

Are users attracted by playlist titles and covers? Understanding playlist selection behavior on a music streaming platform



Zhi Li^{a,*}, Min Song^a, Shen Duan^b, Zhe Wang^{c,*}

^a School of Economics and Management, Tiangong University, Tianjin, 300387, China

^b School of Business, Renmin University of China, Beijing 100872, China

^c School of Economics and Management, Sichuan Tourism University, Chengdu, 610100, Sichuan province, China

ARTICLE INFO

Article History:

Received 8 February 2022

Accepted 2 June 2022

Available online 22 June 2022

Keywords:

User-generated playlist (UGP)

Playlist selection

Language expectancy theory (LET)

Dual coding theory

Music streaming platform

ABSTRACT

Playlists have become the main entry point for users to obtain music resources. This study aimed to investigate the features of playlist titles and covers that attract users and the consequences of playlist selection on music streaming platforms. In this study, 7,606 playlist data were collected from the NetEase Cloud Music platform. The linguistic style of playlist titles was classified based on the language expectancy theory (LET). Artificial intelligence technology was used to recognize and classify the image style of playlist covers according to the content of image features. We then employed hierarchical multiple linear regression and grouping regression to study the moderating effects of playlist features (linguistic style of titles and image style of covers) on the relationship between the number of playlist comments and the number of plays. The findings revealed that (1) the number of playlist comments (comment-count) has a stronger positive relationship with the number of plays (play-count) when the linguistic style of titles, such as concrete, perceptual, interactive, and social languages, appears more frequently and (2) the image style of playlist covers, such as natural, non-natural, painting and text, and portrait images, strengthens the relationships between playlist comment-count and linguistic style of playlist titles on play-count.

© 2022 The Author(s). Published by Elsevier España, S.L.U. on behalf of Journal of Innovation & Knowledge.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Introduction

The International Federation of the Phonographic Industry's Global Music Report 2021¹ revealed that the global recorded music market grew by 7.4% in 2020 (sixth consecutive year of growth) and the growth was driven by streaming, especially by paid subscription streaming revenues, which increased by 18.5%. With the vigorous development of music streaming services, the playlist, a new form of music classification, has become an entry point to help users browse the huge number of available songs (Dias et al., 2017). Whether discovering songs through platform-curated playlists and recommender system, or allowing users to create and share their own personal playlists, it is beneficial for music streaming platforms to enhance user stickiness and form a differentiation advantage. The vast majority of playlists in leading music streaming platforms such as Spotify, Apple Music, QQ Music, and NetEase Cloud Music are amateur playlists made by non-professional music enthusiasts (Bonnin &

Jannach, 2015). Every day, many User-Generated Playlists (UGPs) apply to provide for recommendations to users on the streaming platform. Therefore, it is a strenuous task for platforms to select an excellent UGP to enrich music content, attract users, and enhance the user experience.

Presently, user-generated content has become an important information source to understand users (Kwark et al., 2018). UGP covers a wide range of themes and scenes, and songs are carefully selected to not only fulfill the diversified needs of users but also to facilitate communication between users on the basis of an emotional connection to the same favorite songs. Moreover, the tide of informatization is sweeping across the human society. A structural change is taking place profoundly in consumer behavior and attitude as well as marketing communication. Users are increasingly inclined to obtain commodity information through images, audio, and video, and actively share consumption experience via Internet (Korfatis et al., 2012). Consumers are attracted to vivid and intuitive images, and therefore, combining images with text is more persuasive (Li et al., 2016; Yoo & Kim, 2014). When a song is "commented," "liked," or "shared" many times, the social interaction between users through music medium is established imperceptibly. Moreover, UGP has fancy covers and rich content titles, which greatly attract users' attention during playlist selection.

* Corresponding authors.

E-mail addresses: lizhi@tiangong.edu.cn (Z. Li), wangzhe@sctu.edu.cn (Z. Wang).

¹ https://www.ifpi.org/wp-content/uploads/2020/03/GMR2021_STATE_OF_THE_INDUSTRY.pdf

Most previous studies have focused on the creation of music recommender systems for automatic playlist generation and continuation. Such algorithmic playlists are primarily based on user preferences, interactions, song content, and artists. Moreover, some researchers have considered the users' complex contextual factors (Wang et al., 2020) and song sequence preferences (Liebman et al., 2019) to provide them with a more intelligent recommendation experience. However, from a comprehensive perspective, there has been little research on playlist selection that considers both the social and natural features (titles and covers) of playlists. Without considering the user response to playlist features, music recommendation methods will inevitably reduce the accuracy because of unsatisfactory user needs. This study aimed to investigate the features of playlist titles and covers that attract users and the consequences of playlist selection on music streaming platforms. More specifically, this study attempted to answer the questions at two different levels through an empirical study of music streaming platforms in China: (1) title-cover level: what combination characteristics of playlist titles and covers attract user attention to select the playlist and (2) music streaming platform level: how to select excellent playlists from thousands of UGP. Answering these questions is not only vital to understanding the psychological behavior of users' preferences for playlists but also conducive to implementing a more accurate playlist recommendation mechanism.

The rest of this paper is organized as follows. Section 2 describes the fundamental concept of UGP. Section 3 proposes the theoretical background, classification of the linguistic style of playlist titles, and image style of playlist covers. Section 4 presents the hypotheses. Section 5 elaborates the empirical studies conducted to test the relationship between the number of playlist comments and the number of plays under the influence of playlist titles and covers. Section 6 discusses the findings and results. The last section accounts conclusions and implications.

The concept of UGP

UGPs are created by non-professional music enthusiasts. Such playlists have different purposes, such as sharing music with others at a party, creating a quiet music environment for study and work, and enjoying music for self-expression. Generally, the UGP is published and promoted on a specific interface of a music-streaming platform, such as the daily playlist square of NetEase Cloud Music (see Fig. 1). As shown in Fig. 1, the playlist's title and cover take up much space for playlist information display, which can attract users' attention and inspire them to listen.

Playlist titles

The title is the soul of the playlist, reflecting the key piece of playlist content information. Most playlists on music platforms are created or submitted by users. The title content mainly expresses the creators' understanding, classification, and summary of playlist songs. Its text description does not pursue norms but is more casual. Intuitively, the playlist title created by users usually contains the concrete description of the playlist, such as musical genre, application scene of playlist (e.g., party, Christmas) (Yurekli et al., 2021). We observed that some playlist titles appeared in the form of questions or mentioned one of the recent social hot topics to attract users' attention. On the contrary, some playlist titles may not include any semantic information or their language expression may be more abstract (Faggioli et al., 2018). Psycholinguistic research shows that linguistic style is closely related to measures of people's social and psychological worlds (Tausczik & Pennebaker, 2010). In some sense, the linguistic style of playlist titles can reflect users' needs and preferences for songs. The language of the playlist title is concise and informative and can be expressed in multiple linguistic styles, such as concrete language, perceptual language, and interactive language. Moreover, title informativeness, which is measured by music aspects (i.e., the title highlights characteristics of playlist), has a positive

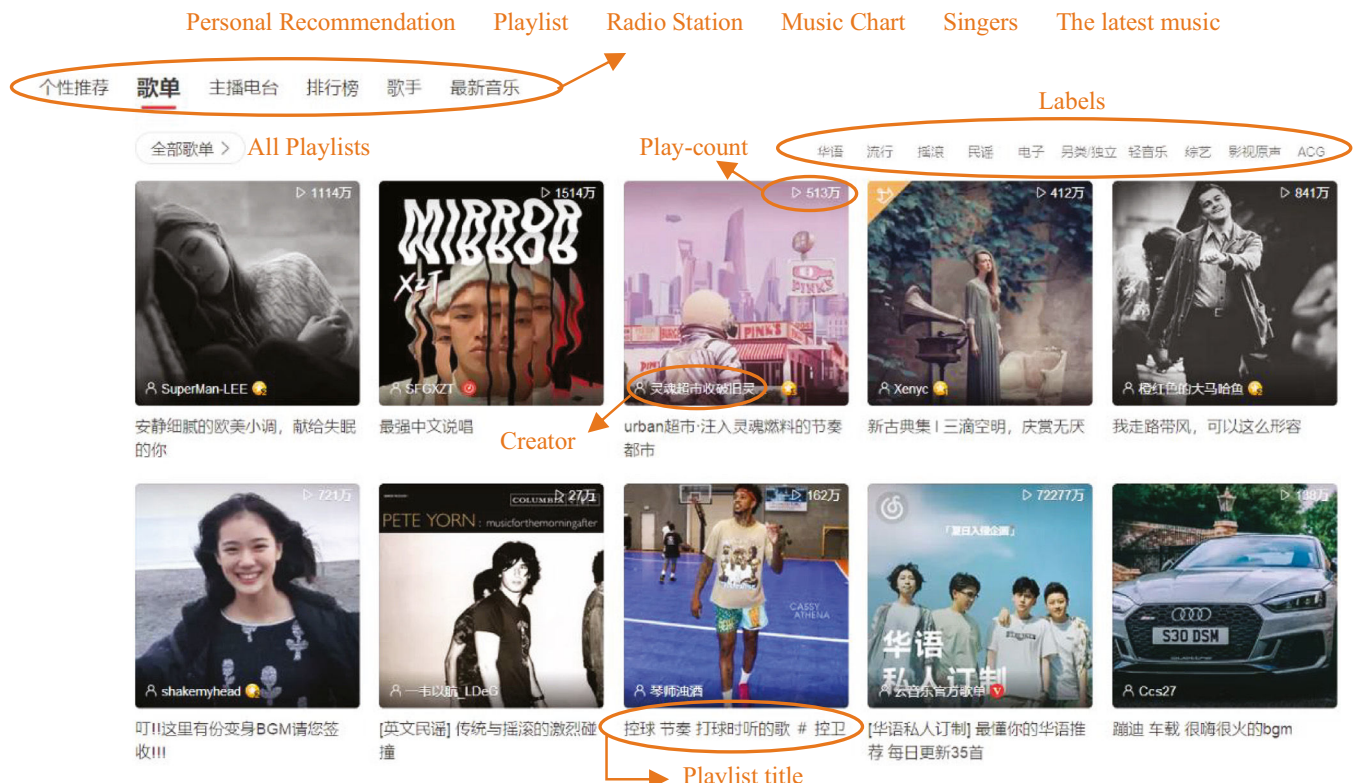


Fig. 1. The playlist square of NetEase Cloud Music.



Fig. 2. Playlist social in NetEase Cloud Music platform.

impact on title usefulness and may influence users' choice of playlist. This study explored the influence of linguistic style of playlist titles on users' playlist selection behaviors.

Playlist covers

According to dual coding theory (Paivio & Allan, 1991), users can gain a more comprehensive understanding of the playlist when providing both textual and visual information. Playlist cover is the explanation and visual extension of its title, helping users better understand the title and inspire their interest and curiosity in the playlist. Based on the visual principle, compared with abstract text, the playlist cover is more attractive to users owing to its vivid, intuitive, and visual characteristics, and therefore has obvious advantages in visual recognition (Lehtiniemi et al., 2012). Additionally, the playlist cover has a higher identification probability than the playlist title in terms of concept. A playlist cover can simply and truly convey sensory information about the playlist content, enhance the authenticity of users' perception, and provide complementary information for users' decision-making on playlist selection (Lehtiniemi et al., 2016). Some studies suggest that pictures are typically associated with a certain mood, which has yielded promising results in using "mood pictures" in playlist selection (Lehtiniemi, Ojala, & Väänänen, 2017) and playlist creation (Lehtiniemi & Ojala, 2013). Currently, the research on visual information of playlist covers focuses on music mood identification, which is usually measured by colors (Carles et al., 2009; Jukka et al., 2009). Among the existing studies, the recognition of playlist cover feature classification is relatively simple, while compared with color, image content contains rich information and is more prominent. Thus, feature dimension extraction based on image content can comprehensively express users' visual responses to playlist cover. This study extracted the image style features of the playlist cover by analyzing its image content of the playlist cover, such as identifying people, objects, and scenes from images. On this basis, the user playlist selection behavior was studied.

Playlist socialization

Music is a social phenomenon, and the idea of social media functionality has been floating around the streaming market for a long time, yet the music streaming experience in the West has become increasingly individualized and non-social, except for the comments section². In China, streaming platforms pay more attention to building social media, especially NetEase Cloud Music, which makes social features the core of its streaming service³ and drives explosive user growth. On music-streaming platforms, users are connected to each other through music relationships with similar concepts or interests. As an essential part of "playlist socialization," the commenting function plays a prominent role in stimulating communication effects. When users follow a playlist, in the comment section, they could share their thoughts about the music and find a group of people of the same taste. The social approach of generating topics and content through common interest points can attract more users to join the discussion and reply (see Fig. 2). Generally, the more comments a post already has, the more comments people will continue to post; this is referred as popularity bias. In turn, the playlist will also be discovered and played by more people on the music streaming platform. This two-way influencing mechanism of users' comments and playlist selection has gradually formed a new social mode of communication-playlist socialization. Moreover, studies on online comments suggest that the greater the number of comments on an item, the more likely it is that consumers will buy it (Siering et al., 2018). Accordingly, this study proposed that the number of comments on a playlist can indirectly have a positive impact on the number of times a song is played.

² <https://www.musicbusinessworldwide.com/tencent-music-uses-tipping-to-rack-up-revenues-why-arent-western-music-streaming-platforms-doing-the-same/>

³ <https://soundcharts.com/blog/chinese-recording-market-streaming>

Theoretical background

Classification of linguistic style of playlist titles

From a psychological perspective, linguistic style is closely linked to measures of people’s social and psychological worlds (Tausczik & Pennebaker, 2010). Thus, the linguistic style of playlist titles (hereafter, linguistic style) reflects how creators communicate with users. According to language expectancy theory (LET) (Burgoon & Miller, 1985), people develop expectations on appropriate communication styles, and the language features of individuals or groups can change the attitude of information recipients. LET indicates that linguistic style features, such as an influencer’s concreteness, preciseness, psychological closeness, and interactivity with viewers, can contribute to an increase in the number of users on social media platforms (Parhankangas & Renko, 2017). Some researchers believe that the linguistic styles are proposed to explore the persuasive effect of affective, sensory, and social processes in information, indicating that perceptual language and social language are crucial in making information appealing and easy to process (Lee et al., 2019). In the fast-paced information era, providing attractive and easy-to-understand information attracts more users. Moreover, as music socialization and interaction have gradually become a new norm, it has become the active exploration direction of playlist operations to engage users in content consumption and guide them to create valuable content. As we observed, questions, rhetorical questions, and interaction-related words appeared frequently in playlist titles. Based on four linguistic styles—concrete, precise, interactive, and low psychological distance languages—proposed by Burgoon and Miller (1985), and considering the characteristics of playlist titles and language expression in NetEase Cloud Music, we derive five linguistic styles of playlist titles: concrete, psychological intimacy, perceptual, interactive, and social languages.

Classification of image style of playlist covers

To extract the image style features of the playlist covers, we used image recognition artificial intelligence (AI) technology. It can not only accurately identify the visual content of images, including thousands of object labels and dozens of common scenes, but also perform scene classification, color recognition, style recognition, and element recognition. The methodology is described in detail in Section 5. In this section, we explain how AI technology allows the recognition of image style of playlist covers. AI technology expands the understanding of user behavior, allowing us to find hidden patterns in “big data”. By identifying associations in playlist selection behavior and the image style of playlist covers, contextualized and effective music streaming operation strategies can be developed.

Currently, many websites offer image recognition AI services, such as Baidu Intelligent Cloud, Google TensorFlow, and Facebook Photo Magic. Ordinarily, image recognition is based on deep learning technology, accompanied by image binarization, gray algorithm, and wavelet algorithm. This study adopted EasyDL image recognition technology to obtain the image style classification of playlist covers, which is a content-based image recognition AI technology provided by Baidu Intelligent Cloud. The intelligent algorithm application programming interface (API) of general object and scene recognition⁴ is used to identify the content of the playlist cover; based on this, we obtain four image styles of playlist covers: natural, non-natural, painting and text, and portrait images. Examples of each image style are listed in Table 1.

⁴ BAIDU AI CLOUD. Image Recognition. Recognition of General Objects and Scenarios <https://cloud.baidu.com/product/imagerecognition/general>

Table 1

The image styles and measure label.

Image Style	Examples
Natural Image	Natural scenery, natural phenomenon, animals, plants, etc.
Non-natural Image	Goods, buildings, transports, etc.
Painting and Text Image	Watercolor, oil painting, art painting, text image, etc.
Portrait Image	Public figures, character activities, cartoon characters, etc.

Hypotheses development

Information transmission and acceptance theory show that when product quality information is asymmetric, consumers may reduce the uncertainty of their decision through an online review search (Sussman & Siegal, 2003). Similarly, when users do not know whether the playlist is good or bad, they may read the comments to obtain information. The richer the comments, the more users will participate in playlist selection. Intuitively, the number of playlist comments (comment-count) is positively related to the number of plays (play-count), whereas the process of the effect of comment-count on play-count is complex. According to the SOR theoretical model of environmental psychology proposed by Mehrabian and Russell (1974), all aspects of the environment play a stimulating role (S), affecting the internal state of organisms (O) and thus driving their behavioral response (R). On the music streaming platform, users’ moods and behaviors in listening to playlists change in response to the stimulation of playlist features. Therefore, in our study, stimuli were playlist features that influenced user behavior of selecting playlists. This study considered the features of playlists (titles and covers) as the stimuli (S). It proposed that these features have a significant impact on user engagement with music-streaming platforms. Their impact on user engagement is regarded as the users’ listening state (O). The latter, in turn, affects the user and leads to a reaction (R), which is divided into two aspects: posting comments and playing music.

More specifically, stimuli (playlist titles and covers) have a moderating effect on behavioral responses (comment-count and play-count). To explore the potential moderators that strengthen the positive relationship between comment-count and play-count, this study first proposed a two-way interaction between comment-count and the linguistic style of playlist titles on play-count, according to persuasion theory (O’Keefe, 2002). Next, by combining the persuasion theory and dual coding theory of playlist covers, this study proposed a three-way interaction between comment-count, linguistic styles, and image styles on play-count. Fig. 3 shows the research model. For simplicity and clarity of exposition, we defined the positive effect of comment-count on play-count as the main effect, throughout this paper.

The effects of linguistic style

Persuasion theory suggests that information transmission generates a “persuasion effect” on audiences. Moreover, persuasion is a process of information transmission. That is, the publisher expresses and transmits information to audiences, trying to change their mental states and corresponding attitudes on a topic, person, or product of interest (Lee & Theokary, 2020). Previous studies show that linguistic style influences information persuasiveness, and further influences audiences’ subsequent attitudes and behaviors (Larrimore et al., 2011; Toma & D’Angelo, 2014). Specifically, words in a concrete language are tangible, special, or imaginable, and can shape user attitudes and behaviors (Packard & Berger, 2021).

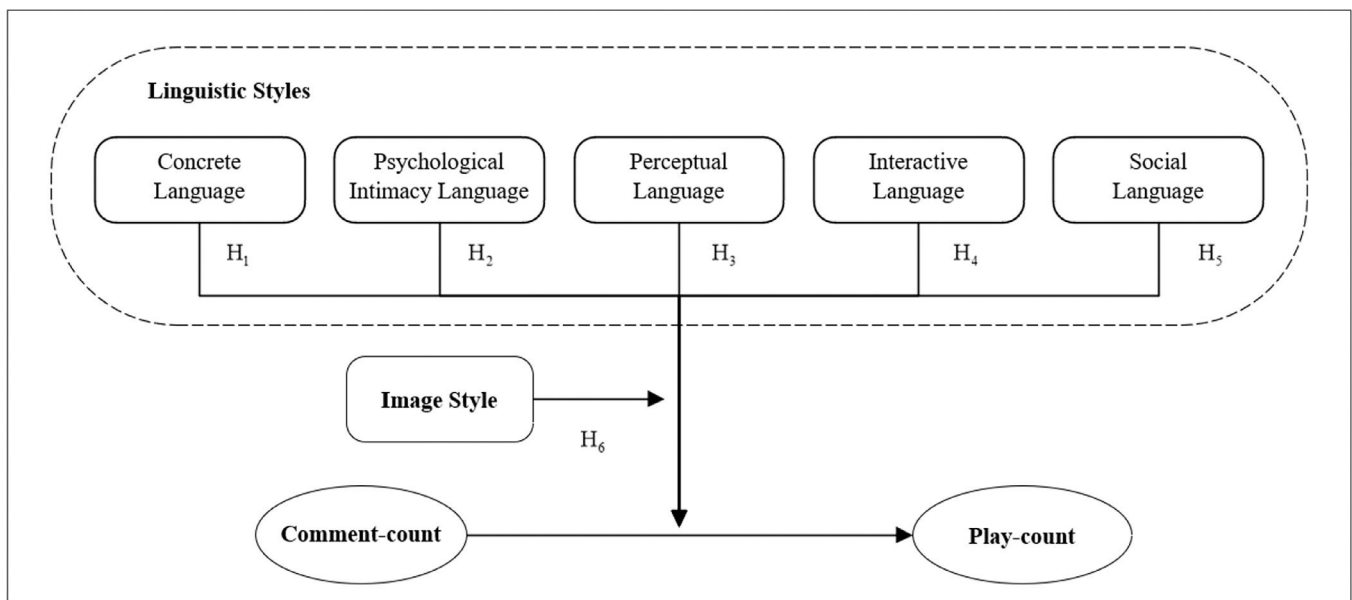


Fig. 3. Conceptual framework.

Concrete language reflects the details and contextualization of information, shaping social perceptions of information publishers' attention and understanding (Toma & D'Angelo, 2014). Concrete language is more specific and vivid, allows faster and deeper information processing for users, and has proven to be a significant factor in increasing viewers (Snefjella & Kuperman, 2015). Thus, the rich details and contextualization information provided by concrete languages make playlist titles more interesting and memorable (Pichl et al., 2016). In sum, we predict that using more concrete language in playlist titles will increase user satisfaction, intention, and behavior toward the playlist. The concrete language of playlist titles can enhance the positive effect of comment-count on play-count. Therefore, we proposed the hypothesis H1.

Hypothesis H1. The effect of playlist comment-count on play-count is stronger with a high level of concrete language.

Psychological distance refers to the degree of cognitive separation between one's perception of an object (events or people) and oneself and is a subjective judgment that has an impact on the credibility and level of interpretation of information (Josef & Zuzana, 2021). Psychological intimacy language has the characteristics of affinity and emotion, and can close the distance with audiences (Connors et al., 2020; Jaeger et al., 2018). Research related to consumer behavior suggests that psychological intimacy language produces informational stimuli that influence consumers' purchase decisions (Bratman et al., 2012). Thus, on the music streaming platform, the use of psychological intimacy language in the design of playlist titles can shorten the psychological distance with users, thus encouraging them to comment and play playlists (Snefjella & Kuperman, 2015). The inclusion of a psychological intimacy language in the playlist title helps users quickly relate to their situation and mood, and thus, to play the song list preferentially. For this reason, the psychological intimacy language of playlist titles can enhance the positive effect of comment-count on play-count. Therefore, we proposed the hypothesis H2.

Hypothesis H2. The effect of playlist comment-count on play-count is stronger with a high level of psychological intimacy language.

Words related to hearing, sight, and other senses can facilitate information comprehension (Federmeier, 2007). In other words, perceptual language is easier for users to process and understand, and statements containing such language are more attractive and persuasive (Kim et al., 2011). Sensory stimulation triggers the human

subconscious so that consumers obtain a figurative view of the abstract concept of the product. Thus, marketing activities that incorporate sensory perceptions will influence users' perceptions and judgment (Krishna, 2012; Krishna et al., 2016). Perceptually relevant cues also have a greater impact on users' perceived enjoyment of online comments; thus, comments containing perceptual language can help users achieve higher levels of enjoyment (Li et al., 2019). Playlist titles with perceptual language will stimulate users' senses, making them more receptive and likely to enjoy the playlist, thus affecting their playing behavior. In this regard, the perceptual language of playlist titles can enhance the positive effect of comment-count on play-count. Therefore, we proposed the hypothesis H3.

Hypothesis H3. The effect of playlist comment-count on play-count is stronger with a high level of perceptual language.

Using more questions in the headline will show a higher level of interaction with users. The study found that in the process of seeking help, the more interactive and interesting the language is, the more likely it is to be supported by others (Dickert et al., 2011). Recent psychological studies suggest that the increased use of questions in interactive language exhibits greater interactivity, which may help information publishers mitigate some of their communication challenges and better understand audiences' likes and needs (Parhankangas & Renko, 2017). When playlist titles ask questions or send invitations to users with interactive language, it arouses users' curiosity and captures their attention. Interaction with users may create a sense of trust and responsibility, which can have a significant impact on attracting users to choose a playlist (Hamby et al., 2015). The interactive language of playlist titles can enhance the positive effect of comment-count on play-count. Therefore, we proposed the hypothesis H4.

Hypothesis H4. The effect of playlist comment-count on play-count is stronger with a high level of interactive language.

The basic function of language is to communicate with the members of the society. Social language that expresses social concerns and social intercourse is conducive to the construction of group identity, which tends to develop positive intergroup attitudes (Yurekli et al., 2021). In recent years, the propaganda value of music streaming platforms has been highlighted, and the quantity and quality of the original sound track (OST) music production has continuously improved. OST is a type of movie soundtrack music that usually refers to the soundtrack (score) or songs in movies but also refers to

music in anime and games. Emerging music variety shows have also released many high-quality hit songs that are highly concerned with music users (Lin et al., 2013). The playlist comment area has become a social community for users to recommend movies and TV shows and support their favorite singers. The users express opinions and attitudes through comments, enhance the sense of participation, and achieve group psychological identity with others (Stsiampkouskaya et al., 2021). According to the playlist characteristics, we attribute the names of popular movies and TV shows, variety shows, and singers to the social language. The social language style of playlist titles reflects the communication and social functions of music platforms (Yoo et al., 2017). The social language of playlist titles can enhance the positive effect of comment-count on play-count. Therefore, we proposed the hypothesis H5.

Hypothesis H5. The effect of playlist comment-count on play-count is stronger with a high level of social language.

Differences in the effects of linguistic style by image style

According to **dual coding theory** in the field of cognitive psychology, an individual cognitive process involves two independent but interrelated systems: nonverbal and verbal systems (Delprato, 2009). The nonverbal system deals with visual images and emotional reactions, whereas the verbal system deals directly with language (Yoo & Kim, 2014). Based on the two coding systems, both nonverbal information, such as images, and verbal information, such as text, can be encoded. From the perspective of information spread, the fusion of image and text information enhances the depth of information, which, in turn, increases consumer confidence and facilitates the entire decision-making process (Mudambi & Schuff, 2010). Based on the above ponder, user-generated content combining images and text on streaming media platforms, has a higher perceived usefulness of information (Banerjee et al., 2018). There is reason to believe that the information conveyed by the playlist cover and title can affect users' comments and playlist selection behavior. Therefore, we proposed the hypothesis H6.

Hypothesis H6. Image style of playlist covers strengthens the relationships between playlist comment-count and linguistic style of playlist titles on play-count.

Considering different playlist image styles and different playlist title linguistic styles, we conducted in-depth discussions. Natural images mainly refer to images containing natural scenery, natural phenomena, animals, and plants. Evolutionary psychology research shows that since human beings originate from the natural environment, natural images make people feel familiar, safe, and comfortable, and it is easy to generate attraction (Bratman et al., 2012). Psychological intimacy language can promote this attraction effect by shortening people's psychological distance. Moreover, the cognitive processing of natural images is faster than that of images depicting artificial content (Kuzinas et al., 2016). Natural images can arouse users' emotional responses more quickly. When there is more psychologically intimate language in playlist titles, users will emotionally identify with them and have the desire to visit and comment on playlists. Therefore, we proposed the hypothesis H6a.

H6a: For the natural image, psychological intimacy language is more strongly associated with the relationship between playlist comment-count and play-count.

Non-natural images contain concrete objects such as buildings, vehicles, or toys, which are intuitive and informative. As discussed, concrete language refers to a specific description of an object. It provides adequate details and context-specific information to simplify the information processing. For non-natural images, people need more directed attention to identify and analyze the information in

the image; thus, a more concrete language in playlist titles can help users filter the "junk" information in the image and better understand the theme and content of the playlist (Kuzinas et al., 2016). According to cue consistency theory, when images are consistent with text descriptions, people's cognition is better promoted, resulting in subsequent intentions and behaviors (Miyazaki et al., 2005). Therefore, we proposed the hypothesis H6b.

H6b: For the non-natural image, concrete language is more strongly associated with the relationship between playlist comment-count and play-count.

Text images contain simple words or sentences with a prominent color background, but limited information. By contrast, painting images are mostly complex visual images with implicit information (Stahl & Kaihovirta, 2019). It is difficult for users to quickly process the information displayed in complex images; therefore, playlist titles are required to provide more clues to improve users' understanding of the playlist. Perceptual language is emotionally infectious, which compensates for the shortcoming of barely intelligible information on painting and text images. Understanding the perceptual language in playlist titles can lead to higher perceived usefulness or helpfulness of playlist covers (Li et al., 2016). Therefore, we proposed the hypothesis H6c.

H6c: For painting and text image, perceptual language is more strongly associated with the relationship between playlist comment-count and play-count.

Facial portraiture is one of the most powerful resources for visual images because individuals can study faces and their expressions from infancy (Harrison & Claire, 2003). For this reason, when people see a portrait image, they have a sense of interaction and close intimacy. Portrait images usually show the facial expressions of **characters**, which will be correspondingly communicated to users, triggering their attention and empathy (Bernat et al., 2010). Portrait images can mobilize users' psychological states, increase their support and identification of playlists, and encourage them to respond more positively to interactive language. In addition, we noticed that portraits of celebrities, who are representative of their respective fields and could bring users more intimacy (Argyris et al., 2020), thus attracting fans to participate in the playlist with their portraits on the covers. Therefore, we proposed the hypothesis H6d.

H6d: For the portrait image, interactive language is more strongly associated with the relationship between playlist comment-count and play-count.

Methodology

Data collection and research design

The data for this study were collected from the NetEase Cloud Music Platform in China. Established in 2013, NetEase Cloud Music is the first music platform with "playlist" as the core structure, and has the largest and best quality "playlist" library in China. Currently, it is one of the most popular music platforms, with more than 180.5 million monthly active users. The nucleus of NetEase Cloud Music in terms of product and operation is to discover and share music experiences, build a professional and social music platform, attract users, and retain early adopters. A total of 99484 playlists were gathered from August 2020 to December 2020, including playlist details (title, comments, covers, created date, etc.) and the basic information of the playlist creator (user gender, user level, user label, number of fans etc.). The number of playlists updated daily is approximately 1290, among which some playlists will be repeated and recommended in the playlist square. We deleted the playlists that were duplicate, those that we failed to identify from covers, or those that had some missing or abnormal data. We also eliminated playlists that were later cancelled by creators. After data cleaning, a sample of 7606 playlists was obtained.

Table 2
The linguistic styles and measure label.

Linguistic Styles	Language components
Concrete language	Proposition, quantifier, tag words for playlist
Psychological intimacy language	First-person pronouns (I and we), sentiment words (positive and negative emotions)
Perceptual language	Perceptual words (percept, see, hear, and feel)
Interactive language	Interactive word
Social language	Social and internet buzzwords

Recognition and analysis of playlist title features

To understand how the linguistic style of playlist titles affects users' choice of playlists, each title was analyzed using TextMind, a Chinese language psychological analysis system designed by Computational Cyber-Psychology Lab, Institute of Psychology, and Chinese Academy of Sciences, to extract the language components of linguistic style. Inspired by the dictionaries of linguistic inquiry and word count ([LIWC], LIWC2007 and C-LIWC), TextMind was developed based on the characteristics of Simplified Chinese, providing an integrated solution from automatic Chinese words segmentation to psychoanalysis⁵. There are 102 linguistic styles in TextMind, including psychological categories (e.g., emotion, social, affect), basic linguistic categories (e.g., pronouns, prepositions, conjunctions), cognitive categories (e.g., work, religion, hear), and so on, which cover most types of words needed to reflect linguistic styles. For research on the English language, LIWC is the most commonly used language analysis tool for finding and counting words in psychology-relevant categories across multiple texts (Tausczik & Pennebaker, 2010). Similar to LIWC, TextMind uses a word frequency count to quantify Chinese text information, and its output is expressed as a percentage of the total words in the text content. Based on previous research in sociolinguistics, linguistic psychological analysis, narrative and discourse analysis, and LET, the language components of each linguistic style were defined and summarized in Table 2, which is described in detail below.

Concrete language refers to a linguistic style that provides a detailed and context-specific description of an object, making information processing easier and more comprehensive (Parhankangas & Renko, 2017). Extant studies mostly used three linguistic cues about concrete language: articles (e.g., a, an, the), prepositions (e.g., to, with) and quantifiers (e.g., many, few). Unlike English, there are no articles in Chinese; therefore, articles were not considered in this study. Concreteness indicates the degree to which a word can refer to a tangible and actual entity, describing objects and behaviors in a way that looks more specific, imaginable, and easier to perceive (Packard & Berger, 2021). A playlist is a music collection with multiple attributes, and it also has several foundational attributes, such as genre, language, and theme, as tags for sorting and searching. These tags are basic and essential information for music streaming users to match their listening habits, and we believe that they are well suited to represent concrete language in playlist titles. Table 3 shows examples of each tag and their numbers in all title samples. In summary, in this study, concrete language is calculated as the ratio of prepositions, quantifiers, and tag words for playlists within the total title words.

Psychological intimacy language can be construed as a linguistic style that reduces the psychological distance to the audience. Psychological distance reflects subjective cognition of how far an object or event is perceived to be from the self (Alter & Oppenheimer, 2010). The use of personal pronouns is significant for manifesting the

Table 3
The examples and numbers of the playlist tag.

Tag	Word	Number
Language	European, American, Japanese, Korean and Cantonese	315
Genre	Pop, rock, folk, electronic, dance music...	756
Theme	Variety show, campus, games, KTV, ACG...	361

Table 4
The examples of the interactive word dictionary.

Word	Number
吗 (A modal auxiliary word in Chinese, usually placed at the end of a sentence to express the question.)	107
一起 (Together)	105
遇见 (Meet)	35
是否 (Whether)	29
给你 (...For you/ Give you...)	29
为什么 (Why)	35
有没有 (Did you ...?/ Have you ever ...?)	10

distance of intimacy. Greater use of first-person pronouns (e.g. "I," "me," and "we") would close the psychological distance between communicators (Lee & Theokary, 2020). There is also a strong association between psychological distance and positive or negative evaluations (Josef & Zuzana, 2021). We suggest that psychological intimacy may also be related to positive or negative emotions. In conclusion, psychological intimacy language was measured as the ratio of first-person pronouns to positive and negative emotion words within the total title words. A high value of this construct is linked to a more psychologically intimate language.

Perceptual language refers to a linguistic style described by sensory experience. In the era of sensory marketing, visual, haptic, and olfactory cues have the potential to influence consumer choices and evaluations by stimulating their perceptions, emotions, and memories (Ghosh & Sarkar, 2016). The dictionaries of perception, vision, auditory, and feeling in TextMind were considered as perceptual words in this study to measure perceptual language and it was calculated as the ratio of perceptual words to the total title words.

A high level of interactivity is reflected in the use of questions and feedback requests in the text (Ki & Kim, 2019). Considering the playlist title is brief, usually a sentence or just a few words, we elaborated an interactive word dictionary, which is composed of phrases and words that reflect interactive behaviors and situations extracted from the collected dataset to identify interactive language; there are 198 interactive words in total. Moreover, the typical interrogative mood word "吗 (Do you)" is an important symbol for identifying questions other than the question mark in the Chinese, so it was added to the interactive word dictionary. Examples of interactive word dictionaries are presented in

Table 4. The NLPir-ICTCLAS Chinese lexical analysis system⁶ was then used for playlist title text segmentation based on the interactive word dictionary, and the percentage of words in the interactive word dictionary was calculated using Python. Python is an interpretative, interactive, and object-oriented computer language with a large number of network-oriented function libraries that make data processing and analysis more efficient (Li et al., 2021). Interactive language was measured as the ratio of words in the interactive word dictionary to the question mark within the total title words.

Prior studies posit that social language is closely related to social concerns (e.g., poverty or employment) and social intercourse-related concerns (e.g., neighbors or children) (Lee et al., 2019). On the music streaming platform, users will be interested in social hot topics (e.g., stay up later, youth) as well as the contents about pop singers,

⁵ Computational Cyber-Psychology Lab, Institute of Psychology, Chinese Academy of Sciences. TextMind <http://ccpl.psych.ac.cn/textmind/>

⁶ NLPir-ICTCLAS Chinese lexical analysis system <http://ictclas.nlpir.org/>

Table 5
The examples of the playlist hot word dictionary in social language.

Word	Number
网抑 (Online Depression expressions)	25
网愈 (Online Emotional-healing)	8
高考 (National College Entrance Examination)	16
七夕 (Chinese Valentine's Day)	17
赵英俊 (Yingjun Zhao, a well-known Chinese singer)	23
如此可爱的我们 (Lovely Us, a Chinese drama)	26
乐队的夏天 (The Big Band, a Chinese music variety show)	13

popular TV shows, and TV Series. In particular, the popularity of music entertainment programs in China has dramatically increased users' listening behavior on music streaming platforms, which has also driven playlist updates. Thus, we obtained network buzzwords and the names of pop singers, popular TV shows, and TV series that appeared frequently in playlist titles and the Internet. We integrated them into a playlist hot-word dictionary, and there were 161 hot words. Examples of the hot-word dictionary are listed in Table 5. Social language was measured as the ratio of words from the hot word dictionary to social words from the dictionary of TextMind within the total title words.

Recognition and analysis of playlist cover features

Several studies have demonstrated that visual attention and emotions can be affected by the content division of images (Kuzinas et al., 2016). To recognize the content of the playlist cover and obtain relevant labels for the cover image, we invoked the API services of general object and scenario recognition in the Baidu AI cloud. The API can identify over 100 thousand kinds of objects and scenarios, and return the names and details of the items in the image. As shown in Fig. 4, the output of this image recognition technology lists all the results of the uploaded image and its confidence and is displayed on the upper right side of the image. Then, the cover images were classified by the results with the highest confidence and divided into four categories: natural, non-natural, portrait, paintings, and text images. We generated three dummy variables to account for the category to which the cover image belonged, with portrait images as the reference group.

Results and discussion

Analysis strategy

First, zero-mean normalization was applied to three variables: comment-count, play-count, and fan-count to reduce the multicollinearity problem. SPSS 23 was used to perform the descriptive statistics and correlation analysis of each variable. Next, hierarchical regression analysis was conducted to test the hypotheses using Python. Specifically, control and independent variables were entered in Model 1, and the main effect was tested (the effect of comment-count on play-count). In Model 2, the five independent variables related to the linguistic style and three dummy variables that represent the image style with portrait images as the reference group, were added. In Model 3, the interaction terms with the respective comment-count were added to test whether the linguistic style of playlist title or the image style of playlist cover had an impact on the main effect. In Model 4, the moderating effect of image style was tested by analyzing the three-way interaction terms of comment-count, linguistic style, and image style. Additionally, the two-way and three-way interaction effects were graphically reconfirmed. We calculated the variance inflation factor (VIF) of the variables for each model to rule out multicollinearity.

Correlation and collinearity analyses

The descriptive statistics and correlations of all measures studied are shown in Table 6. We found that research variables such as the number of comments and the number of plays and control variables (user-level, fan-count, interval, playlist-track-count, and title-words-count) were significantly related to each other ($p < 0.01$). The correlation coefficients between the independent and control variables were mostly low, from which multicollinearity was preliminarily excluded. Furthermore, we examined the value of the variance inflation factor (VIF); the majority of variables were below the recommended threshold of 10. Only two interaction terms of the moderators had a slightly higher VIF ($VIF_{CC+IL} = 10.092$, $VIF_{CC+SL} = 10.373$) in Model 4 of Table 7, indicating that collinearity concerns are not serious.



Fig. 4. The recognition analysis of playlist covers.

Table 6
Correlations, means, standard deviations minimum and maximum.

	(1) User-level	(2) Fan-count	(3) Interval	(4) Playlist-track-count	(5) Title-words-count	(6) Comment-count	(7) Concrete language	(8) Psychological intimacy language	(9) Perceptual language	(10) Interactive language	(11) Social language	(12) Image category	(13) Play-count	Mean	Std. Dev.	Min	Max
13														0	1	-0.16	36.34
12														2.56	0.79	1	4
11														0.07	0.12	0	1
10														0.04	0.09	0	1
9														0.11	0.18	0	1
8														0.06	0.10	0	1
7														0.06	0.09	0	1
6														0	1	-0.20	45.04
5														7.56	3.30	1	22
4														96.46	290.68	1	10000
3														57.87	124.59	0	2395
2														0	1	-0.15	36.40
1														7.19	2.09	0	10

Notes: N = 7606; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed test)

Hypothesis testing on the effect of different linguistic styles

Table 7 reports the results of the hierarchical regression analysis. As shown in Model 3, the interaction terms of concrete language and comment-count ($\beta = 3.4328, p < 0.01$), perceptual language and comment-count ($\beta = 1.0617, p < 0.01$), interactive language and comment-count ($\beta = 0.731, p < 0.01$), and social language and comment-count ($\beta = -1.394, p < 0.01$) were statistically significantly related to play-count. However, the interaction term of psychological intimacy language and comment-count was not significant, which indicated that psychological intimacy language has no significant moderating effect on the relationship between playlist comment-count and play-count, based on a 95% confidence interval. Thus, hypothesis H2 is not supported. This finding suggests that the connection between the psychological intimacy language and comments may not translate into the sustained playing of a playlist. Different from the influence of psychological intimacy language in civic crowdfunding campaign and advertising marketing, it may be difficult to show the advantages in a short text like the playlist title.

To demonstrate the two-way interaction of the above four linguistic styles with a significant moderating effect, we plotted the results of their simple slope analysis for each of them (see Fig. 5). When concrete language, perceptual language, and interactive language are more common in playlist titles, the positive relationship between playlist comment-count and play-count is stronger. Thus, Hypotheses H1, H3, and H4 are supported. For music users, picking a playlist mostly depends on concrete and perceptual information, especially the tags provided by platforms in a concrete language, which has been widely accepted by the general public as the primary message to match the expected playlist. Moreover, interactive language may attract users' attention toward playlists by stimulating their desire for expression. Users may pursue self-identity and group identity through comments, which promotes rapid growth in play-clicks.

Contrary to the positive prediction in H5, we found that there was a stronger positive correlation between playlist comment-count and play-count for titles with fewer social languages. Thus, H5 is not supported. This shows that the social language of the title has a reverse effect on the positive correlation between playlist comment-count and play-count. This finding concurs with the study that examined the role of linguistic style in civic crowdfunding (Lee et al., 2019), in which the authors indicated that extensive use of social language reduces the likelihood of funding performance. Several possible interpretations of this finding are as follows: first, the information updates rapidly in the Internet media age, and it would be difficult for users to keep up with the hot topics of social language and focus on it for a long time; and second, users listen to music mainly for entertainment and relaxation, and they want to escape from the language that focuses on social processes and practice.

Hypothesis testing on the difference by image style of playlist cover

Model 4 in Table 7 shows that most three-way interaction terms of image styles, linguistic styles, and comment-count had a statistically significant effect on play-count, which means that there was a significant quadratic interaction between image style and linguistic style. To further analyze the three-way moderating effect of the styles of playlist cover and linguistics on the relationship between playlist comment-count and play-count (the main effect), we built a grouping regression analysis to control the effect of different image styles, and the results are presented in Fig. 6.

As shown in Fig. 6(a), for playlists with natural images as their cover, concrete language, psychological intimacy language, and perceptual language have positive moderating effects on the main effect, while social language has a negative moderating effect and interactive language has no significant effect. In particular, psychological intimacy language is more strongly associated with the main effect.

Table 7
Summary of regression analysis results.

	Model 1	Model 2	Model 3	Model 4
Variables		Dependent variable = Play-count		
Intercept	-0.1005***	-0.1220***	-0.0876**	-0.1026***
Controls				
User-level	0.0046	0.0054	0.0037	0.0054
Fan-count	0.0449***	0.0440***	0.0411***	0.0419***
Interval	0.0008***	0.0008***	0.0007***	0.0007***
Playlist-track-count	-0.0001***	-0.0001***	-0.0001***	-0.0001***
Title-words-count	0.0038	0.0013	0.0007	0.0003
Independent				
Comment-count (CC)	0.6027***	0.6040***	0.4046***	0.4226***
Linguistic Styles				
Concrete language (CL)		0.2726***	0.1252	-0.0048
Psychological intimacy language (PIL)		0.0777	-0.0177	0.0827
Perceptual language (PL)		0.1726***	0.2286***	0.2441***
Interactive language (IL)		0.1655*	0.1311	0.2006*
Social language (SL)		-0.2074***	-0.2096**	-0.2735***
Moderator				
Image category 1 (IC-1)		0.0167	-0.0353	-0.0432
Image category 2 (IC-2)		0.0211	0.0009	-0.0406*
Image category 3 (IC-3)		-0.0486	-0.0461	-0.0342
Two-way interaction terms				
CC*CL			3.4328***	0.5519***
CC*PIL			-0.1562	0.7103***
CC*PL			1.0617***	1.4507***
CC*IL			0.7310***	1.8321***
CC*SL			-1.3940***	-1.8443***
CC*IC-1			-0.1138***	-0.2804***
CC*IC-2			0.2934***	-0.1296***
CC*IC-3			-0.1806***	-0.1648***
IC-1*CL			-0.0210	0.2552
IC-1*PIL			0.5494*	0.3945
IC-1*PL			0.1420	0.1255
IC-1*IL			0.0434	-0.1024
IC-1*SL			-0.0425	0.0757
IC-2*CL			0.6944***	1.3372***
IC-2*PIL			0.0824	0.2478
IC-2*PL			-0.1097	-0.1159
IC-2*IL			-0.0707	-0.2579
IC-2*SL			-0.0387	-0.0043
IC-3*CL			0.0528	-0.0629
IC-3*PIL			0.0870	0.0697
IC-3*PL			-0.1302	-0.1149
IC-3*IL			0.3172	0.0389
IC-3*SL			0.3037	0.2409
Three-way interaction terms				
CC*IC-1*CL				1.3575***
CC*IC-1*PIL				2.6123***
CC*IC-1*PL				1.0412***
CC*IC-1*IL				-2.0339***
CC*IC-1*SL				-0.8960**
CC*IC-2*CL				8.8195***
CC*IC-2*PIL				1.4392***
CC*IC-2*PL				-1.3117***
CC*IC-2*IL				-2.3622***
CC*IC-2*SL				-0.1548
CC*IC-3*CL				-0.6595
CC*IC-3*PIL				-1.2789***
CC*IC-3*PL				-0.7074**
CC*IC-3*IL				-1.1661***
CC*IC-3*SL				1.7780***
R-squared	0.409	0.412	0.538	0.603
Adj. R-squared	0.409	0.411	0.535	0.600
F	877.459***	379.611***	237.704***	220.518***
VIF	1.01~1.072	1.012~1.146	1.019~6.5	1.021~10.373

Notes: N = 7606; * p<.1, ** p<.05, ***p<.01, two-tailed test; Image category 1 (IC-1) = Natural images; Image category 2 (IC-2) = Non-natural images; Image category 3 (IC-3) = Paintings and text pictures; The baseline for the image category is portrait images.

Thus, H6a is supported. As shown in Fig. 6(b), for playlists with non-natural image as the cover, except that perceptual language has no significant effect on the main effect, the other four linguistic styles have positive moderating effect on the relationship between playlist comment-count and play-count. Clearly, concrete language has stronger positive moderating effect. Thus, H6b is supported. As

shown in Fig. 6(c), for playlists with painting and word image as their cover, only psychological intimacy and perceptual languages have a significant positive moderating effect on the main effect, and the effect of psychological intimacy language is stronger. Thus, H6c is supported. As shown in Fig. 6(d), for playlists with portrait image as the cover, all five linguistic styles have positive moderating effects on

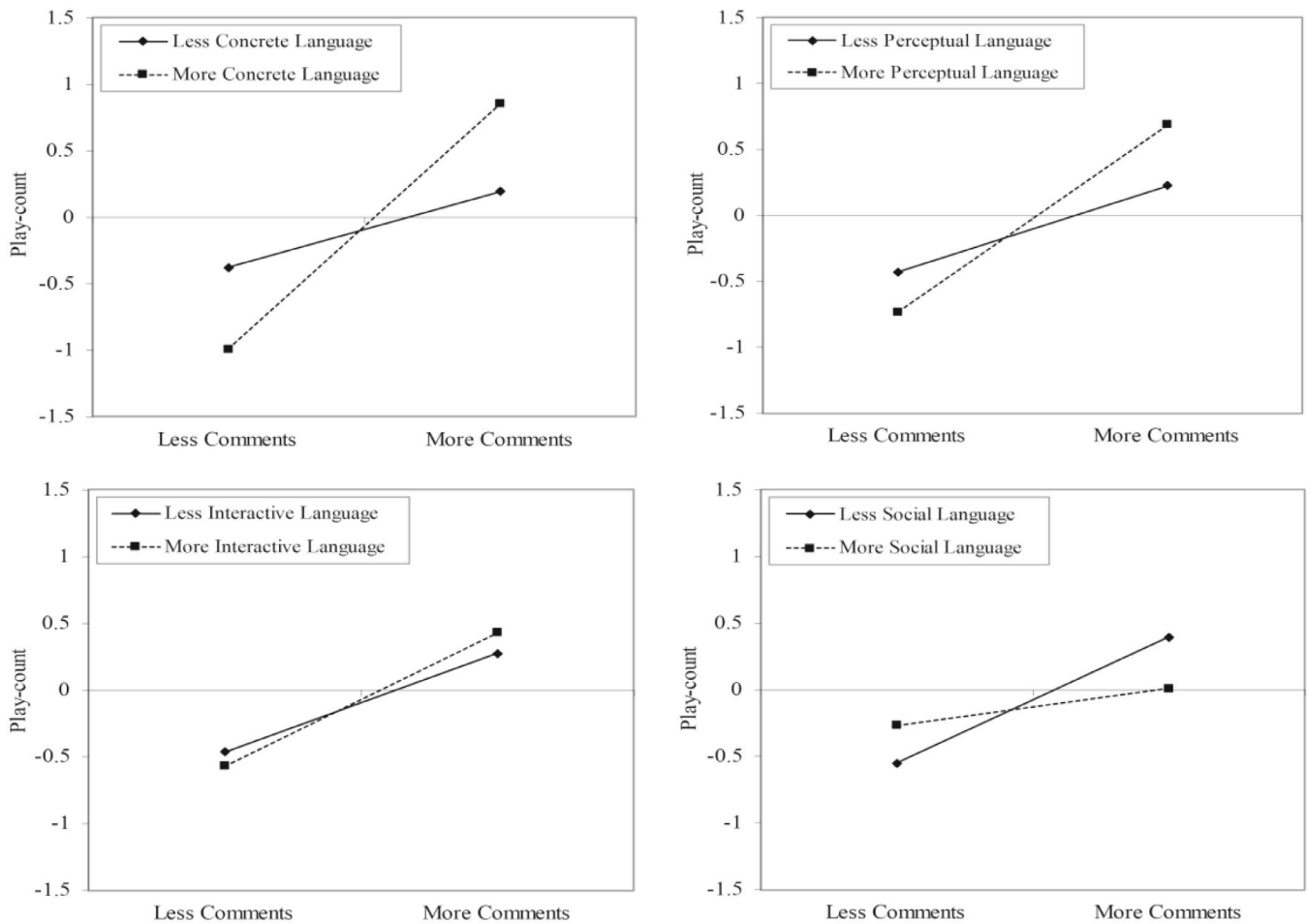


Fig. 5. Moderating effect of linguistic style on the relationship between comment-count and play-count.

the main effect, and interactive language is the greatest. Thus, H6d is supported.

All the above results suggest that the influence of title linguistic style may vary according to different image styles of playlist cover. In particular, the results of the grouping regression analysis suggest that psychological intimacy language has a significant moderating effect for each image style, while it is not significant when image style is not controlled. Thus, H6 is supported. People's emotions can be influenced by the content of images, which can compensate for the lack of emotional expression in words. For different image styles, strengthening the use of corresponding linguistic styles with stronger effects is more conducive to stimulating the users' enthusiasm for music socialization and their long-term love and repeated playing of playlists.

Conclusion and implications

This study applies persuasion theory and dual coding theory to measure the influence of verbal features (title) and visual features (cover) of playlists on users' comments and playlist selection behavior. We derived five linguistic styles and four image styles of playlist titles and covers and investigated their interaction effect on the relationships between the number of playlist comments and plays. The findings indicated that the number of playlist comments had a stronger positive relationship with the number of plays when concrete, perceptual, interactive, and social languages appeared more frequently in playlist titles. Furthermore, after the three-way interaction model analysis, we found that the effect of title linguistic style may

vary according to different image styles of playlist cover. The image style of the playlist strengthens the relationship between playlist comment-count and the linguistic style of playlist titles on play-count.

Theoretical implications

The dual coding theory suggests that the simultaneous presentation of text and images has a positive effect on cognition. The present study reveals that users' interest in a playlist and its final choice are mainly influenced by the visual cues of cover images and linguistic cues of titles. Based on LET, we developed new scales to measure the linguistic style of playlist titles, namely concrete, psychological intimacy, perceptual, interactive, and social languages. Especially, considering the characteristics of playlist titles and language expression in NetEase Cloud Music, we expanded and refined the linguistic elements of interactive, social, and concrete linguistic languages. AI image recognition technology was used to classify playlist cover images into four categories: natural, non-natural, portrait, paintings, and text images. Further, combined with linguistic style, this study analyzed the image and text information of music streaming platforms and verified the applicability of double coding theory in streaming media platforms.

Managerial implications

This study emphasizes the significance of capturing users' visual attention, which should be regarded as the primary goal of marketing

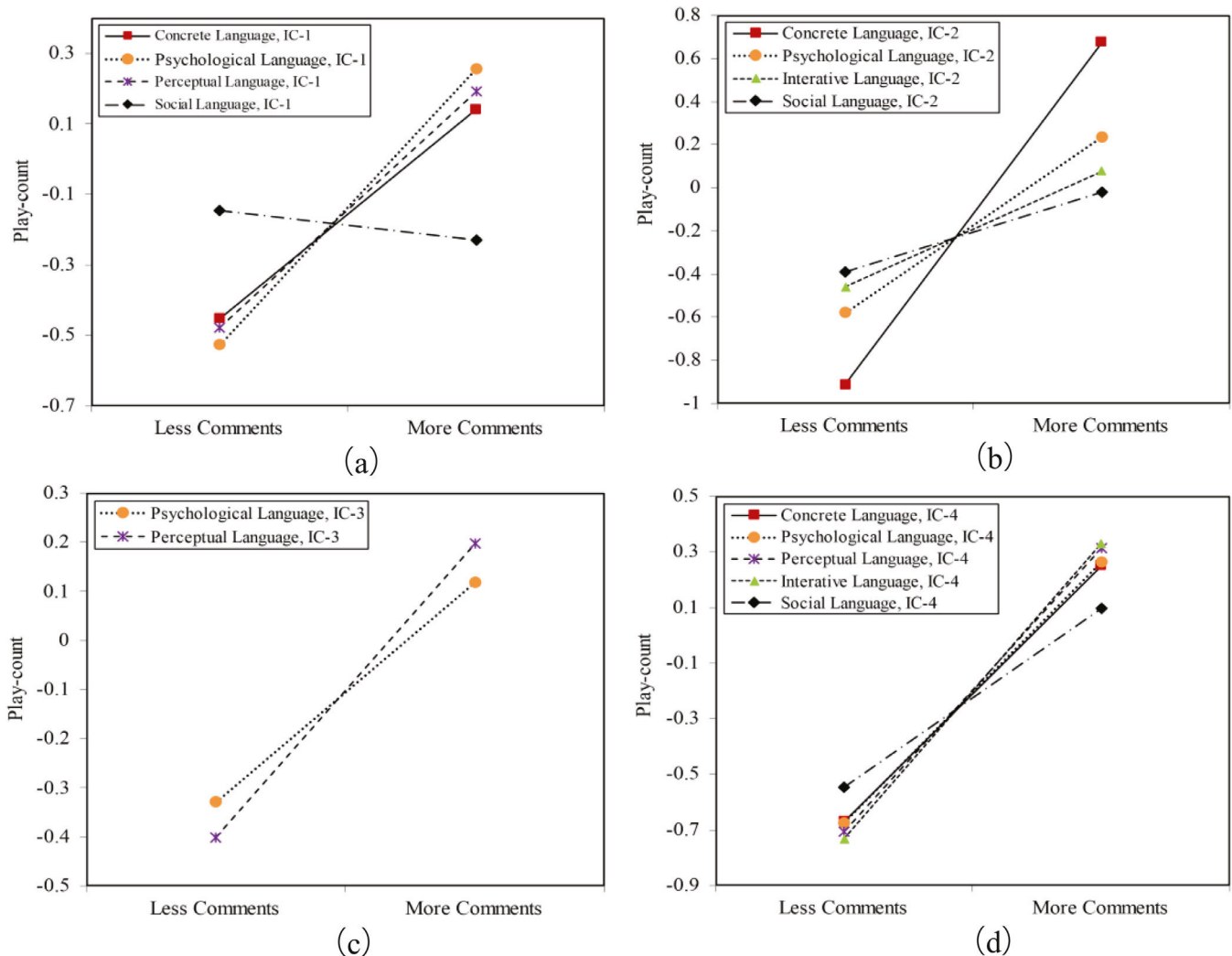


Fig. 6. Interaction diagrams of the three-way interaction.

communication in today's information overload marketplace. The implications are not only applicable to operators of music platforms, such as NetEase Cloud Music, but also to marketing managers of other streaming media platforms (e.g., video platforms). First, it is obvious that titles and user-generated pictures affect users' perceptions of the attributes of a playlist. These linguistic and visual cues enable users to easily compare the attractiveness of different playlists and facilitate playlist-selection decisions. Second, surprisingly, the results yield additional insights into the persuasion process when interpreted in light of a specific "playlist socialization" setting. Therefore, a recommendation for music platform marketing managers would be to encourage users co-create the music image, inviting them to share their life photos on "playlist socialization" community. Third, our findings provide suggestions for developing effective visual materials that capture users' visual attention and elicit positive responses to playlists. The results provide a filtering criterion for music stream platform operators to recommend front-page playlists and inspire music creators to focus on matching playlist titles with covers when creating playlists.

Limitations and future research

This study has several limitations and scope for future research. While there are benefits from using text-image analysis to measure the interaction effect of linguistic style, image style, and comments

on users' playlist selection behavior, the analysis has limitations related to exploring the internal psychological mechanism of user behavior. Future research could collect users' behavioral data through [questionnaire survey](#) and experiments to provide an analysis of psychological behavior that expands the results of this study. Moreover, the study focused on NetEase Cloud Music, a prominent music platform in China, whose [language system](#) is Chinese and user group is mainly Chinese. Since cultural differences may affect users' information processing ([Dwivedi et al., 2020](#); [Mariani & Predvoditeleva, 2019](#)), future research could expand to English music platforms to investigate how culturally different users process visual and verbal cues and further influence their playlist selection behavior. Finally, future studies can examine other streaming platforms such as YouTube and Oasis APP. While the current research has generated insights into "playlist socialization," we look forward to additional research exploring different social media platforms, such as Instagram, Snapchat, and TikTok, to fully understand how users react in various platform environments.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this study.

Funding

This work was sponsored by the National Nature Science Foundation of China NSFC under Grant no. (72002153, 41971249). The authors gratefully acknowledge the financial support and greatly appreciate all the reviewers and editors who contributed to this study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jik.2022.100212.

References

- Alter, A. L., & Oppenheimer, D. M. (2010). Effects of fluency on psychological distance and mental construal (or why New York is a large city, but New York is a civilized jungle). *Psychological Science*, 19(2), 161–167.
- Argyris, Y. A., Wang, Z., Kim, Y., & Yin, Z. (2020). The effects of visual congruence on increasing consumers' brand engagement: An empirical investigation of influencer marketing on Instagram using deep-learning algorithms for automatic image classification. *Computers in Human Behavior*, 112(1), 106443.
- Banerjee, S., Dellaroas, C. N., & Zervas, G. (2018). Interacting user-generated content technologies: How questions and answers affect consumer reviews. *Ssrn Electronic Journal*, 58(4), 742–761.
- Bernat, E., Patrick, C. J., Benning, S. D., & Tellegen, A. (2010). Effects of picture content and intensity on affective physiological response. *Psychophysiology*, 43(1), 93–103.
- Bonnin, G., & Jannach, D. (2015). Automated generation of music playlists: Survey and experiments. *Acm Computing Surveys*, 47(2), 1–35.
- Bratman, G. N., Hamilton, J. P., & Daily, G. C. (2012). The impacts of nature experience on human cognitive function and mental health. *Annals of the New York Academy of Sciences*, 1249(1), 118–136.
- Burgoon, M., & Miller, G. R. (1985). An expectancy interpretation of language and persuasion. In Howard Giles, & Robert N. St. Clair (Eds.), *In Language, Communication, and Social Psychology* (p. pp. 31). London: Routledge.
- Carles, F., Julià, & Sergi, J. (2009). SongExplorer: A tabletop application for exploring large collections of songs. the 10th International Society for Music Information Retrieval Conference (pp. 675–680).
- Connors, S., Khamitov, M., Thomson, M., & Perkins, A. (2020). They're just not that into you: How to leverage existing consumer-brand relationships through social psychological distance. *Journal of Marketing*, 85(5), 92–108.
- Delprato, D. J. (2009). Mind and its evolution: A dual coding theoretical approach. *Psychological Record*, 59(2), 295–300.
- Dias, R., Goncalves, D., & Fonseca, M. J. (2017). From manual to assisted playlist creation: A survey. *Multimedia Tools and Applications*, 76(12), 14375–14403.
- Dickert, S., Sagara, N., & Slovic, P. (2011). Affective motivations to help others: A two-stage model of donation decisions. *Journal of Behavioral Decision Making*, 24(4), 361–376.
- Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., & Wang, Y. (2020). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, 59(1), 1–37.
- Faggioni, G., Polato, M., Aiolfi, F., & Acm (2018). Efficient similarity based methods for the playlist continuation task. the Recsys Challenge'18: Proceedings of the Acm Recommender Systems Challenge (pp. 1–6).
- Federmeier, Kara D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491–505.
- Ghosh, T., & Sarkar, A. (2016). To feel a place of heaven: Examining the role of sensory reference cues and capacity for imagination in destination marketing. *Journal of Travel & Tourism Marketing*, 33, 25–37.
- Hamby, A., Daniloski, K., & Rinberg, D. B. (2015). How consumer reviews persuade through narratives. *Journal of Business Research*, 68(6), 1242–1250.
- Harrison, & Claire (2003). *Visual social semiotics: Understanding how still images make meaning*: 50 (pp. 46–61). Technical Communication.
- Jaeger, S. R., Spinelli, S., Ares, G., & Monteleone, E. (2018). Linking product-elicited emotional associations and sensory perceptions through a circumplex model based on valence and arousal: Five consumer studies. *Food Research International*, 109(7), 626–640.
- Josef, K., & Zuzana, R. (2021). Psychological distance as a means of evaluation. *New Ideas in Psychology*, 63(1), 1–9.
- Jukka, H., Antti, A., & Harri, S. (2009). Associating colours with musical genres. *Journal of New Music Research*, 38(1), 87–100.
- Ki, C-W., & Kim, Y-K. (2019). The mechanism by which social media influencers persuade consumers: The role of consumers' desire to mimic. *Psychology & Marketing*, 36(10), 905–922.
- Kim, K., Bae, J., Nho, M. W., & Chang, H. L. (2011). How do experts and novices differ? Relation versus attitude and thinking versus feeling in language use. *Psychology of Aesthetics Creativity & the Arts*, 5(4), 379–388.
- Korfiatis, N., García-Barriocanal, E., & Sánchez-Alonso, S. (2012). Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications*, 11(3), 205–217.
- Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology*, 22(3), 332–351.
- Krishna, A., Cian, L., & Sokolov, T. (2016). The power of sensory marketing in advertising. *Current Opinion in Psychology*, 10(1), 142–147.
- Kuzinas, A., Noiret, N., Bianchi, R., & Laurent, E. (2016). The effects of image hue and semantic content on viewer's emotional self-reports, pupil size, eye movements, and skin conductance response. *Psychology of Aesthetics Creativity & the Arts*, 10(3), 360–371.
- Kwak, Y., Chen, J. Q., & Raghunathan, S. (2018). User-generated content and competing firms' product design. *Management Science*, 64(10), 4608–4628.
- Larimore, L., Jiang, L., Larimore, J., Markowitz, D., & Gorski, S. (2011). Peer to peer lending: The relationship between language features, trustworthiness, and persuasion success. *Journal of Applied Communication Research*, 39(1), 19–37.
- Lee, C. H., Bian, Y. Y., Karaouzene, R., & Suleiman, N. (2019). Examining the role of narratives in civic crowdfunding: Linguistic style and message substance. *Industrial Management & Data Systems*, 119(7), 1492–1514.
- Lee, M. T., & Theokary, C. (2020). The superstar social media influencer: Exploiting linguistic style and emotional contagion over content? *Journal of Business Research*, 132, 860–871.
- Lehtiniemi, A., & Ojala, J. (2013). Evaluating moodPic - A concept for collaborative mood music playlist creation. the 17th Information Conference on Information Visualisation (pp. 86–95).
- Lehtiniemi, A., Ojala, J., & Väänänen, K. (2012). Using animated mood pictures in music recommendation. the 16th International Conference on Information Visualisation (pp. 143–150).
- Lehtiniemi, A., Ojala, J., & Toukoma, H. (2016). Design and evaluation of mood pictures in social music discovery service. *Personal and Ubiquitous Computing*, 20(1), 97–119.
- Lehtiniemi, A., Ojala, J., & Väänänen, K. (2017). Socially augmented music discovery with collaborative Playlists and mood pictures. *Interacting with Computers*, 29(3), 416–437.
- Li, H. Y., Wang, R., Meng, F., & Zhang, Z. L. (2019). Making restaurant reviews useful and/or enjoyable? The impacts of temporal, explanatory, and sensory cues. *International Journal of Hospitality Management*, 83(1), 257–265.
- Li, Q., Huang, Z., & Christianson, K. (2016). Visual attention toward tourism photographs with text: An eye-tracking study. *Tourism Management*, 54(7), 243–258.
- Li, Y., Liu, Z., Zhang, Y., Jiang, L., Cai, Y., Chen, X., et al. (2021). Investigation and probabilistic health risk assessment of trace elements in good sale lip cosmetics crawled by python from Chinese e-commerce market. *Journal of Hazardous Materials*, 405, 124279.
- Lieberman, E., Saar-Tsechansky, M., & Stone, P. (2019). The right music at the right time: Adaptive personalized playlists based on sequence modeling. *Mis Quarterly*, 43(3), 765–786.
- Lin, K. S., Lee, A., Yang, Y. H., Lee, C. T., & Chen, H. H. (2013). Automatic highlights extraction for drama video using music emotion and human face features. *Neurocomputing*, 119, 111–117.
- Mariani, M., & Predvoditeleva, M. (2019). How do online reviewers' cultural traits and perceived experience influence hotel online ratings? An empirical analysis of the Muscovite hotel sector. *International Journal of Contemporary Hospitality Management*, 31(12), 4543–4573.
- Miyazaki, A. D., Grewal, D., & Goodstein, R. C. (2005). The effect of multiple extrinsic cues on quality perceptions: A matter of consistency. *Journal of Consumer Research*, 32(1), 146–153.
- Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on amazon.com. *Mis Quarterly*, 34(1), 185–200.
- O'Keefe, D. J. (2002). *Persuasion theory and research* (3rd ed.). Thousand Oaks, CA: SAGE.
- Packard, G., & Berger, J. (2021). How concrete language shapes customer satisfaction. *Journal of Consumer Research*, 47(5), 787–806.
- Paivio, & Allan (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45(3), 255–287.
- Parhankangas, A., & Renko, M. (2017). Linguistic style and crowdfunding success among social and commercial entrepreneurs. *Journal of Business Venturing*, 32(2), 215–236.
- Pichl, M., Zangerle, E., & Specht, G. (2016). Understanding playlist creation on music streaming platforms. the Proceedings of 2016 IEEE International Symposium on Multimedia (pp. 475–480).
- Siering, M., Muntermann, J., & Rajagopalan, B. (2018). Explaining and predicting online review helpfulness: The role of content and reviewer-related signals. *Decision Support Systems*, 108(4), 1–12.
- Sneffella, B., & Kuperman, V. (2015). Concreteness and Psychological Distance in Natural Language Use. *Psychological Science*, 26(9), 1449–1460.
- Stahl, M., & Kaihoviirta, H. (2019). Exploring visual communication and competencies through interaction with images in social media. *Learning, Culture and Social Interaction*, 21, 250–266.
- Stsiampkouskaya, K., Joinson, A., Piwek, L., & Ahlbom, C. P. (2021). Emotional responses to likes and comments regulate posting frequency and content change behaviour on social media: An experimental study and mediation model. *Computers in Human Behavior*, 124(1), 106940.
- Sussman, S. W., & Siegal, W. S. (2003). Informational Influence in organizations: An integrated approach to knowledge adoption. *Information Systems Research*, 14(1), 47–65.

- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54.
- Toma, C. L., & D'Angelo, J. D. (2014). Tell-tale words: Linguistic cues used to infer the expertise of online medical advice. *Journal of Language & Social Psychology*, 34(1), 25–45.
- Wang, R. R., Ma, X., Jiang, C., Ye, Y., & Zhang, Y. (2020). Heterogeneous information network-based music recommendation system in mobile networks. *Computer Communications*, 150, 429–437.
- Yoo, C. W., Goo, J., Huang, C. D., Nam, K., & Woo, M. (2017). Improving travel decision support satisfaction with smart tourism technologies: A framework of tourist elaboration likelihood and self-efficacy. *Technological Forecasting & Social Change*, 123(1), 330–341.
- Yoo, J., & Kim, M. (2014). The effects of online product presentation on consumer responses: A mental imagery perspective. *Journal of Business Research*, 67(11), 2464–2472.
- Yurekli, A., Bilge, A., & Kaleli, C. (2021). Exploring playlist titles for cold-start music recommendation: An effectiveness analysis. *Journal of Ambient Intelligence and Humanized Computing*, 12(11), 10125–10144.