for the degree of Master of Science in de psychologie by

Sara Samadi

Supervisor: Prof. Dr. Peter Kuppens

2025

KU LEUVEN

FACULTEIT PSYCHOLOGIE EN
PEDAGOGISCHE WETENSCHAPPEN

What Does Your Spotify Wrapped Say About Your Personality? A Lens Model Investigation Into Personality Impressions Based on Music Preferences

Master's thesis submitted for the degree of Master of Science in de psychologie by

Sara Samadi

Supervisor: Prof. Dr. Peter Kuppens

2025

Summary

Many people listen to music on Spotify, where they can freely choose what they listen to. Based on these choices, which reflect their listening behaviour, Spotify creates an annual compilation known as Spotify Wrapped. This raises the question of whether people are able to form accurate impressions of someone's personality based on their Spotify Wrapped. While many studies have examined the accuracy of personality impressions in other areas, only one study to date has addressed this in the context of music. This is the first study to examine the link between music preferences, personality and personality judgements using the popular Spotify Wrapped feature, which is primarily guided by the Lens Model of Brunswik (1955).

The aim of this thesis is investigating: (1) the relationship between a person's music preferences, specifically Spotify Wrapped and self-reported preferences for genres, and their personality, (2) the relationship between music preferences and personality judgements by others, and 3) the accuracy of personality judgements based on this information. To answer the first two questions, 14 genres and two other music features (quantity of preferred genres and the tendency to listen to the same artist) served as a total of 16 cues to make correlations. The study was conducted in two parts. For the first part, a questionnaire was created in which participants (N = 96) were asked to report their age group, gender, to upload a screenshot of their top 5 Spotify Wrapped songs, their preferences across 14 genres, and to complete a personality questionnaire. The first part of the study served as material for the second part of the study. 93 music profiles were created, which included the participant's gender, age group, top 5 Spotify Wrapped songs, and the genres they enjoy listening to. The 93 music profiles (referred to as targets) were divided across questionnaire A and questionnaire B, each questionnaire containing half of the music profiles. Six observers rated the personality of targets in questionnaire A, while another six did so for questionnaire B.

Regarding the first research question, there is weak evidence for the association between personality and music preferences. After applying the Bonferroni correction, most of the associations become negligible. Regarding the second research question, several associations are found, and this remains partly the case even after applying the Benjamini-Hochberg correction. People indeed form personality impressions based on multiple musical cues. This is especially true for the traits openness, conscientiousness, and agreeableness. Regarding the last research question, the results show that observers are modestly accurate in rating conscientiousness, neuroticism, and openness for experiences based on people's music preferences. People are indeed capable to form impressions that somewhat correspond to the actual personality based on looking at targets' music preferences as reflected in their Spotify Wrapped and preferred genres. This is a noteworthy finding given the limited information provided and the weak association between personality and music preferences in this study, as well as in others.

Acknowledgements

I would like to start by thanking Professor Dr. Peter Kuppens, my supervisor, for his support and guidance throughout my thesis. He was always there to give me feedback, to redirect me if necessary and to give me advice. Furthermore, he did not just point out what needed improvement, he also recognized and encouraged me when I got things right.

I also want to thank my godfather. He has been an important part of my life, especially during my university years. He and his wife provided me with a warm and stimulating environment where I could spend countless hours at your desk, studying and writing. I thank him for his unconditional belief in me and his unwavering support.

I also thank my kind mother, father and sisters for their loving support. It is especially to my mother that I owe my love for music, cultures, art, and beauty.

I want to thank my friends, with special thanks to Jirka, who has provided me with thoughtful and critical feedback.

Furthermore I want to thank my participants, especially those who took part in the second part of the study, which was a time consuming process to complete.

Last but not least, I want to thank my dear partner for his unwavering optimism and support throughout this entire journey. He truly has one of the most beautiful personalities I know.

Approach and Own Contribution

During the thesis topic selection phase, I was assigned the subject "cars and personality: inferring personality traits from car-related behaviour." However, during my first meeting with my supervisor, I was given the opportunity to adjust the topic and pursue something related to music. He handed me scientific articles on the Lens Model and some articles on the association between personality and music preferences. From there, I was tasked with applying the Lens Model to the domain of music. Due to the scarcity of studies applying the Lens Model to music, I engaged with a broad range of literature. In parallel with engaging extensively with the existing literature, I began observing the role of music in everyday life. As an avid Spotify user, I found inspiration in the platform's Spotify Wrapped feature, a yearly compilation of one's listening behaviour, which is widely shared on social media. I noticed that this compilation often revealed something new or unexpected about a person, which sparked the idea of linking Spotify listening behaviour to perceptions of personality. This observation eventually became the central inspiration for my thesis. I presented this idea to my supervisor, along with the method I was considering using. We brainstormed how the compilation should look like, and on my supervisor's advice, I decided to include genre preferences alongside participants' top tracks, so that cues would be included. After that, I made the final decisions on how the methodology would look and began drafting the ethics application for the study. My supervisor provided me with an example of it. He also provided me with several personality questionnaires, which I then selected and implemented in the study.

After approval by the Social and Societal Ethics Committee of KU Leuven, I was able to start creating the first questionnaire. Once created, I distributed it via my social media channels. After collecting the first dataset, I designed the second questionnaire based on the data from the initial study. This second survey was then distributed among my family and friends on social media. Following data collection, I met with my supervisor to discuss the analysis phase. I addressed the issue of two different observer groups, and he advised centering a part of my variables. He also recommended correcting for the risk of Type I errors. For each analytic method, I chose myself which one to use. Since I had read across a broad range of literature, it was difficult to define clear boundaries and establish a coherent storyline in my literature review. However, thanks to the constructive feedback and redirection from my supervisor toward the research questions, I was ultimately able to succeed in this task.

Finally, I have used ChatGPT as a language assistance, short-form input assistance and search engine. My GenAI code of conduct can be found in Appendix A.

Table of Contents

List of Tables

List Of Figures

"Music is our oldest form of expression, older than language or art; it begins with the voice, and with our overwhelming need to reach out to others." (Menuhin & Davis, 1979, p. 1)

People are interested in understanding others' personalities, often relying on various "cues" to form impressions (Hinds & Joinson, 2024). One such cue is how people present themselves online, such as through social media. One increasingly popular way to self-express online is by sharing one's Spotify Wrapped. This is an annual compilation of a user's music listening behaviour that provides information on the top five artists, songs, and genres (Wikipedia contributors, n.d.) of a user. Sharing one's Spotify Wrapped potentially allows others to form personality judgements based on musical taste. Despite the widespread visibility of Spotify Wrapped, relatively little research has been conducted on how it may reveal someone's personality. Historically, music preferences were less immediately observable compared to cues like behaviour or personal environments. With Spotify Wrapped increasing the visibility of music preferences, it prompts questions about what they reveal about personality, the personality impressions that are formed, and the accuracy of these judgements.

To address these gaps, this thesis investigates: (1) the relationship between a person's music preferences, specifically their Spotify Wrapped and genre preferences, and their personality (2) the relationship between music preferences (Spotify Wrapped and preferred genres) and personality judgements by others, and (3) the then combined accuracy of personality judgements based on this information. The Lens Model (Brunswik, 1955) and the Realistic Accuracy Model (Funder, 1995) will guide this investigation.

Personality and its Associations

The Five-Factor Model of Personality

The Five-Factor Model of Personality is the dominant framework for describing personality traits (Costa & McCrae, 1999; John & Srivastava, 1999) due to its strong empirical evidence (John & Srivastava, 1999). This dimensional framework describes the five broad, latent personality factors: extraversion, agreeableness, conscientiousness, neuroticism and openness. The dimensions are defined by John and Srivastava (1999) as follows: extraversion represents an energetic orientation toward the world, characterized by traits like sociability, activity, assertiveness, and positive emotions. Agreeableness represents a prosocial and cooperative disposition, encompassing qualities such as altruism, kindness, trust, and modesty. Conscientiousness reflects socially guided impulse control, behaviours that are task-oriented

and goal-focused, including careful decision-making, delaying gratification, adherence to rules, and effective planning and organization. Neuroticism captures a contrast with emotional stability and reflects tendencies toward negative emotional states such as anxiety, sadness, nervousness, and tension. Finally, openness to experience describes the richness, originality, and depth of an individual's mental and experiential world, as opposed to a more closed-minded or conventional outlook.

Associations between Personality and Life Domains

Personality is important because it plays a role in many life domains. A large body of research has demonstrated that the Big Five traits play crucial roles in physical and mental health (Hyatt et al., 2024; Luo et al., 2023; Malouff et al., 2005; Stephan et al., 2022), longevity (Chapman et al., 2011) and happiness (Malinka et al., 2025). Meta-analyses have also shown that the Big Five dimensions are associated with academic and job performance (Barrick & Mount, 1991; Mammadov, 2022; Poropat, 2009; van Aarde et al., 2017), relationships (Fransson et al., 2013; Malouff et al., 2010), intelligence (Anglim et al., 2022) and so on. Since personality influences numerous aspects of an individual's life, getting an idea of someone's personality is of significant importance.

Personality Judgements

For many people it is important to get an idea of someone's personality. Inferring personality traits of others is part of the broader domain of social cognition, which involves psychological processes that allows us to make inferences about people's intentions, feelings and thoughts (Adolphs, 2009).

The Lens Model

The Lens Model of Brunswik (1955) and Realistic Accuracy Model of Funder (1995) serve as the theoretical framework in this thesis to examine personality judgements based on behavioural cues. Personality judgement studies involve a target, this is the person who is being judged, and an observer, this is the person making a personality judgement about the target (Hinds & Joinson, 2024).

According to the Lens Model of Brunswik (1955), observable information or cues act as a lens through which observers indirectly perceive and judge an underlying construct, which, in this thesis, is a personality trait. Accurate personality judgements then arise when valid cues are used (Brunswik, 1955; Hinds & Joinson, 2024). According to Brunswik (1955), personality

judgement happens by a two-step inference process: cue validity and cue utilization (Gosling et al., 2002) which will be explained with Figure 1, which describes the Lens Model of Brunswik (1955).

Figure 1

The Lens Model of Brunswik (1955) (Gosling et al., 2002, p. 380)

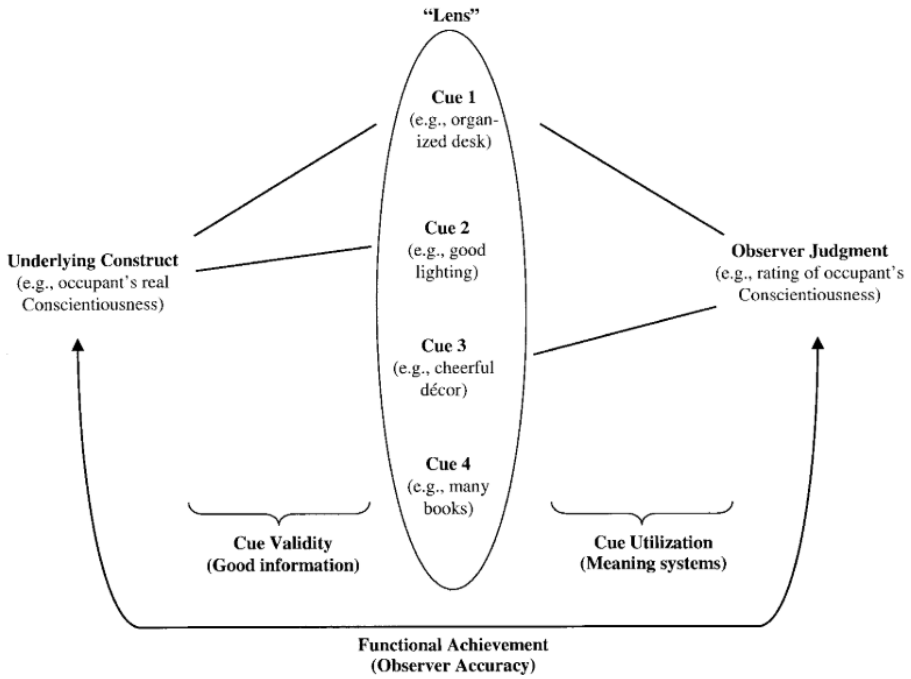


Figure 1 illustrates cue validity on the left side, representing the relationship between observable cues (e.g. the distinctiveness of a room) and underlying personality traits (e.g. self-rated openness) (Gosling et al., 2002). The right side of the figure illustrates *cue utilization,* representing the relationship between observable cues (e.g. the distinctiveness of a room) and the perceived underlying personality trait (e.g. observer-rated openness) (Gosling et al., 2002). The bottom line of the figure describes *observer accuracy*, this is the relation between self-rated personality traits (e.g., self-reported openness) and observer-rated personality traits (e.g., observer-rated openness) based on the cues (Gosling et al., 2002). High cue validity and cue utilization lead to high observer accuracy (Gosling et al., 2002).

Taking Gosling et al. (2002) research as an example, Brunswik's (1955) Lens Model suggests that observable information, such as rooms and offices, serves as a lens through which observers indirectly perceive and judge personality traits. Accurate personality judgement then

arises when observers rely on valid cues such as distinctive features, cluttered or clean room, many books,…

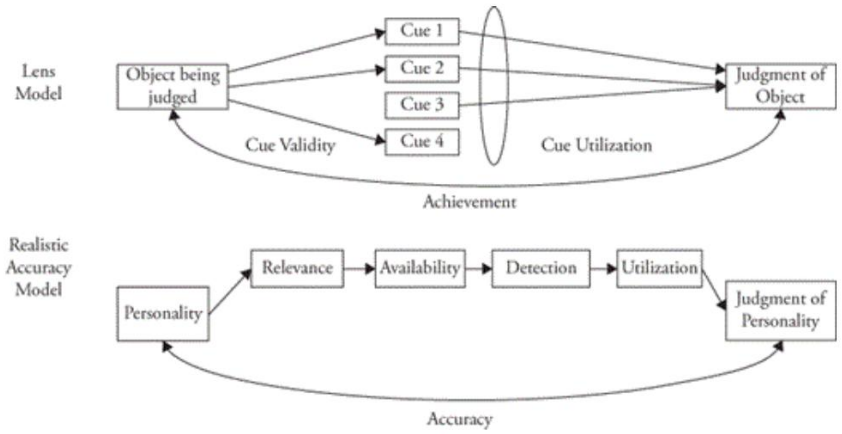
The Lens Model has been applied to different research areas, not only in psychology but also in business, medicine, and education (Karelaia & Hogarth, 2008; Kaufmann et al., 2013). Within psychology, the model has been applied to constructs such as pain (Ruben & Hall, 2016) lying (Hartwig & Bond Jr., 2011) and well-being (Choi et al., 2024). In personality research, cue validity has been adapted to describe personality expression, and cue utilization to describe impression formation, terms that more accurately reflect its application in personality context (Osterholz et al., 2023). The Lens Model has inspired further theory development (Letzring & Funder, 2021) which led to the Realistic Accuracy Model (RAM) of Funder (1995).

The Realistic Accuracy Model

The Lens Model (Brunswik, 1955) is broken down by the Realistic Accuracy Model of Funder (1995) (Letzring & Funder, 2021) (Figure 2) and provides a more extensive understanding of the process of accurate personality judgement. Accurate personality judgement happens in a four-stage process: relevance, availability, detection and utilization (Letzring & Funder, 2021).

Figure 2

The Lens Model of Brunswik (1955) and The Realistic Accuracy Model of Funder (1995) (Letzring & Funder, 2021, p. 12)



Within the RAM, every stage is necessary to observer accuracy (Letzring & Funder, 2021). The first stage *relevance* states that there must be behavioural cues that are relevant to the trait (Letzing & Funder, 2019). *Availability* emphasizes that that these behavioural cues have

to available in the external environment (Letzring & Funder, 2021). The two stages correspond to cue validity of the Lens Model (Letzring & Funder, 2021). Next, the relevant and available cues should be *detected* by the observer, which involves paying attention to the cue (Letzring & Funder, 2021). The fulfilling of this stage is primarily influenced by the judge itself (Letzring & Funder, 2021). Finally it should also be *utilized* by the judge to make accurate personality judgement, which includes determining which cues are relevant, giving appropriate weights to different cues, combining the cues and accounting for other characteristics such as the situation (Letzring & Funder, 2021).

Forming accurate personality judgements also involves several factors or moderators: good judge, good target, good trait and good information (Letzring & Funder, 2021). *A good judge* reflects the idea that some individuals are better at judging others accurately due to their own qualities such as intelligence, knowledge on the relationship between personality and behaviour, agreeableness and social skills, empathy, perspective taking and psychological adjustment (Letzring & Funder, 2021). *A good target* reflects the idea that some individuals are easier to judge accurately because of the targets having a high psychological adjustment, behavioural consistency, social skills, extraversions and expressiveness (Letzring & Funder, 2021). *The good trait* suggests there is something about traits that can be accurately judged across judges, targets and situation (Letzring & Funder, 2021). In general, good traits are easily observable (Letzring & Funder, 2021). Finally, *good information* emphasizes that observer accuracy depends on information available, containing two aspects, information quantity and information quality (Letzring & Funder, 2021). The RAM proposes that each moderator can interact with every other moderator (Letzring & Funder, 2021). In this thesis, we are interested in the interaction between good information and good trait, which is called diagnosticity (Letzring & Funder, 2021). This refers to the idea that some information is more useful for judging certain traits (Letzring & Funder, 2021). Letzring and Funder (2021) provide the example that neuroticism is compared to other traits the least accurately judged trait, however observers are able to make accurate judgements of neuroticism after they observe a target in a stressful situation.

Research on Personality Judgement

Traditionally the focus has been on observer accuracy in personality judgement research (Hinds & Joinson, 2024) with less research testing the Lens Model as a whole. Therefore we will discuss personality judgement research with an emphasis on observer accuracy.

Observer Accuracy at Zero-Acquaintance

In this thesis, observers are strangers (or "zero-acquainted"), a context that presents a unique challenge for personality judgement. Frequent interaction and interpersonal intimacy enhance accuracy by providing access to targets' feelings, attitudes, and behaviours (Connelly & Ones, 2010; Hinds & Joinson, 2024). Furthermore, a comparison in presented information with varying quality (richness) and quantity revealed that an increasing quantity and quality of information improves accuracy (Letzring et al., 2006). However, strangers must rely on limited information (Hinds & Joinson, 2024). Nevertheless, research suggests they can still make accurate judgements (Connelly & Ones, 2010), this raises the question: which cues do they use to form their judgements?

Cues from Behaviour and Environments

Observers can use cues from behaviour such as physical appearance, e.g. a stylish look (Naumann et al., 2009; Vazire et al., 2008) or language, e.g. swearing (Koutsoumpis et al., 2022). Observers can also use cues from physical environments, e.g. a distinctive room (Gosling et al., 2002) or online environments, such as smiling on TikTok (Dong & Xie, 2023).

Gosling et al. (2002) is an example of a study that investigates all the relationships of the Lens Model of Brunswik (1955). In their research, observers made personality judgements based on physical environments, including the organization and features of rooms and offices. The researchers hypothesized that individuals select and shape their physical environments in ways that reflect and reinforce their personalities. They proposed that observers can detect meaningful cues from these environments and infer aspects of an individual's personality. The study identified a wide range of valid cues, particularly for the traits of conscientiousness and openness. A cue was considered valid if it was both utilized by observers (i.e. impression formation) and related to the target individual's personality (i.e. personality expression). Examples of such cues include tidiness, cleanliness, distinctiveness, and the variety of books present.

The digitalization of society has made it possible to perceive others not only in physical environments, but also in online environments (Hinds & Joinson, 2024). In the same way the Lens Model has been applied to digital cues in online social networks. Digital cues include Facebook cues such as profile picture friendliness, outgoing photos, information pages, and vocabulary (Darbyshire et al., 2016; Hall et al., 2014). Digital cues also include TikTok cues such as acting in videos, smiling, and background music (Dong & Xie, 2023). Another example

of digital cues are Instagram cues, such as post diversity and privacy settings (Osterholz et al., 2023).

Observer accuracy varies across different digital platforms, though general trends emerge (Hinds & Joinson, 2024). A meta-analysis by Tskhay and Rule (2014) examined 20 studies on online social networks and found that accuracy rankings followed this order as measured by the effect size: extraversion ($\bar{r}$ = .42), agreeableness ($\bar{r}$ = .21), openness ($\bar{r}$ = .19), conscientiousness ($\bar{r}$ = .18), and neuroticism ($\bar{r}$ = .11). A more recent meta-analysis by Hinds and Joinson (2024), which included 30 samples from online social networks and smartphone data, revealed modest observer accuracy across all five traits, with the highest accuracy for openness ($\bar{r}$ = .42), followed by extraversion ($\bar{r}$ = .40), agreeableness ($\bar{r}$ = .34), conscientiousness ($\bar{r}$ = .29), and finally neuroticism ($\bar{r}$ = .29). The meta-analyses of Tskhay and Rule (2014) and Hinds and Joinson (2024) show comparable findings, with openness and extraversion having the highest observer accuracy scores, followed by agreeableness, conscientiousness and neuroticism.

In summary, online social networks generally allow for moderate accuracy of personality judgements at zero-acquaintance, where accuracy varies depending on the platform, each providing different cues (Hall et al., 2014; Hinds & Joinson, 2024; Osterholz et al., 2023; Tskhay & Rule, 2014). Music preferences are another source of information that people can rely on, as they can reveal significant insights into a person's personality.

Music Preferences and Associations with Self-Reported Personality and Observer-Reported Personality

The Association Between Music Preferences and Self-Reported Personality

Music preferences are another possible source of information that could be used to infer someone's personality. But are there sufficient indications that music preferences can also serve as a relevant lens (Funder, 1995)? Our discussion of the relationship between music preferences and personality will be based on five studies. The first is the meta-analysis by Schäfer and Mehlhorn (2017), which covers research from 2003 to 2015. Then four additional studies are reviewed published from 2015, investigating personality traits at the level of the Big Five dimensions.

The relationship between music preferences and personality has been studied using two methods. The first relies on self-reported music preferences, while the second analyses actual listening behaviour (using for instance machine learning techniques).

Self-*reported music preferences*

Schäfer and Mehlhorn (2017) did a meta-analysis of 28 studies conducted between 2003 and 2015 regarding how personality traits are correlated to self-reported music preferences. The researchers aimed to bring clarity about the heterogeneity of the previous results, so reliable conclusions could be drawn. Following Cohen's conventions (Cohen, 1988), they found five positive but small correlations in relation to the MUSIC model, a model that categorizes music (Rentfrow et al., 2011). The categorization of music preferences can be done in different ways, but one categorization is the MUSIC model (Nave et al., 2018; Rentfrow et al., 2011). It encompasses five factors: Mellow (smooth, relaxing), Unpretentious (country, singer-songwriter), Sophisticated (complex, intellectual), Intense (loud, energetic), and Contemporary (rhythmic, percussive) (Rentfrow et al., 2011). This model transcends genre, as some styles (e.g., jazz) span multiple factors (Rentfrow et al., 2011). The five small positive correlations between the Big Five traits and the MUSIC model are as follows, as measured by the effect size: openness correlates with sophisticated ($\bar{r}$ = .21), mellow ($\bar{r}$ = .16), and intense music ($\bar{r}$ = .12); extraversion with contemporary music ($\bar{r}$ = .12); and agreeableness with unpretentious music ($\bar{r}$ = .10). Schäfer and Mehlhorn (2017) consider the results regarding the trait openness trivial because these people seem to be open to all kinds of music, which is known as an omnivorous taste (Elvers et al., 2015). They conclude that the direct influence of personality on music preferences has been overemphasized in the literature.

Greb et al. (2018) (N = 587, M = 25.4 years) also found openness to experience to be positively correlated with the selection of complex music and agreeableness and neuroticism negatively associated with the selection of complex music. Greb et al. (2018) conclude that their research is in line with the meta-analysis of Schäfer and Mehlhorn (2017).

Nave et al. (2018) (N = 22 252) found four small correlations with dimensions of the MUSIC model: openness was positively correlated to sophisticated music but negatively to both mellow and contemporary music, while extraversion showed a positive correlation with unpretentious music. Furthermore, the MUSIC model in its whole is associated with all of the Big Five traits. Finally, openness and extraversion are associated with general liking of music. The researchers conclude that there is strong evidence supporting the association between music preferences and personality.

All of the research discussed above has examined music preferences through self-reports of participants. But what about research that relies on music preferences based on actual behaviour of people? This will be discussed next.

Music Preferences Based on Actual Behaviour

Anderson et al. (2021) (N = 5808, M = 26 years) studied the association between the Big Five traits and music preferences based on actual behaviour. They collected Spotify data from users, including their Big Five personality traits. As a widely used music streaming service, Spotify provided three months of user data, covering streaming history, age and gender, music preferences (genres and moods), and habitual listening behaviours (e.g., playing songs on repeat).

Anderson et al. (2021) challenge the meta-analysis of Schäfer and Mehlhorn (2017). Using machine learning algorithms, they found associations between music listening on Spotify and each personality trait. Openness to experience was positively correlated with listening to atmospheric music (r = .14), folk (r = .15), reggae (r = .09), and Afropop (r = .11). Conscientiousness was positively correlated with funk (r = .12), easy listening (r = .06), and "romantic" music (r = .09) and negatively correlated with listening to rock (r = -.08), comedy (r = -.13), alternative genres (r = -.08), "energizing" (r = -.06), and "excited" moods (r = -.06). Extraversion was positively correlated with listening to funk (r = .12), reggaeton (r = .11), and "sensual" music (r = .08) and negatively correlated with listening to rock (r = -.11), metal (r = -.10), and "urgent" music (r = -.10). Agreeableness was positively correlated with listening to jazz (r = .12), soul (r = .12), and "sophisticated" music (r = .08). The trait is negatively correlated with listening to punk (r = -.10), death metal (r = -.10), or other "aggressive" music (r = -.12). Emotional stability (opposite of neuroticism) correlated positively with blues (r = .09), old country (r = .09), soul (r = .08), and music with "lively" moods (r = .05). The trait correlated negatively with "brooding" (r = -.09) and "defiant" moods (r = -.08), as well as indie (r = .10), emo (r = -.16), and regional music from Korea (r = -.08). In summary, a lot of correlations were found, however they are all small in effect size.

Similar to Anderson et al. (2021), Sust et al. (2023) collected a smartphone sensing dataset (N = 330, M = 22.42 years) spanning between 3 and 85 days along with the Big Five personality of each user. The dataset included apps such as Spotify, Android Music and Google Play Music. Using machine learning algorithms, the researchers found only openness to be significantly predicted from music listening variables (technical audio characteristics, textual

lyrics, listening duration which all together resulted in 844 variables) with a correlation of .25. They did not find significant associations with the other four traits. The researchers conclude that their results align more closely with the meta-analysis of Schäfer and Mehlhorn (2017), despite the similar behaviour based approach as Anderson et al. (2021).

In sum, in self-reported studies, the numerous relationships between the MUSIC dimensions and the trait openness are considered trivial, as it is suggested that individuals high in openness tend to enjoy a wide variety of music (Schäfer & Mehlhorn, 2017). However, Nave et al. (2018) provide evidence for an association between the MUSIC model in its whole and each of the Big Five Personality traits. The behaviour-based study by Anderson et al. (2021) contradicts the findings of Schäfer and Mehlhorn (2017) and it also provides evidence of associations between each Big Five Personality trait and music listening. However, a similar behaviour-based study by Sust et al. (2023) presents conflicting findings but acknowledge that Anderson et al.'s (2021) study is more methodologically robust due to its large sample size and inclusion of demographic variables. Nevertheless, the associations in Anderson et al. (2021) are small in effect size. Taken together, these findings suggest that music preferences can offer at least some insight into people's personalities (Anderson et al., 2021; Nave et al., 2018). Overall, this suggests that music preferences can serve as relevant information or cues for making personality judgements, which will be discussed next.

The Association Between Music Preferences and Observer-Reported Personality and Observer Accuracy

In the study of Burroughs et al. (1991), four observers (N = 4) rated the personalities of twenty targets based on a list of ten record albums that each target considered most representative of their personality. Personality ratings were based on 20 bipolar adjectives. Following adjectives were accurately judged: energetic/lazy (r = .40), optimistic/pessimistic (r = .50), inhibited/uninhibited (r = .49), formal/informal (r = .61), and efficient/wasteful (r = .42).

In the study of Rentfrow and Gosling (2006) eight observers (M = 20.4 years) listened independently to the top-10 favourite songs of 74 targets (M = 18.9 years) and rated their personality. Regarding observer accuracy, the following traits were accurately judged based on the top-10 favourite songs of the targets: extraversion (r = .27), agreeableness (r = .21), emotional stability (r = .23), and openness to experience (r = .47). Observer accuracy was not significant for conscientiousness (r = -.01). The correlation table for impression formation was not reported in the research article. This table was requested from the first author, P. Rentfrow

(personal communication, 2025). Significant correlations ($p < .05$) were found between extraversion and jazz ($r = -.29$), country ($r = .24$), electronica ($r = .27$), soul ($r = .48$) and hip-hop ($r = .33$). Correlations were found between agreeableness and classical music ($r = .28$), heavy metal ($r = -.32$), country ($r = .29$), pop ($r = .25$) and religious music ($r = .24$). Conscientiousness was correlated with classical music ($r = .49$), heavy metal ($r = -.36$), pop ($r = .23$) and religious music ($r = .27$). Neuroticism correlated significantly with country music ($r = -.32$) and finally, openness correlated significantly with classical music ($r = .29$), alternative music ($r = .37$), rock ($r = .28$), country ($r = -.32$), pop ($r = -.24$) and soul ($r = -.24$).

A follow-up study (Rentfrow & Gosling, 2007) examined personality judgements based on 14 prototypical music fans, meaning that no actual individuals were judged, only the prototypical music fan of 14 genres. Observers generally perceived classical and religious music fans as high in agreeableness, conscientiousness, and emotional stability, while rock and rap fans were seen as high in extraversion but lower in conscientiousness. They also found observers to make accurate personality judgements based on the genres blues, folk, jazz, heavy metal, rock, religious and electronic music (Rentfrow & Gosling, 2007). Observer accuracy is not known from this study.

Nevertheless, more research is needed. In the Lens Model (Brunswik, 1955), the focus is on cues that are expressed and observed. However, there has been little research on music preferences that are expressed and observed in real life, which leads us to the possibility of using Spotify Wrapped as cues.

Spotify Wrapped: A Widely Shared Insight into Music Preferences and Personality Judgements

Recently, music preferences have been communicated through sharing Spotify Wrapped. Spotify Wrapped is a compilation of the user its music listening behaviour, created by Spotify. This feature presents a vibrant summary of a user's top five artists, songs, and genres (Wikipedia contributors, n.d.). Users frequently share these compilations on social media platforms like Facebook, Instagram, and TikTok. In 2022, nearly 160 million people shared their Spotify Wrapped on their Instagram stories and there were over 400 million mentions of Spotify Wrapped on X within three days after release (Nagy, 2023). Inspired by Funder's (1995) Realistic Accuracy Model (RAM), the theoretically essential steps to achieve observer accuracy are evaluated. Music preferences appear to reveal something about a person's personality, thereby fulfilling the stage of *relevance* (Funder, 1995). Individual music preferences have

traditionally been scarcely available, but with the emergence of Spotify Wrapped and the possibility to share this on social media, the *availability* stage (Funder, 1995) is also fulfilled. Therefore, music preferences do not only appear to be relevant, but also available for personality judgement (RAM; Funder, 1995). The remaining question is whether this information is also *detected* by observers, and whether they actually *utilize* music preferences to form impressions of personality, these are the final two stages in Funder's model that are theoretically necessary to achieve observer accuracy. These last two stages represent the impression formation association of Brunswik's Lens Model (1955).

Current Research

While several studies have investigated the association between personality and music preferences (personality expression), far less research has examined impression formation processes and observer accuracy from music preferences. Rentfrow and Gosling (2006) had observers listen to the preferred music of targets. However, no research has been done on music preferences that are expressed and observed similar to a real life situation. Given the recent availability of music preferences through Spotify Wrapped, understanding how these cues contribute to personality judgement is crucial. The present study aims to study 1) personality expression (cue validity), 2) impression formation (cue utilization) and 3) observer accuracy.

The current study consists of two parts. In the first part, a questionnaire was distributed in which participants were asked to upload a screenshot of their top five Spotify Wrapped songs, rate their preference for 14 music genres, and complete a Big Five personality questionnaire. Only the music preferences from the Spotify Wrapped are behaviour-based. For the second part of the study, music profiles were created using the material collected in the first part. Each profile included the participant's top five Spotify songs, preferred genres, age group, and gender. These profiles were then presented to observers, who were asked to form a personality impression based on them. To answer the first two research questions, 14 genres and two other music features (quantity of preferred genres and the tendency to listen to the same artist) were used as a total of 16 cues to be associated with personality.

Based on the self-report studies (Greb et al., 2018; Nave et al., 2018; Schäfer & Mehlhorn, 2017), it is hypothesized that self-reported openness will associate with multiple genres. Similarly, based on Schäfer and Mehlhorn (2017) and Nave et al. (2018), extraversion is also expected to associate with some genres. It is further hypothesized that agreeableness will associate with some genres, based on the findings of Schäfer and Mehlhorn (2017) and Greb et

al. (2018). Based on Greb et al. (2018), neuroticism is also expected to associate with some genres. Finally, no associations are expected for conscientiousness, as this trait showed no associations in the self-report studies. It is expected that openness will yield the most associations with music genres, followed by extraversion, agreeableness, and neuroticism. Since in this thesis music preferences are mostly operationalized as genres, the self-reported studies do not allow to hypothesize which specific associations can be expected. However, comparisons will be made with the behaviour-based study of Anderson et al. (2021), which also made associations with genres. Finally, based on Nave et al. (2018), it is expected that openness and extraversion will be associated with a general liking for music or a preference for a wide variety of genres.

Based on the study by Rentfrow and Gosling (2006), it is expected that each observer-reported trait will be associated with some music cues. For example, it is hypothesized that soul will positively associate with observer-reported extraversion, classical music will positively associate with observer-reported conscientiousness, heavy metal will negatively associate with observer-reported agreeableness, country will negatively associate with both neuroticism and openness to experience. Nevertheless, it remains a difference to note that observers in Rentfrow and Gosling (2006) got to know people's music preferences by listening to songs, whereas the observers in this study looked at music profiles.

Based on the study of Rentfrow and Gosling (2006), it is expected that observer accuracy will be significant for all the traits except for conscientiousness.

Method

Part 1 of the study: Personality Expression

Participants

A total of 104 individuals participated in the study. The participants were selected according to the convenience sampling method. Flyers were distributed on my personal account on Facebook, Instagram and LinkedIn with the flyer (available in Appendix B) containing information about the study, the QR-code and a link to the questionnaire. The data was collected from March 2024 until July 2024. Individuals were eligible to participate if they met the following inclusion criteria: speaking Dutch, being at least 18 years old, having a Spotify account primarily used individually, and being able to retrieve a screenshot from their Spotify Wrapped for 2023. From the 104 individuals that participated in the study, 8 were excluded for failing the attention checks (see below, Muszyński, 2023). The sample consisted of 74 females, 21 males, and one nonbinary person. Table 1 provides an overview of the participants' age group distribution. The compensation was a raffle for 3 Bol.com gift cards worth €20 each. To ensure anonymity, participants were asked to fill in their email address in a separate link.

Table 1

Age and Gender Distribution of Participants

Age group (years)	Female (freq.)	Male (freq.)	Nonbinary (freq.)	Total (freq.)	Percentage of total (%)
18 – 24	47	9	1	57	59,38
24 – 29	19	6	0	25	26,04
30 – 34	0	2	0	2	2,08
35 – 39	1	1	0	2	2,08
40 – 44	4	2	0	6	6,25
45 – 49	1	0	0	1	1,04
50 – 54	1	1	0	2	2,08
55 – 59	0	0	0	0	0
60 – 64	0	0	0	0	0
65 – 69	1	0	0	1	1,04
>= 70	0	0	0	0	0
N	74	21	1	96	100%

Materials

Personality

Personality was measured using the Big Five Inventory (BFI-44; John & Srivastava, 1999), which consists of 44 items representing the Five-Factor-Model dimensions: openness, conscientiousness, extraversion, agreeableness, and neuroticism. Each item was rated on a 5-point Likert scale ranging from 1 ("Disagree strongly") to 5 ("Agree strongly"), with 3 indicating neutrality. An example of an item measuring openness is "*I see myself as someone who is original, comes up with new ideas*." An item measuring conscientiousness is "*I see myself as someone who does a thorough job.*" An item assessing extraversion is "*I see myself as someone who is talkative.*" Agreeableness was measured with items such as "*I see myself as someone who is helpful and unselfish with others*," and neuroticism was assessed with statements like "*I see myself as someone who is depressed, blue.*" The internal consistency of the scale, as measured by Cronbach's alpha (α), was .82 for openness, .72 for conscientiousness, .88 for extraversion, .74 for agreeableness, .88 for neuroticism. In the BFI-44 (John & Srivastava, 1999) an attention check was included, where participants were required to select 'agree' as a response to the item 'Indicate 'agree''. If they failed to do so, it was assumed they had not read the items properly, and they were subsequently removed from the sample.

Self-Reported Music Preferences: Genres

The selection of genres was predominantly based on the Short Test of Music Preferences (STOMP; Rentfrow & Gosling, 2003), which asks participants to rate their preferences across 14 music genres. Some genre names were adjusted to better fit the Belgian population, which has been inspired by Wikipedia's discussion of genres in Dutch (Wikipedia contributors, n.d.). For example, "pop" was adjusted to "top 40" and "electronica/dance" was modified to "trance/techno." Additionally, "blues," "alternative music," and "soundtracks" were removed, while "punk/hardcore," "gothic," and "reggae" were added. This led to the following 14 genres: heavy metal/hardrock, punk/hardcore, gothic, rock, jazz, classical music, religious music, hip-hop/rap, soul/R&B, trance/techno, top 40, country, reggae, and traditional/folk music. The preference of music genres were measured on a 5-point scale ranging from 1 ("Disagree strongly") to 5 ("Agree strongly"), an example item was "I like to listen to country". In the ratings of music genres, another attention check was included, where participants were required to select 'agree' as a response to the item 'Indicate 'agree''. If they failed to do so, it was assumed they had not read the items properly, and they were subsequently removed from the sample.

Although music preferences can be categorized into five broad factors as outlined by the MUSIC model (Rentfrow et al., 2011), previous research suggests that major music genres are familiar to 97% of people (Rentfrow & Gosling, 2007). This familiarity makes genres as the optimal approach for studying music preferences for both the targets and observers (Rentfrow & Gosling, 2003, 2007). Furthermore, the aim in creating the music profiles was to simulate Spotify Wrapped, which does not involve the MUSIC model of Rentfrow et al. (2011).

Procedure

The study was reviewed and approved by the Social and Societal Ethics Committee (SMEC) of KU Leuven. Personal data was processed in accordance with the General Data Protection Regulation (GDPR). Participants were required to provide informed consent (available in Appendix C) and agree to the information letter (available in Appendix D), both approved by SMEC. They also had to confirm that their Spotify account was primarily used individually and that they were at least 18 years old.

After providing consent, participants were asked to report their gender and age group. They then uploaded a screenshot of their Top 5 Spotify Wrapped from 2023. Instructions were provided on how to retrieve the playlist and take a screenshot. Participants had to answer the statement *"The top 5 of my Spotify Wrapped represents my music preferences"* on a 5-point scale ranging from 1 ("Disagree strongly") to 5 ("Agree strongly"). This is to rule out that Spotify Wrapped is only for marketing purposes, which would lose the 'relevance' aspect of music preferences (Letzring & Funder, 2021). Following this, they rated their preference for 14 music genres. Finally, they completed the BFI-44 personality questionnaire (John & Srivastava, 1999). This process took approximately 10–15 minutes. The full questionnaire can be found in Appendix E.

Part 2 of the study: Impression Formation

Participants

In the second part of the study, the participants are observers. A total of 14 individuals participated in the second part of the study. Participants were selected using a convenience sampling method. Flyers containing information about the study (available in Appendix F) and links to two different questionnaires were distributed through my personal Facebook and Instagram accounts. One questionnaire (Questionnaire A) consisted of 46 music profiles, while the other (Questionnaire B) included 47 different music profiles. Two participants were excluded from the analysis: one for failing the attention check and the other for participating

after the study's end date, when data analysis had already been completed. As a result, the final sample consisted of 12 participants. Six individuals completed Questionnaire A, while the remaining six completed Questionnaire B. Data collection took place from September 2024 to November 2024. Participants were compensated with the chance to win a €20 Bol.com gift card in a raffle. Tables 2 and 3 show the age groups and gender of the observers in Questionnaire A and Questionnaire B.

Table 2

Observers Who Participated in Questionnaire A

Age group	Female	Male	Total
18 – 24	1	1	2
25 – 29	2	1	3
30 – 34	1	0	1
N	4	2	6

Table 3

Observers Who Participated in Questionnaire B

Age group	Female	Male	Total
18 – 24	3	0	3
25 – 29	2	1	3
N	5	1	6

Materials

Music Profiles (Targets)

In this part of the study, participants of the first part of the study are referred to as 'targets'. Three targets were excluded because they uploaded incorrect screenshots (for example a screenshot of different playlists). However, two individuals with unusual uploads, a screenshot containing only four numbers and another from the 2022 Spotify Wrapped, were retained. As a result, 93 music profiles were created.

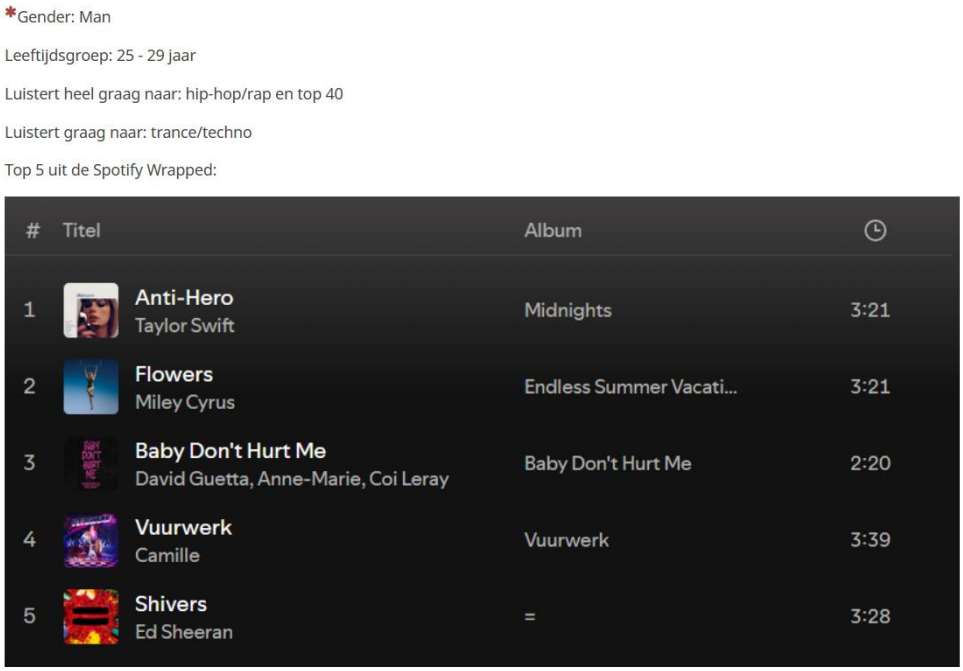
In the creation of music profiles, an attempt was made to simulate Spotify Wrapped, which typically includes the top 5 artists, songs, and genres (Wikipedia contributors, n.d.). However, this widely shared format on social media eventually disappears, leaving users with only a 100-song Spotify Wrapped playlist. That is why targets were asked to self-report their genre preferences, so that a similar format could be created. In the music profiles, the genres

that the targets rated as a 4 (agree) or 5 (strongly agree) were included. Since the uploaded screenshots had different layouts, they were converted to a standardized format.

The music profile thus consisted of: 1) the gender of the target 2) the age group of the target 3) the preference of genres that the targets rated as ‘agree’ or ‘strongly ‘agree’ and 4) their top 5 of the Spotify Wrapped (Figure 3).

Figure 3

Example of a Music Profile



Rating impressions of 93 music profiles is a large task, so to diminish response burden (Rolstad et al., 2011), the decision was made to split them into two questionnaires that was each administered to half of the observers: Questionnaire A, which included music profile 1 – 46, and Questionnaire B, which included music profile 47 – 93. In the same way as in the first part of the study, two attention checks were included in each of the questionnaire. The process for completing the questionnaire took approximately 80 minutes. Figure 3 shows the participant flow: the first six observers saw music profiles 1–46, and six other observers saw music profiles 47–93.

Figure 4

Participant Flow

Observer	Seen Questionnaire
Observer 1	Questionnaire A
…	Questionnaire A
Observer 6	Questionnaire A
Observer 7	Questionnaire B
…	Questionnaire B
Observer 12	Questionnaire B

Personality

The observers made personality judgements with the Ten Item Personality Inventory (TIPI, Hofmans et al. (2008)), a short scale representing the Five-Factor-Model dimensions: openness, conscientiousness, extraversion, agreeableness, and neuroticism. Each item was rated on a 7-point scale ranging from 1 (“Does not describe the person at all”) to 7 (“Describes the person very well”). The inventory includes both positively and negatively worded items, with the latter being reverse-scored (denoted by "R"). Extraversion was measured with *"Extravert, enthusiastic" (item 1)* and *"Reserved, quiet" (item 6; R)*. Agreeableness was measured with *"Critical, quarrelsome" (item 2; R)* and *"Sympathetic, warm" (item 7)*. Conscientiousness was measured through *"Dependable, self-disciplined" (item 3)* and *"Disorganized, careless" (item 8; R)*. Neuroticism was assessed with *"Anxious, easily upset" (item 4)* and *"Calm, emotionally stable" (item 9; R)*. Finally, openness to experience was measured by *"Open to new experiences, complex" (item 5)* and *"Conventional, uncreative" (item 10; R)*. The Cronbach’s α is averaged across the observers and was .86 for extraversion, .69 for agreeableness, .52 for conscientiousness, .62 for neuroticism and .75 for openness.

Procedure

The study was reviewed and approved by the SMEC of KU Leuven. The personal data has been processed in accordance with GDPR. The participants agreed on the informed consent (available in Appendix G) and information letter (available in Appendix H) which were approved by the SMEC.

After providing consent on the informed consent and information letter, participants were asked to report their gender and age group (starting from at least 18 years old). Then they got the following explanation:

"*Spotify releases an annual list of its users' favourite songs, called the Spotify Wrapped. This is widely shared on Instagram, TikTok, Facebook, etc., around November. In this questionnaire, you are asked to form a personality impression based on someone's music listening behaviour, specifically their top 5 from Spotify Wrapped. You will also be provided with their gender, age group, and the genres they enjoy listening to. This information is self-reported. Note: A distinction is made between "very much like" and "like" when it comes to listening preferences. For example, person X may very much enjoy listening to soul/R&B and classical music, and like listening to Top 40, punk, and rock genres. Furthermore, the genres in a list are considered equal. This means that person X enjoys Top 40, punk, and rock equally, and Top 40 is not preferred over punk and rock. Good luck!*"

Next, participants viewed a music profile and were asked to form a personality impression using the TIPI questionnaire (Hofmans et al., 2008). The process of impression formation was carried out for 46 or 47 music profiles, depending on the questionnaire completed. An example of this can be found in Appendix I.

Coding of Cues

In order to be able to calculate correlations between personality and music preferences, music preferences were coded into cues. This process of coding will be discussed for the two associations separately.

Personality Expression

As previously mentioned, genre preferences were measured on a 5-point scale, so the following coding's were used:

- The participant does not enjoy listening at all = 1.
- The participant does not enjoy listening = 2.
- The participant neither enjoys nor dislikes listening = 3.
- The participant does enjoy listening = 4.
- The participant does really enjoy listening = 5.

Some participants indicated that they enjoy many different music genres (rating them 4 or 5), while others rated only a few or no genres that highly. This may indicate their overall interest in music. Therefore, the cue ‘genre quantity’ was created to reflect how many genres a participant enjoys. For example, if someone rated 3 out of 14 genres with a 4 or higher, they received a score of 3 on ‘genre quantity.’ If they rated 10 genres with a 4 or higher, their score would be 10, and so on.

Based on the top 5 of Spotify Wrapped, one cue was created: the ‘same artist’ cue. In some top 5 lists, the same artist appeared five times (available in Appendix I), while in others, all five artists were different (Figure 3). The following coding approach was used for this cue:

- Five different artists (Artist A, Artist B, Artist C, Artist D, Artist E) = 1;
- Two instances of the same artist and three different artists (Artist A, Artist A, Artist B, Artist C, Artist D) = 2;
- Three instances of the same artist and two different artists (Artist A, Artist A, Artist A, Artist B, Artist C) = 3;
- Two artists appear twice, and one different artist (Artist A, Artist A, Artist B, Artist B, Artist C) = 3;
- Three instances of one artist and two instances of another artist (Artist A, Artist A, Artist A, Artist B, Artist B) = 4;
- Four instances of the same artist (Artist A, Artist A, Artist A, Artist A, Artist B) = 4;
- Five instances of the same artist (Artist A, Artist A, Artist A, Artist A, Artist A) = 5.

Impression Formation

As previously mentioned, the genres shown on the music profiles represent the target's preferred genres, with a distinction between 'enjoys listening to' and 'really enjoys listening to.' To establish 'impression formation' correlations, the genres were coded as follows:

- The target does not enjoy listening at all = 0.
- The target does not enjoy listening = 0.
- The target neither enjoys nor dislikes listening = 0.
- The target does enjoy listening = 1.
- The target does really enjoy listening = 2.

The first three categories are coded as 0 because they were not shown in the music profiles.

The cue ‘genre quantity’ has been coded the same way as described in the section ‘personality expression’.

A ‘same artist’ cue based on the top 5 Spotify Wrapped has been coded the same way as described in the section ‘personality expression’

Design and Analytic Strategy

The analytic strategy is the following:

1. Personality expression (cue validity) will be calculated using Pearson’s correlation and additionally, to correct for multiplicity, the Bonferroni correction will be applied.
2. Next, impression formation (cue utilization) will also be calculated using Pearson’s correlation, with both the Bonferroni and Benjamini-Hochberg corrections (Streiner, 2015) applied.
3. Then, for interobserver agreement we will be using Fleiss’ Kappa (Conger, 1980). On the contrary of Cohen’s Kappa which measures interobserver agreement between two raters, Fleiss’ Kappa measures interobserver agreement for multiple raters on categorical data (Conger, 1980). Finally, Pearson’s correlation will also be used to calculate observer accuracy.

Results

Personality Expression

As response to the statement "The top 5 of my Spotify Wrapped represents my music preferences" people rather agreed with the compilation made by Spotify ($M = 3.80$, $SD = 0.85$), ruling out that this compilation is (purely) for marketing purposes.

Personality expression (Table 4) is presented using Pearson's correlation coefficients between self-reported personality traits and music preferences coded through the following cues: 'quantity of preferred genres', and the 'same artist in the top 5'.

Extraversion had a small, positive correlation with a preference for religious music, indicating that individuals who enjoy religious music tend to score higher on extraversion.

Agreeableness had a small negative correlation with both punk/hardcore and gothic music, suggesting that individuals who enjoy these genres tend to score lower on agreeableness. Conscientiousness was negatively correlated with heavy metal/hardrock, punk/hardcore, gothic, rock and reggae. These findings suggest that individuals who prefer these genres tend to score lower on conscientiousness. Except for reggae, all the correlations reflect a moderate effect size.

Neuroticism was positively correlated with a preference for gothic music, implying that individuals that like to listen to gothic music tend to score higher on neuroticism. Neuroticism was negatively correlated with a preference for jazz music, indicating that individuals who enjoy jazz tend to score lower on neuroticism, or in other words, tend to be more emotionally stable. Both correlations reflect a small effect size.

Openness to experience showed a small positive correlation with jazz, classical music, religious music, and traditional/folk music, suggesting that individuals who enjoy listening to these genres tend to score higher on openness. Additionally, openness showed a small negative correlation with top 40 music, indicating that individuals who enjoy listening to top 40 songs tend to score lower on openness to experience.

For the quantity of preferred genres, a moderate and negative correlation was found with conscientiousness. This suggests that individuals who score higher on conscientiousness tend to prefer a narrower range of musical genres.

Finally, regarding the presence of the same artist in the top 5 of Spotify Wrapped, agreeableness showed a small positive correlation. This suggests that individuals with higher levels of agreeableness are more likely to have the same artist appear multiple times in their top 5. In contrast, neuroticism showed a small negative correlation with this cue, indicating that

individuals with higher neuroticism scores are less likely to have repeated artists in their top 5. In other words, people who are more emotionally stable tend to have more occurrences of the same artist in their top 5 Spotify Wrapped.

Table 4

Personality Expression and Impression Formation

Personality expression (N = 96)						Impression formation (targets = 93)				
Ext.	Agr.	Cons.	Neur.	Open.	Cues	Ext.	Agr.	Cons.	Neur.	Open.
.25	-.05	-.31**	.06	.01	Heavy Metal/Hardrock	-.15	**-.32****	-.17	.21*	.16
-.15	-.24*	**-.45****	.07	.06	Punk/Hardcore	-.19	**-.29****	-.15	.13	.12
-.06	-.28**	-.42**	.25*	.07	Gothic	-.11	**-.45****	-.19	.18	.03
-.04	.00	-.30**	-.03	.17	Rock	.04	-.16	.11	-.02	.27**
.06	.10	-.16	-.26*	.28**	Jazz	**-.31****	.00	.18	-.23*	.19
-.02	.08	-.12	-.15	.21*	Classical music	-.23*	.01	**.28****	-.27**	.20
.27**	.01	-.02	-.03	.29**	Religious music	**-.29****	.08	**.36****	.06	.02
.17	-.05	-.12	-.10	.03	Hip-hop/Rap	.24*	-.06	-.06	-.10	.19
.19	.03	-.12	-.04	.10	Soul/R&B	-.02	.04	.14	-.07	.16
.05	-.14	.03	.06	-.19	Trance/Techno	.11	-.23*	**-.29****	.00	.03
.03	-.02	.14	.16	-.25*	Top 40	.20*	.24*	-.10	.17	-.16
.17	.15	-.02	-.10	.06	Country	-.11	.13	0.19	.02	.05
.17	-.01	-.22*	-.06	.18	Reggae	.11	-.04	-0.08	-.08	**.29****
.09	.02	-.19	-.05	.26*	Traditional music/folk	-.17	-.07	.21*	-.04	.22*
.14	-.08	**-.39****	-.05	.16	Quantity of preferred genres	-.07	-.12	.03	-.10	**.38****
.01	.24*	-.01	-.22*	.03	Same artist in the top 5	-.09	.00	.08	.08	**-.44****

Note. Ext. = Extraversion; Agr. = Agreeableness; Cons. = Conscientiousness; Neur. = Neuroticism; Open. = Openness. * p < 0.05, two-tailed. ** p < 0.01, two-tailed.

Same artist cue is calculated with 93 participants for personality expression. Personality expression correlations in bold indicate significance after Bonferroni correction. Impression formation correlations in bold indicate significance after Benjamini-Hochberg correction.

Since multiple correlations have been calculated, there is an increased chance for finding a significant effect by chance, this is the Type I error (Streiner, 2015). Therefore, the Bonferroni correction was applied to the five traits × the 14 genres and the cue genre quantity. A separate correction was applied for the correlation between the five traits and the cue same artist, as these calculations were performed on 93 targets instead of 96. A Bonferroni correction on five

traits x 15 cues yielded a new alpha of .05/75 = .00067. After this correction, only two correlations remained significant, which were between conscientiousness and punk/hardcore and conscientiousness and cue genre quantity. A Bonferroni correction on the five traits x cue same artist yielded a new alpha op 0.05/5 = 0.01. There were no significant correlations after this correction.

Assumption Check for Combining Participant Groups

The two groups of observers have rated two different profiles presented in questionnaires A and B. We want to combine the two groups of observers in the statistical analyses of impression formation and observer accuracy. However, according to the RAM of Funder (1995), detection and utilization of cues is primarily influenced by the judge itself, which affects observer accuracy (Letzring & Funder, 2021). Furthermore, based on an overview of Letzring et al. (2021) on personality judgement research, it is known that observers differ from one another in their personality judgements. Therefore we will have a look at the way personality ratings of the two observer groups differ (Table 5).

Table 5

Descriptives of the Two Observer Groups their FFM ratings

	Observers' ratings Who Rated Questionnaire A		Observers' ratings Who Rated Questionnaire B	
FFM personality dimensions	*M*	*SD*	*M*	*SD*
Extraversion	4.41	.78	4.12	.77
Agreeableness	4.83	.68	4.49	.51
Conscientiousness	4.61	.72	4.16	.60
Neuroticism	3.81	.51	4.01	.65
Openness	4.68	.72	4.03	.87

Calculated with the independent samples T-test, there is a statistically significant difference between the two observer groups in their ratings of agreeableness $t(91) = 2.81$, $p = .006$, $g = .58$, 95% CI [.16, .99], conscientiousness $t(91) = 3.31$, $p < .001$, $g = .67$, 95% CI [.26, 1.09] and openness $t(91) = 3.93$, $p < .001$, $g = .81$, 95% CI [.39, 1.23]. There was no significant difference in the ratings for extraversion, $t(90.80) = 1.78$, $p = .078$, $g = .37$, 95% CI [-.04, .77] and neuroticism $t(91) = -1.67$, $p = .099$, $g = -.34$, 95% CI [-.75, .07].

However, it is not known what the reason is for these differences. Is this because of major differences in presented music profiles or are there differences between the observer groups, or both of them? We will have a look at possible differences in presented music profiles.

Comparison Between Questionnaire A and B

Observers who rated questionnaire B saw more variety in gender and age groups (Table 6). They also saw more of the genres heavy metal/hardrock and rock (Table 7). On the contrary, observers who rated questionnaire A saw more of the genres classical music and traditional music/folk (Table 7). However, other genres such as gothic and jazz did not differ, and hip-hop/rap and the top 40 seemed to differ only little (Table 7). Nevertheless, each screenshot was unique, and therefore judges who rated questionnaire A versus questionnaire B really saw different music profiles.

Table 6

Gender and Age Group Frequencies in Questionnaire A and Questionnaire B

Demographics		Frequency in Questionnaire A	Frequency in Questionnaire B
Gender	Females	37	34
	Males	9	12
	Nonbinary	0	1
Age group	18 – 24	30	24
	25 – 29	16	9
	30 – 34	0	2
	35 – 39	0	2
	40 – 44	0	6
	45 – 49	0	1
	50 – 54	0	2
	55 – 59	0	0
	60 – 64	0	0
	65 – 69	0	1
	70 or older	0	0

Table 7

Music Genre Frequencies in Questionnaire A and Questionnaire B

Genres	Frequency in Questionnaire A	Frequency in Questionnaire B
Heavy Metal/Hardrock	7	13
Punk/Hardcore	6	8
Gothic	4	4
Rock	20	33
Jazz	15	18
Classical music	13	9
Religious music	4	4
Hip-hop/rap	24	26
Soul/R&B	22	27
Trance/techno	15	14
Top 40	22	21
Country	11	14
Reggae	9	12
Traditional music/folk	18	7

Since the two questionnaires differ from each other and the two observer groups also differ on their ratings on three traits, we want to remove the average differences between the two observer's group scores so we can combine them in our next analyses. The mean differences between the two observer groups have been erased by centering their scores around each observer group's average scores. Centering was done by subtracting the means (as shown in Table 5) from the observer ratings on each trait within each group. This centering effectively removed the average differences between the two observer's group scores: Extraversion, $t(91) = .00$, $p = 1.00$, 95% CI [–0.32, 0.32], Agreeableness, $t(91) = .00$, $p = 1.00$, 95% CI [–0.25, 0.25], Conscientiousness, $t(91) = .00$, $p = 1.00$, 95% CI [–0.27, 0.27], Neuroticism, $t(91) = .00$, $p = 1.00$, 95% CI [–0.24, 0.24], and Openness, $t(91) = .00$, $p = 1.00$, 95% CI [–0.33, 0.33].

These centered scores will be used in the impression formation correlations and observer accuracy correlations.

Impression Formation

Impression formation (Table 4) is presented using Pearson's correlation coefficients between observer-reported personality traits and music preference cues.

In terms of extraversion, people who enjoy listening to jazz, classical music and religious music were perceived as less extraverted presenting a small effect size. In contrast,

listeners of hip-hop/rap and top 40 music were perceived as more extraverted, both with small effect sizes.

For agreeableness, several music genres were significantly related to lower perceptions of this trait. Listeners of punk/hardcore and trance/techno were perceived as less agreeable, both showing small effect sizes. Notably, heavy metal/hardrock listeners and gothic music listeners were viewed as less agreeable, both reflecting a moderate effect size. In contrast, top 40 music listeners were perceived as more agreeable, reflecting a small effect size.

In terms of conscientiousness, those who enjoy listening to trance/techno were perceived as less conscientious, representing a small effect size. On the other hand, classical music listeners, religious music listeners and traditional music/folk listeners were perceived as more conscientious. Among these, religious music was associated with a moderate effect size, while the others reflected small effects.

Regarding neuroticism, heavy metal/hardrock listeners were perceived as more neurotic which represents a small effect size. In contrast, jazz listeners and classical music listeners were viewed as less neurotic, both reflecting small effect sizes.

Regarding openness, people who enjoy listening to rock, reggae and traditional music/folk were all perceived as higher in openness, with each correlation reflecting a small effect size.

In addition to the correlations between music genres and perceived personality traits, specific music-related cues were also associated with perceived openness. Perceived openness correlated positively with genre quantity, suggesting that participants who enjoyed listening to a greater number of music genres were perceived as more open to experiences. This correlation reflects a moderate effect size. A moderate, significant negative correlation was found between perceived openness and the cue same artist, indicating that when the top 5 Spotify Wrapped more frequently contained the same artist, the person was perceived as less open to experiences.

Punk/hardcore and gothic were valid cues to rate a target's lower agreeableness. Jazz was a valid cue to rate a target's lower neuroticism. Finally, traditional music/folk was a valid cue to rate a target's openness to experience.

Again, to correct for the occurrence of Type I error, a Bonferroni correction was applied to the correlations between the 5 traits and 16 cues, yielding a new alpha of .05/80 = .000625. This resulted in four remaining significant correlations: a negative correlation between the cue

same artist and perceived openness, a positive correlation between the cue genre quantity and perceived openness, a negative correlation between gothic and perceived agreeableness, and a positive correlation between religious music and perceived conscientiousness. However, the Bonferroni correction is very conservative, the decreased probability of the Type I error has the downside of an increased probability of the Type II error (Streiner, 2015). Since studies of impression formation on music preferences are rather scarce, we do not want to miss out on any important effects. The Benjamini-Hochberg correction is better at separating the important few effects from the many trivial effects and has a higher power (Streiner, 2015). This correction yielded seven additional significant correlations: heavy metal/hardrock and lower perceived agreeableness, jazz and lower perceived extraversion, trance/techno and lower perceived conscientiousness, reggae and higher perceived openness, religious music and lower perceived extraversion, punk/hardcore and lower perceived agreeableness, classical music and higher perceived conscientiousness.

Interobserver Agreement and Observer Accuracy

Interobserver agreement among the six observers that rated personality traits was calculated using Fleiss' Kappa. Fleiss' Kappa varies from -1 to +1, a negative value indicates poorer than chance agreement, zero indicates exactly chance agreement and a positive value indicates a better than chance agreement (Fleiss & Cohen, 1973). Six observers rated the targets' personality from questionnaire A (the first half of the targets) and six from questionnaire B (the second half of the targets). The results are as follows for observers rating targets in questionnaire A: extraversion, $\kappa = -.012$ ($p = .335$); agreeableness, $\kappa = .022$ ($p = .116$); conscientiousness, $\kappa = .044$ ($p = .002$); neuroticism, $\kappa = .001$ ($p = .959$); and openness, $\kappa = -.016$ ($p = .199$). For observers rating targets in questionnaire B, the results are as follows: extraversion, $\kappa = .010$ ($p = .449$), agreeableness, $\kappa = -.007$ ($p = .637$), conscientiousness, $\kappa = -.015$ ($p = .268$), neuroticism, $\kappa = -.020$ ($p = .095$), and openness, $\kappa = -.007$ ($p = .575$). These results indicate chance agreement between observers on the music profiles.

To calculate observer accuracy, the average ratings from the six observers for each personality trait were correlated with the self-reported personality traits using Pearson's correlation. The correlations were interpreted following Cohen's (1988) guidelines. The results showed very small and non-significant correlations for extraversion ($r = .03$, $p = .780$) and agreeableness ($r = .05$, $p = .618$), indicating that observer ratings did not align with self-reports for these traits. In contrast, conscientiousness ($r = .22$, $p = .033$), neuroticism ($r = .24$, $p = .019$)

and openness (r = .21, p = .046) showed small but statistically significant correlations, suggesting that observers had some degree of accuracy in judging these traits.

Discussion

This thesis aimed to study personality expression, impression formation, and observer accuracy based on music preferences of the top 5 Spotify Wrapped and preferences for music genres. For personality expression, participants' music preferences and their personality were collected. To study impression formation and observer accuracy, music profiles were made, consisting of the targets' genre preferences, their top 5 songs from Spotify Wrapped, and their gender and age group. These profiles were then shown to observers, which viewed targets' music profiles and made personality impressions.

Summary of the Results

Personality Expression

Regarding personality expression, it was hypothesized that openness would be associated with preferences for multiple music genres. This was also hypothesized for the traits extraversion, agreeableness, and neuroticism, although a lower number of associations was expected compared to openness. It was also hypothesized that conscientiousness would not be associated with any of the genres. These hypotheses were based on self-report studies (Greb et al., 2018; Nave et al., 2018; Schäfer & Mehlhorn, 2017). Furthermore, for the specific associations, comparisons will be made with the behaviour based study of Anderson et al. (2021), since the self-reported studies studied music preferences on other levels than of genres.

Multiple significant correlations with the trait openness are found, which aligns with the hypothesis that openness would associate with multiple music genres. Specifically, openness correlates positively with jazz, classical music, religious music, and folk/traditional music, and negatively with top 40 music. The association with folk music is found in the study by Anderson et al. (2021), but the other associations were not.

For extraversion, one significant correlation is found. The occurrence of at least one association is as expected from the self-report studies. Specifically, extraversion is positively associated with a preference for religious music, an association that was not found in Anderson et al. (2021).

Furthermore, agreeableness correlates negatively with a preference for punk/hardcore and gothic music. The occurrence of some associations is as expected from the self-report studies. The first association is also found in the study of Anderson et al. (2021).

Neuroticism is positively associated with a preference for gothic music and negatively with a preference for jazz. The occurrence of some associations is as expected from the self-report studies. Furthermore, the associations are not found in the study of Anderson et al. (2021).

However, the trait conscientiousness shows more associations than openness, while it was hypothesized that there would not be a single association (based on the self-report studies). There is a correlation between lower conscientiousness and preference for heavy metal/hardrock, punk/hardcore, gothic, rock and reggae. The negative correlation between conscientiousness and rock aligns with the study of Anderson et al. (2021).

It was also hypothesized that openness and extraversion would be associated with the cue 'preference for a large quantity of genres', however, this was not found. Instead, the cue was associated with lower conscientiousness. This may reflect a tendency to switch frequently, and therefore an indication of lower goal-directedness.

Furthermore, a hypothesis for the association between the cue 'same artist' and personality traits was not formulated. However, the same artist cue is associated with a higher agreeableness and lower neuroticism.

After the Bonferroni correction (Steiner, 2015), two correlations remain significant: a preference for many genres appears to be a robust indication of someone who is less conscientious. Furthermore, enjoying listening to punk/hardcore is a robust indication of someone lower in conscientiousness. It is interesting that a preference for a wide range of genres, which does not reflect the specific content of music preferences, is associated with lower conscientiousness.

In summary, the results generally suggest that personality is somewhat expressed through music preferences, however, these associations are not robust. After applying the Bonferroni correction, only two associations remain, both involving the trait conscientiousness. This suggests that, in this thesis, there is limited evidence for personality being expressed through music preferences.

Impression Formation

Regarding impression formation, it was hypothesized that each observer-reported trait would associate with some music genre cues. The hypotheses of specific associations were

based on the study by Rentfrow and Gosling (2006). Indeed, there is a correlation for each observer-reported trait with some genres.

There is a negative correlation between observers' ratings of extraversion and jazz, classical, and religious music. The first association aligns with Rentfrow and Gosling (2006), but the others do not. Furthermore, ratings of extraversion are positively correlated with hip-hop/rap and top 40 music. The first correlation aligns with Rentfrow and Gosling (2006), whereas the second does not.

Next, observers' ratings of agreeableness were negatively correlated with heavy metal/hardrock, punk/hardcore, gothic, and trance/techno. Only the first association aligns with the study by Rentfrow and Gosling (2006). Observer-rated agreeableness is also positively correlated with top 40 music, which aligns with the findings by Rentfrow and Gosling (2006).

Observers' ratings of conscientiousness are negatively correlated with trance/techno. This association does not align with the study by Rentfrow and Gosling (2006). Furthermore, ratings of conscientiousness are positively correlated with classical music, religious music, and traditional music/folk. The first two positive correlations align with the study by Rentfrow and Gosling (2006).

Observers' ratings of neuroticism were negatively correlated with jazz and classical music, and positively correlated with heavy metal/hardrock. None of these associations were found in the study by Rentfrow and Gosling (2006).

Finally, observers' ratings of openness were positively correlated with rock, reggae, and traditional music/folk. Only the first association aligns with the findings of Rentfrow and Gosling (2006).

No hypotheses were made regarding the associations between traits and the cues genre quantity and same artist. Observers rate people who listen to many genres as high in openness and people who tend to listen to the same artist as low in openness. This can be explained by the expectation that people high in openness enjoy a broad variety of genres and tendency to listen to different artists. However, these cues do not provide information about openness to experience.

After applying the Benjamini-Hochberg correction (Streiner, 2015), eleven correlations remain significant: religious music and lower perceived extraversion, jazz and lower perceived extraversion (which aligns with Rentfrow and Gosling, 2006), gothic and lower perceived

agreeableness, heavy metal/hardrock and lower perceived agreeableness (which aligns with Rentfrow and Gosling, 2006), punk/hardcore and lower perceived agreeableness, trance/techno and lower perceived conscientiousness, religious music and higher perceived conscientiousness (which aligns with Rentfrow and Gosling, 2006), classical music and higher perceived conscientiousness (which aligns with Rentfrow and Gosling, 2006), reggae and higher perceived openness, the cue "same artist" and lower perceived openness, the cue "genre quantity" and higher perceived openness.

In summary, multiple associations are found, and this remains partly the case even after applying the Benjamini-Hochberg correction (Streiner, 2015). People indeed form personality impressions based on multiple musical cues. This is especially true for the traits openness, conscientiousness, and agreeableness, each of which shows a robust association with three different cues. Extraversion also remains associated with two cues after the correction, but no robust associations remain with the trait neuroticism.

Observer Accuracy

It was hypothesized that observer accuracy would be significant for all the traits except for conscientiousness. The results show that observers are modestly accurate in rating conscientiousness, neuroticism, and openness based on people's music preferences. The first correlation does not align with the hypothesis, whereas the last two do. It appears that people are indeed able to form impressions that somewhat correspond to the actual (self-reported) personality traits based on their music preferences as reflected in their Spotify Wrapped and preferred genres. It is noteworthy that, based on such minimal information, observers are still able to form a modestly accurate impression of three personality traits. The information is not only very minimal, from research it is known that music preferences are correlated with personality, but weakly, which is the case in Schäfer and Mehlhorn (2017) and Anderson et al. (2021). This makes the results even more impressive. It is possible that this finding occurred by chance, particularly given the increased risk of Type I errors due to the family-wise error rate.

Linking to Theory

The results suggest that personality is only weakly expressed in music preferences, so there is a weak personality expression (Brunswik, 1955) and that music preference cues serve as a lens through which observers form personality impressions, so there occurs impression formation (Brunswik, 1955). For four associations, both personality expression ànd impression formation were observed (so valid cues were used). This was the case for the associations

between punk/hardcore and lower agreeableness, gothic and lower agreeableness, jazz and lower neuroticism, and traditional music/folk and higher openness. For these associations, all stages of Funder (1995) have been met. The number of valid cues is small, especially considering that observers were able to accurately judge conscientiousness, neuroticism and openness. This can be explained by not accounting for other possible cues, such as the language in titles.

Furthermore, the cues utilized by observes could have been valid, but the personality expression associations were perhaps not found in this study. For example, listeners of reggae were rated as higher in openness, but this does not appear in the personality expression in this study. However, in the study by Anderson et al. (2021), the researchers found reggae to correlate with openness to experience. In this sense, the utilization of the cue reggae to form an impression of openness seems to be valid. Another example is that listeners who enjoy top 40 music were considered higher in extraversion by observers, which does not appear in the personality expression in this study. In the meta-analysis by Schäfer and Mehlhorn (2017), people who are more extraverted listen more to contemporary music. In that sense, the utilization of the cue top 40 to rate extraversion seems to be valid.

Next, from the RAM theory explained by Letzring and Funder (2021), observers do not only determine which cues are relevant or valid, but also have to give appropriate weights to different cues and combine them. This process of giving weights and combining different cues have not been considered in this study.

While it is not very clear what specific cues have contributed to significant observer accuracy of conscientiousness, neuroticism and openness, the significant, but small correlations, indicate that observers are modestly capable at forming an impression of conscientiousness, neuroticism and openness from people's music preferences. This is very interesting, especially since the information is minimal and the correlations between personality and music preferences are weak (Anderson et al., 2021; Schäfer & Mehlhorn, 2017) and therefore of rather less relevance in personality judgement. A replication study would be important to confirm this, since the correlations are small.

Limitations

An important limitation is that interobserver agreement occurred by chance among observers. There is evidence showing that observers differ from one another in judging personality traits (Letzring et al., 2021). However, the influence of observer variance on

personality judgements is estimated at 30%, and when there is no observer-target interaction, the estimation is considerably lower (Letzring et al., 2021). So, despite the differences between observers, the lack of observer-target interaction may have downsized these differences. Why these differences do occur yields for further investigation.

The role of observers in such studies is relatively minor (Letzring et al., 2021), however, the observer still plays a role in personality judgments. In line with this, a second limitation is that the sample of observers is small. The observer ratings were averaged across six observers. This is a very small sample, and therefore this average number cannot be generalized to what observers in the general population would, on average, answer.

Furthermore, there was only one Spotify-related cue (same artist cue). More cues could provide greater insight into the cues people use to make personality judgements. Finally, gender and age group information was included in the music profiles, which could also have contributed to observer accuracy. However, gender and age group are basic pieces of information that any stranger on social media can infer, so this is representative of the external environment and part of realistic judgement as described in Funder's 'Realistic' Accuracy Model (1995).

Finally, in both parts of the study, there may have been a lack of diversity, especially since the convenience sampling method was used.

Practical Implications and Future Research

This study can be seen as an addition to research on personality judgement based on digital data. Furthermore, this research can be useful in dating apps or other new encounters where people have never seen each other and want to get an idea of who someone is. In line with this, a systematic review of all the cues that are associated with personality would be useful to know what to rely on in new encounters. Next, it is recommended to use larger groups of observers in future research. Finally it is recommended to conduct a replication study.

Conclusion

This is the first study to examine the link between music preferences, personality and personality judgements using the popular Spotify Wrapped feature. There is weak evidence for the association between personality and music preferences. After applying the Bonferroni correction, most of the associations become negligible. Next, observers appeared to rely on a variety of genre and Spotify related cues when forming personality impressions, even after the Benjamini-Hochberg correction. Finally, observers demonstrated a modest capability to

accurately judge openness, neuroticism and conscientiousness based on individuals' Spotify Wrapped and genre preferences, which is a noteworthy finding given the limited information provided and the weak association between personality and music preferences in this study, as well as in others.

References

Adolphs, R. (2009). The Social Brain: Neural Basis of Social Knowledge. *Annual Review of Psychology*, *60*(1), 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

Anderson, I., Gil, S., Gibson, C., Wolf, S., Shapiro, W., Semerci, O., & Greenberg, D. M. (2021). "Just the Way You Are": Linking Music Listening on Spotify and Personality. *Social Psychological and Personality Science*, *12*(4), 561–572. https://doi.org/10.1177/1948550620923228

Anglim, J., Dunlop, P. D., Wee, S., Horwood, S., Wood, J. K., & Marty, A. (2022). Personality and Intelligence: A Meta-Analysis. *Psychological Bulletin*, *148*(5–6), 301–336. https://doi.org/10.1037/bul0000373

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44(1)*, 1–26. https://doi.org/https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, *62*(3), 193–217. https://doi.org/10.1037/h0047470

Burroughs, W. J., Drews, D. R., & Hallman, W. K. (1991). Predicting Personality from Personal Possessions: A Self-presentational Analysis. *Journal of Social Behavior and Personality*, *6*(6), 147–163. https://kuleuven.e-bronnen.be/scholarly-journals/predicting-personality-personal-posessions-self/docview/1292260552/se-2?accountid=17215

Chapman, B. P., Roberts, B., & Duberstein, P. (2011). Personality and longevity: knowns, unknowns, and implications for public health and personalized medicine. *Journal of aging research*, *2011*, 759170. https://doi.org/10.4061/2011/759170

Choi, H., Diener, E., & Oishi, S. (2024). Do We Know How Happy Strangers Are? Accuracy in Well-Being Judgments at Zero Acquaintance. *Social Psychological and Personality Science*, *15*(6), 691–701. https://doi.org/10.1177/19485506231197844

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Conger, A. J. (1980). Integration and generalization of kappas for multiple raters. *Psychological Bulletin, 88*(2), 322–328. https://doi.org/10.1037/0033-2909.88.2.322

Connelly, B. S., & Ones, D. S. (2010). An Other Perspective on Personality: Meta-Analytic Integration of Observers' Accuracy and Predictive Validity. *Psychological Bulletin*, *136*(6), 1092–1122. https://doi.org/10.1037/a0021212

Costa, P. T., & McCrae, R. R. (1999). A five-factor theory of personality. *Handbook of personality: Theory and research*, *2*(01), 139–153. https://www.researchgate.net/publication/284978581

Darbyshire, D., Kirk, C., Wall, H. J., & Kaye, L. K. (2016). Don't Judge a (Face)Book by its Cover: Exploring judgement accuracy of others' personality on Facebook. *Computers in Human Behavior*, *58*, 380–387. https://doi.org/10.1016/j.chb.2016.01.021

Dong, Z., & Xie, T. (2023). How to know who you are through your short video selfies?-Capturing personality via short video selfies. *Frontiers in psychology*, *14*, 1–9. https://doi.org/10.3389/fpsyg.2023.1072344

Elvers, P., Omigie, D., Fuhrmann, W., & Fischinger, T. (2015). Exploring the musical taste of expert listeners: musicology students reveal tendency toward omnivorous taste. *Frontiers in psychology*, *6*, 1–11. https://doi.org/10.3389/fpsyg.2015.01252

Fleiss, J. L., & Cohen, J. (1973). The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educational and Psychological Measurement*, *33*(3), 613–619. https://doi.org/10.1177/001316447303300309

Fransson, M., Granqvist, P., Bohlin, G., & Hagekull, B. (2013). Interlinkages between attachment and the Five-Factor Model of personality in middle childhood and young adulthood: A longitudinal approach. *Attachment and Human Development*, *15*(2), 219–239. https://doi.org/10.1080/14616734.2013.754985

Funder, D. C. (1995). On the Accuracy of Personality Judgment: A Realistic Approach. *Psychological Review*, *102*(4), 652–670. https://doi.org/10.1037/0033-295X.102.4.652

Gosling, S. D., Ko, S., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology*, *82*(3), 379–398. https://doi.org/10.1037/0022-3514.82.3.379

Greb, F., Steffens, J., & Schlotz, W. (2018). Understanding music-selection behavior via statistical learning: Using the percentile-Lasso to identify the most important factors. *Music & Science*, *1*, 1–17. https://doi.org/10.1177/2059204318755950

Hall, J. A., Pennington, N., & Lueders, A. (2014). Impression management and formation on Facebook: A lens model approach. *New Media and Society*, *16*(6), 958–982. https://doi.org/10.1177/1461444813495166

Hartwig, M., & Bond, C. F. (2011). Why Do Lie-Catchers Fail? A Lens Model Meta-Analysis of Human Lie Judgments. *Psychological Bulletin*, *137*(4), 643–659. https://doi.org/10.1037/a0023589

Hinds, J., & Joinson, A. N. (2024). Digital Data and Personality: A Systematic Review and

Meta-Analysis of Human Perception and Computer Prediction. *Psychological Bulletin*, *150*(6), 727–766. https://doi.org/10.1037/bul0000430

Hofmans, J., Kuppens, P., & Allik, J. (2008). Is short in length short in content? An examination of the domain representation of the Ten Item Personality Inventory scales in Dutch language. *Personality and Individual Differences*, *45*(8), 750–755. https://doi.org/10.1016/j.paid.2008.08.004

Hyatt, C. S., Reddi, P. J., Sharpe, B. M., Michopoulos, V., van Rooij, S. J. H., House, S. L., Beaudoin, F. L., An, X., Stevens, J. S., Zeng, D., Neylan, T. C., Clifford, G. D., Linnstaedt, S. D., Germine, L. T., Bollen, K. A., Rauch, S. L., Haran, J. P., Lewandowski, C., Musey, P. I., … Maples-Keller, J. L. (2024). Longitudinal associations between five factor model and impulsive personality traits and PTSD symptoms: Findings from the AURORA study. *Journal of Research in Personality*, *112*, 1–12. https://doi.org/10.1016/j.jrp.2024.104524

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). Guilford Press. https://jenni.uchicago.edu/econ-psych-traits/John_Srivastava_1995_big5.pdf

Karelaia, N., & Hogarth, R. M. (2008). Determinants of Linear Judgment: A Meta-Analysis of Lens Model Studies. *Psychological Bulletin*, *134*(3), 404–426. https://doi.org/10.1037/0033-2909.134.3.404

Kaufmann, E., Reips, U.-D., & Wittmann, W. W. (2013). A Critical Meta-Analysis of Lens Model Studies in Human Judgment and Decision-Making. *PloS One*, *8*(12), 1–16. https://doi.org/10.1371/journal.pone.0083528

Koutsoumpis, A., Oostrom, J. K., Holtrop, D., van Breda, W., Ghassemi, S., & de Vries, R. E.

(2022). The Kernel of Truth in Text-Based Personality Assessment: A Meta-Analysis of the Relations Between the Big Five and the Linguistic Inquiry and Word Count (LIWC). *Psychological Bulletin*, *148*(11–12), 843–868. https://doi.org/10.1037/bul0000381.supp

Letzring, T. D., & Funder, D. C. (2021). The realistic accuracy model. In T. D. Letzring & J. S. Spain (Eds.), *The Oxford handbook of accurate personality judgment* (pp. 9–24). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190912529.001.0001

Letzring, T. D., Murphy, N. A., Allik, J., Beer, A., Zimmermann, J., & Leising, D. (2021). The judgment of personality: An overview of current empirical research findings. *Personality Science*, *2*. https://doi.org/10.5964/ps.6043

Letzring, T. D., Wells, S. M., & Funder, D. C. (2006). Information quantity and quality affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology*, *91*(1), 111–123. https://doi.org/10.1037/0022-3514.91.1.111

Luo, J., Zhang, B., Graham, E. K., & Mroczek, D. K. (2023). Does Personality Always Matter for Health? Examining the Moderating Effect of Age on the Personality-Health Link From Life Span Developmental and Aging Perspectives. *Journal of Personality and Social Psychology*, *125*(5), 1189–1206. https://doi.org/10.1037/pspp0000485

Malinka, J., Mitte, K., & Ziegler, M. (2025). Who (Really) Wins with Basic Income: Personality and Values as Predictors of Happiness Trajectories. *Journal of Happiness Studies*, *26*(1), 1–34. https://doi.org/10.1007/s10902-024-00831-x

Malouff, J. M., Thorsteinsson, E. B., & Schutte, N. S. (2005). The relationship between the five-factor model of personality and symptoms of clinical disorders: A meta-analysis. *Journal of Psychopathology and Behavioral Assessment*, *27*(2), 101–114. https://doi.org/10.1007/s10862-005-5384-y

Malouff, J. M., Thorsteinsson, E. B., Schutte, N. S., Bhullar, N., & Rooke, S. E. (2010). The Five-Factor Model of personality and relationship satisfaction of intimate partners: A meta-analysis. *Journal of Research in Personality*, *44*(1), 124–127. https://doi.org/10.1016/j.jrp.2009.09.004

Mammadov, S. (2022). Big Five personality traits and academic performance: A meta-analysis. *Journal of Personality*, *90*(2), 222–255. https://doi.org/10.1111/jopy.12663

Marrero, Z. N. K., Gosling, S. D., Pennebaker, J. W., & Harari, G. M. (2022). Evaluating voice samples as a potential source of information about personality. *Acta Psychologica*, *230*, 1–10. https://doi.org/10.1016/j.actpsy.2022.103740

Menuhin, Y., & Davis, C. W. (1979). *The music of man*. Macdonald General Books.

Muszyński, M. (2023). Attention checks and how to use them: Review and practical recommendations. *Ask: Research and Methods*, *32*(1), 3–38. https://doi.org/10.18061/ask.v32i1.0001

Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance. *Personality and Social Psychology Bulletin*, *35*(12), 1661–1671. https://doi.org/10.1177/0146167209346309

Nagy, D. (2023, January 6). *Spotify Wrapped: To share or not to share?* The Tower. https://sbadamsthetower.com/10721/opinion/spotify-wrapped-to-share-or-not-to-share/

Nave, G., Minxha, J., Greenberg, D. M., Kosinski, M., Stillwell, D., & Rentfrow, J. (2018). Musical Preferences Predict Personality: Evidence From Active Listening and Facebook Likes. *Psychological Science*, *29*(7), 1145–1158. https://doi.org/10.1177/0956797618761659

Osterholz, S., Mosel, E. I., & Egloff, B. (2023). #Insta personality: Personality expression in

Instagram accounts, impression formation, and accuracy of personality judgments at zero acquaintance. *Journal of Personality*, *91*(3), 566–582. https://doi.org/10.1111/jopy.12756

Poropat, A. E. (2009). A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance. *Psychological Bulletin*, *135*(2), 322–338. https://doi.org/10.1037/a0014996

Rentfrow, P. J., Goldberg, L. R., & Levitin, D. J. (2011). The structure of musical preferences: A five-factor model. *Journal of Personality and Social Psychology*, *100*(6), 1139–1157. https://doi.org/10.1037/a0022406

Rentfrow, P. J., & Gosling, S. D. (2006). Message in a Ballad: The Role of Music Preferences in Interpersonal Perception. *Psychological Science*, *17*(3), 236–242. https://doi.org/10.1111/j.1467-9280.2006.01691.x

Rentfrow, P. J., & Gosling, S. D. (2007). The content and validity of music-genre stereotypes among college students. *Psychology of Music*, *35*(2), 306–326. https://doi.org/10.1177/0305735607070382

Rolstad, S., Adler, J., & Rydén, A. (2011). Response burden and questionnaire length: Is shorter better? A review and meta-analysis. *Value in Health*, *14*(8), 1101–1108. https://doi.org/10.1016/j.jval.2011.06.003

Ruben, M. A., & Hall, J. A. (2016). A Lens Model Approach to the Communication of Pain. *Health Communication*, *31*(8), 934–945. https://doi.org/10.1080/10410236.2015.1020261

Schäfer, T., & Mehlhorn, C. (2017). Can personality traits predict musical style preferences?

A meta-analysis. *Personality and Individual Differences*, *116*, 265–273. https://doi.org/10.1016/j.paid.2017.04.061

Stephan, Y., Sutin, A. R., Luchetti, M., Canada, B., & Terracciano, A. (2022). Personality and fatigue: meta-analysis of seven prospective studies. *Scientific Reports*, *12*(1), 1–8. https://doi.org/10.1038/s41598-022-12707-2

Streiner, D. L. (2015). Best (but oft-forgotten) practices: the multiple problems of multiplicity—whether and how to correct for many statistical tests. *The American Journal of Clinical Nutrition*, *102*(4), 721–728. https://doi.org/10.3945/ajcn.115.113548

Sust, L., Stachl, C., Kudchadker, G., Bühner, M., & Schoedel, R. (2023). Personality Computing With Naturalistic Music Listening Behavior: Comparing Audio and Lyrics Preferences. *Collabra. Psychology*, *9*(1). https://doi.org/10.1525/collabra.75214

Tskhay, K. O., & Rule, N. O. (2014). Perceptions of personality in text-based media and OSN: A meta-analysis. *Journal of Research in Personality*, *49*(1), 25–30. https://doi.org/10.1016/j.jrp.2013.12.004

van Aarde, N., Meiring, D., & Wiernik, B. M. (2017). The validity of the Big Five personality traits for job performance: Meta-analyses of South African studies. *International Journal of Selection and Assessment*, *25*(3), 223–239. https://doi.org/10.1111/ijsa.12175

Vazire, S., Naumann, L. P., Rentfrow, P. J., & Gosling, S. D. (2008). Portrait of a narcissist: Manifestations of narcissism in physical appearance. *Journal of Research in Personality*, *42*(6), 1439–1447. https://doi.org/10.1016/j.jrp.2008.06.007

Wikipedia contributors. (n.d.). *Music genre*. Wikipedia. Retrieved December 31, 2024, from

https://en.wikipedia.org/wiki/Music_genre

Wikipedia contributors. (n.d.). *Spotify Wrapped*. Wikipedia. Retrieved January 20, 2025, from https://en.wikipedia.org/wiki/Spotify_Wrapped