"Listen to Music, Listen to Yourself": Design of a Conversational Agent to Support Self-Awareness While Listening to Music

Wanling Cai*
Department of Computer Science
Hong Kong Baptist University
Hong Kong, China
cswlcai@comp.hkbu.edu.hk

Yucheng Jin*
Department of Computer Science
Hong Kong Baptist University
Hong Kong, China
yuchengjin@hkbu.edu.hk

Xianglin Zhao
Department of Computer Science
Hong Kong Baptist University
Hong Kong, China
csxlzhao@comp.hkbu.edu.hk

Li Chen
Department of Computer Science
Hong Kong Baptist University
Hong Kong, China
lichen@comp.hkbu.edu.hk

ABSTRACT

Music can affect the human brain and cognition. Melodies and lyrics that resonate with us can awaken our inner feelings and thoughts; being in touch with these feelings and expressing them allow us to understand ourselves better and increase our self-awareness. To support self-awareness elicited by music, we designed a novel conversational agent (CA) that guides users to become self-aware and express their thoughts when they listen to music. Moreover, we investigated two prominent design factors in the CA, *proactive guidance* and *social information*. We then conducted a 2x2 between-subjects experiment (N = 90) to investigate how the two design factors affect self-awareness, user acceptance, and mental well-being. The results of a five-day user study reveal that high *proactive guidance* and *social information* increased self-awareness, but high *proactive guidance* tended to influence perceived autonomy and usefulness negatively. Further, users' subjective feedback revealed the CA's potential to support mental well-being.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in interaction design**; **User studies**; • **Applied computing** → **Health informatics**.

KEYWORDS

Listening to music, chatbot, user engagement, self-determination theory, emotional well-being

ACM Reference Format:
Wanling Cai, Yucheng Jin, Xianglin Zhao, and Li Chen. 2023. "Listen to Music, Listen to Yourself": Design of a Conversational Agent to Support Self-Awareness While Listening to Music. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3544548.3581427

*Both authors contributed equally to this research.

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3581427

1 INTRODUCTION

Music is ubiquitous in everyday life [8, 102]. People often listen to music as a way to self-regulate their emotions (e.g., reduce stress) and achieve self-awareness (e.g., focus on themselves to understand their feelings) [91, 99]. Research in psychology and neuroscience has shown that music affects our brain and cognition [39, 108], and enhances our positive experience and well-being [91, 107]. For example, when people listen to music, particularly their preferred music, the interconnectivity in their brains can associate music and self-awareness (i.e., an inner state when one becomes the focus of attention [25]), and invoke their personal stories and emotional memories (also referred to as resonance with music). Being self-aware allows us to recognize and understand our feelings and thoughts, which is associated with good mental well-being [13, 62].

Mental well-being has drawn increasing attention from human-computer interaction (HCI) researchers. In recent decades, they have proposed a variety of technologies to enhance mental well-being, such as conversational agents [51, 54], tangible interfaces [23], and virtual reality [6]. Given the benefits of listening to music and its relation to self-awareness, we are interested in the design of technology to support self-awareness while listening to music, with the goal to promote mental well-being.

Among various technologies, a conversational agent (CA) or chatbot, in particular, is designed to mimic human-like conversations in the form of either text or voice. The two-way communication that the CA enables can increase the rapport between users and agents, which has also been shown to enhance user engagement and affect user experience in various contexts positively [15, 55, 69]. In the well-being field, CA has been used increasingly to help track daily information (e.g., mood and activities) [47, 59] and provide guidance to motivate healthy lifestyle behavior change (e.g., diet and journaling) [51, 54]. In the music context, CA has been adopted to recommend personalized songs and help people explore music genres [15, 42]. However, few studies thus far have investigated CA's potential to increase people's self-awareness while listening

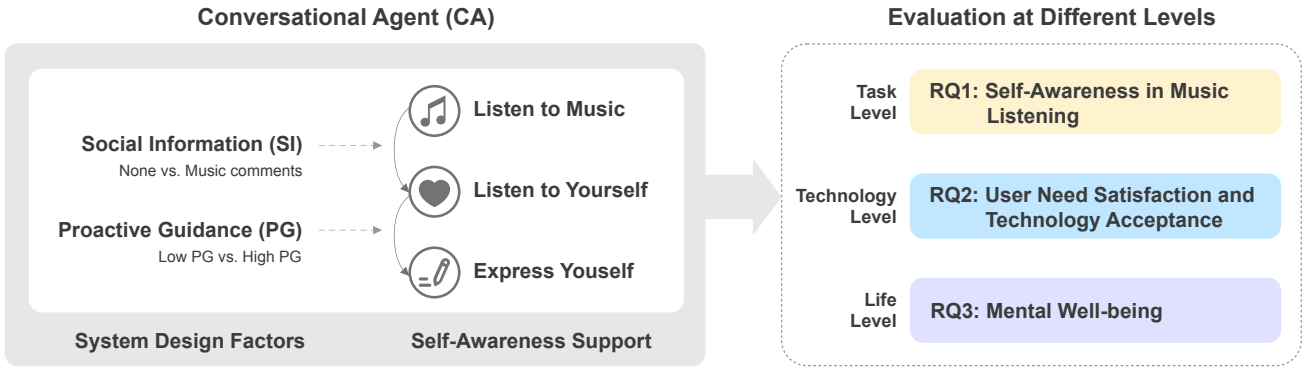


Figure 1: Our research questions.

to music. In light of the merits of CA above, we believe that by integrating conversation interaction into music listening, CA can be suited well to act as a virtual companion that can deliver guidance to support people's self-awareness while listening to music.

Therefore, in this study, we designed a CA to foster people's self-awareness while listening to music by guiding them to listen to themselves (i.e., feel the resonance with music) and express their personal feelings and thoughts that the music might evoke. Previous works on CA have suggested that a CA's proactivity level, e.g., proactive or reactive manner of offering assistance, can influence user experience and interaction behavior in the decision-making process [15, 75]. Therefore, we considered the level of *proactive guidance (PG)* in our CA design and examined whether high *PG* in the listening process (i.e., CA proactively nudges users to sense their resonance with music and express themselves) or low *PG* (i.e., CA only guides users to express their thoughts when they indicate explicitly that the music resonates with them) influences users' self-awareness when listening to music. In addition, previous studies in social music systems have revealed that *social information (SI)*, such as music comments that other listeners in the music community posted, contains other listeners' emotional stories and personal thoughts, which may cultivate emotional resonance as well and help people think about themselves [14, 41, 56]. As a CA's information delivery has been shown to affect users' perceptions and behaviors [53, 55, 64], for example, a CA's self-disclosure can affect users' perceived intimacy positively and encourage deeper self-disclosure [55], we were also interested to know whether the presentation of *SI* (e.g., music comments) in CA can enhance users' self-awareness while listening to music.

In particular, drawing on a model of well-being-supportive design proposed by Peters et al. [78], which demonstrates that technology design may influence well-being within various spheres of experience, we investigated the effect of our CA design at three specific levels: *task*, *technology*, and *life*. At the *task* level, we examined the way the two system design factors (i.e, *PG* and *SI*) in the CA influenced users' self-awareness while listening to music. At the *technology* level, we attempted to understand users' acceptance (i.e., intention to use [80]) of the CA, which can be useful to further improve its design to support self-awareness while listening to music. Previous research has indicated that users are more willing to engage with technology and accept it when their interaction with the system satisfies their psychological needs [77, 78], particularly the three basic needs identified in Self-Determination Theory (SDT) [24]: autonomy, competence, and relatedness. For example, psychological need satisfaction can mediate the effect of system design on users' engagement and acceptance [9, 109]. Thus, in our study, we investigated the way the two system design factors influenced users' need satisfaction, and thereby, affected their acceptance of the CA. Finally, at the *life* level, we evaluated our CA with respect to its benefits for mental well-being, with the goal to determine whether users can reap the benefits of interacting with the CA that supports self-awareness while listening to music.

In summary, this work investigated the following three research questions (see Figure 1):

RQ1: How do the two system design factors (*PG* and *SI*) in our CA affect users' self-awareness during music listening?

RQ2: What factors can influence users' need satisfaction and their acceptance of the CA?

RQ3: What benefits can such a CA offer users to enhance their mental well-being?

To answer our research questions, we conducted a five-day user study (N=90) with a 2 [*proactive guidance (PG)*: low vs. high] X 2 [*social information (SI)*: none vs. music comments] between-subjects design to evaluate our CA. The evaluation of the CA focused on users' self-awareness (at the *task* level), psychological need satisfaction as well as user acceptance (at the *technology* level), and mental well-being (at the *life* level). Our quantitative analyses revealed three major findings: (1) At the *task* level, the level of *PG* in the CA influenced the users' self-awareness, and the presentation of *SI* encouraged a deeper level of user sharing; (2) at the *technology* level, the CA's *PG* influenced users' satisfaction of the need for autonomy and perceived usefulness negatively, and (3) at the *life* level, our CA improved users' emotions slightly and showed a potential benefit to mental well-being.

In short, the contributions of our work are four-fold:

(1) We designed a novel conversational agent (CA) that supports people's self-awareness when they listen to music. To our knowledge, this is the first work that attempts to support self-awareness by integrating conversation interaction into music listening.

(2) We investigated the effects of two prominent design factors (i.e., *proactive guidance* and *social information*) in such a CA with empirical evidence.

(3) We conducted a five-day longitudinal study to evaluate our CA and revealed its influence on well-being at three levels: *task*, *technology*, and *life*.

(4) We provided practical implications for designing a CA that can engage people better while they listen to music, so as to foster their self-awareness and mental well-being.

2 RELATED WORK

In this section, we review the literature that motivated our study, including the work on the psychological benefits of listening to music and self-awareness, technologies for mental well-being as well as the design of conversational agents.

2.1 Psychological Benefits of Listening to music

Music can be seen as a ubiquitous companion in our everyday lives, and listening to music is one of the most popular activities in our leisure time [72, 99, 107]. Previous literature that has investigated the functions of music has revealed three principal dimensions [91]: (1) *self-awareness*, which indicates that music elicits self-related thoughts and helps people think about who they are and how they feel; (2) *social relatedness*, which suggests that people use music to express their own ideas and thoughts to others, to understand the way others think, and to feel closer to their friends; and (3) *arousal and mood regulation*, which shows that music is used as a form of entertainment or to improve mood. Moreover, according to psychology and neuroscience research [39, 108], music can affect our brain and cognition; when people listened to their preferred music, they found themselves having unsolicited self-related thoughts. They also showed increased connectivity in the frontal part of the brain that is involved in cognitive processes, such as conscious thinking and self-reflection, which suggests a relation between music and self-awareness, together with associated personal and emotional memories [39].

2.2 Self-Awareness in Psychology Research

In psychology research, self-awareness can be defined as an inner state in which one becomes the object of one's own attention [25]. It is commonly distinguished into two types: public self-awareness and private self-awareness [29, 92, 100]. Public self-awareness refers to focusing on how other people see us, which often emerges in social situations when we are in front of audiences, such as when we are giving a presentation. In comparison, private self-awareness represents how we see our beliefs, thoughts, and feelings, which usually occurs when we are exposed to self-focusing and reflecting stimuli such as mirrors and photos. Another type of stimuli that can create self-focus is written materials (e.g., books and articles) and various media sources (e.g., movies and music), which convey different perspectives and feelings that may likely elicit self-awareness and lead to self-evaluation [62]. Research has demonstrated many positive effects of self-awareness on mental health and well-being [13, 62, 84]; for example, when we are more self-aware, we can understand our emotions better and reflect on the way they affect our behavior, which helps us regulate our emotions positively, and hence improves our mental health and life satisfaction in several aspects, e.g., job, decision-making, and relationships [84]. To foster self-awareness, different activities (such as yoga, meditation, and self-reflection journals) have been suggested to practice being completely in the present [92, 100]. Given that music can be used suitably as a self-focus stimulus to lead us to self-awareness, as discussed above, we considered supporting and strengthening self-awareness in the process of listening to music to enhance mental well-being.

2.3 Technologies for Mental Well-being

In recent decades, a growing body of literature on HCI and health has designed various technologies (e.g., self-tracking technologies, conversational agents, tangible interfaces) to promote mental health and well-being [21, 89, 96]. For instance, several self-tracking tools have been developed to self-monitor people's physical activities [32, 83], stress [19], mood [50, 67], diet [83], etc., which can help increase self-awareness and motivate health-related behavior change [46, 81, 85]. Conversational agents (CAs) are gaining increasing attention in HCI research. They are useful in helping people monitor their feelings and behaviors [47, 59], encourage deep self-disclosure in journaling practice [53–55, 79], and improve mental health and well-being [27, 54, 93]. Previous studies have suggested that CAs can be a virtual companion, which can not only help track people's behaviors and listen to their sharing without judgment but also proactively offer users' guidance (such as cognitive-behavioral therapy [27, 93]) to improve mental well-being in conversational interactions [27, 93]. Moreover, human-like communication in CAs has been shown to establish rapport between users and agents and foster user engagement and trust[43, 54]. Given these advantages of CAs, our study attempted to design a CA that can provide guidance to support and strengthen people's self-awareness while listening to music.

2.4 Design and Evaluation of Conversational Agents

In the literature, conversational agents (CAs) are considered largely to be one type of application that mimics human-human communication to help users achieve specific tasks, such as finding recommendations [15, 43], expressive writing [74], and peer support chats [71]. In designing a CA, the design of conversation interaction, such as the level of CA proactivity and the information delivery, can affect how users interact with and perceive a CA [15, 71, 75]. For example, in the recommendation context, a proactive CA that guides users to follow the suggestions may offer users new insights and knowledge [15, 71, 73], while a reactive CA that allows users more control may be perceived to be more enjoyable [71, 75]. Moreover, the information delivery of a CA may also affect users' experience [53, 55, 64]. A recent study found that a CA's social sharing, such as its self-disclosure, can improve users' perceived intimacy and encourage them to engage in a deeper level of self-disclosure [55]. Motivated by these observations, we considered two system design factors, i.e., *proactive guidance (PG)* and *social information (SI)*, in our conversational design of a CA and investigated whether and how they affected user interaction and perception.

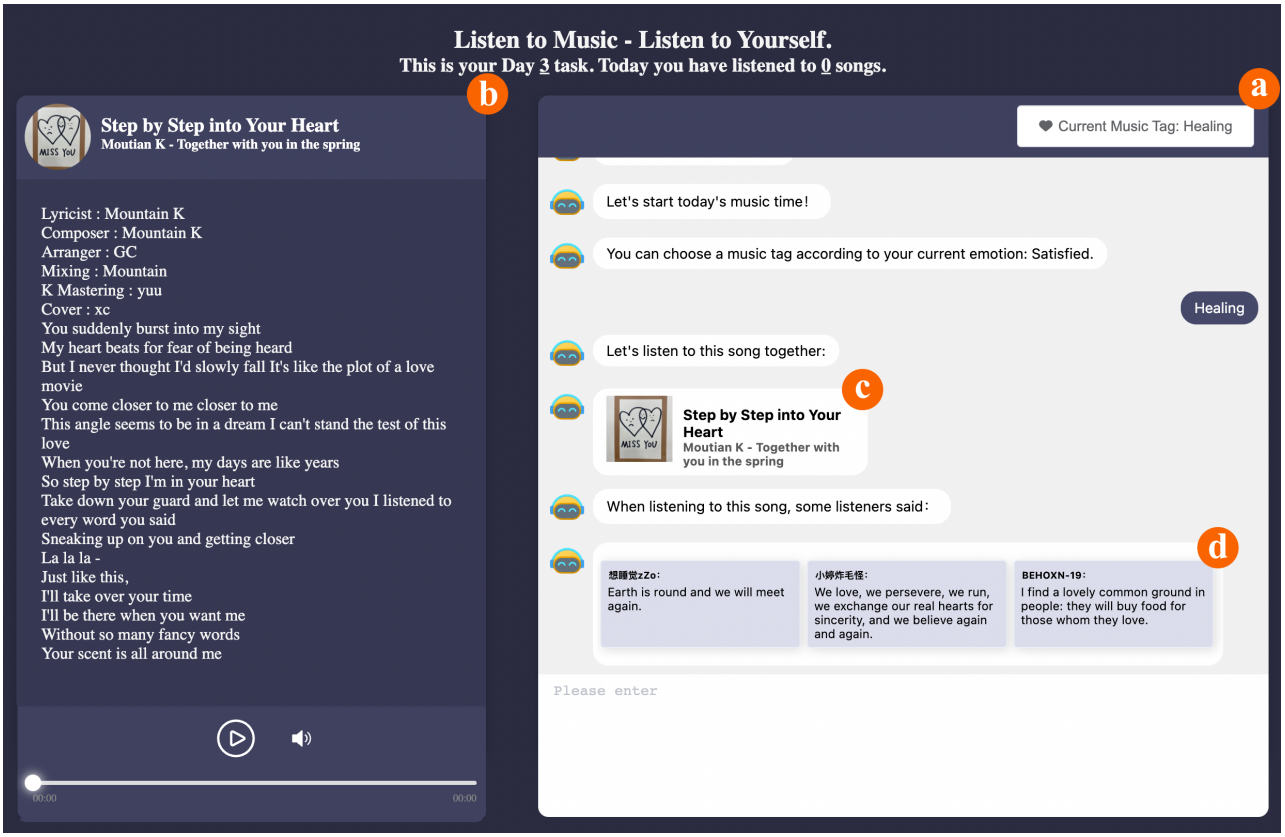


Figure 2: The user interface of our conversational agent includes (a) a dialog window, (b) a music player, (c) music recommendation, and (d) music comments. Note that the content presented in the screenshot was translated from Chinese to English for easy understanding.

Moreover, in the HCI literature, a model for well-being-supportive design, i.e., Motivation, Engagement, and Thriving in User Experience (METUX) proposed by Peters et al. [78], suggests that the design of technology may influence human well-being within different spheres of experience (e.g., task, interface, life, society), which motivated our evaluation of the CA at three different levels: *task* (i.e., how our CA design affected users' self-awareness while listening to music), *technology* (i.e., how our CA design affected users' need satisfaction and user acceptance), and *life* (i.e., how such a CA influenced mental well-being). At the technology level, in particular, it has been shown that positive experiences can be enhanced when the technology can support people's psychological needs [16, 77, 78]. According to self-determination theory (SDT), a theory of human motivation [24] that has been applied in many domains [68, 77], such as health [4] and games [109], all of us have three fundamental psychological needs: autonomy, competence, and relatedness, which can motivate us to seek growth and improve our well-being [78]. With respect to these three needs, autonomy refers to the need to feel able to make choices or behave in accordance with one's interests and values, competence refers to the need to feel confident and effective in a task in which one is engaged, and relatedness refers to the desire to feel connected and have a sense of belonging with others. Prior studies have shown certain relations among these, e.g., higher perceived autonomy is likely to lead to higher perceived competence. Studies have also indicated that users are more willing to engage with a technology and show a greater acceptance when their interaction with the system satisfies their psychological needs [77, 78]. For instance, in studies of mobile applications [9, 109], game elements in the interaction have been shown to improve basic psychological need satisfaction, and hence lead to greater perceived user engagement and acceptance. Given that users' perceived psychological need satisfaction can act as a mediator that affects their acceptance [16], in our study, we evaluated our CA to understand how our conversation design affected users' need satisfaction and hence their acceptance of a CA that supports self-awareness while listening to music.

3 CONVERSATIONAL AGENT DESIGN

To foster people's self-awareness while listening to music, we designed a CA that may guide users to listen to themselves (i.e., sense the resonance with music) and express their personal feelings and thoughts the music has elicited. Specifically, inspired by the music psychology literature [37], the CA is designed to recommend music according to the users' current emotions and music preferences, and to guide them to experience themselves and sense their resonance with music by displaying music comments that

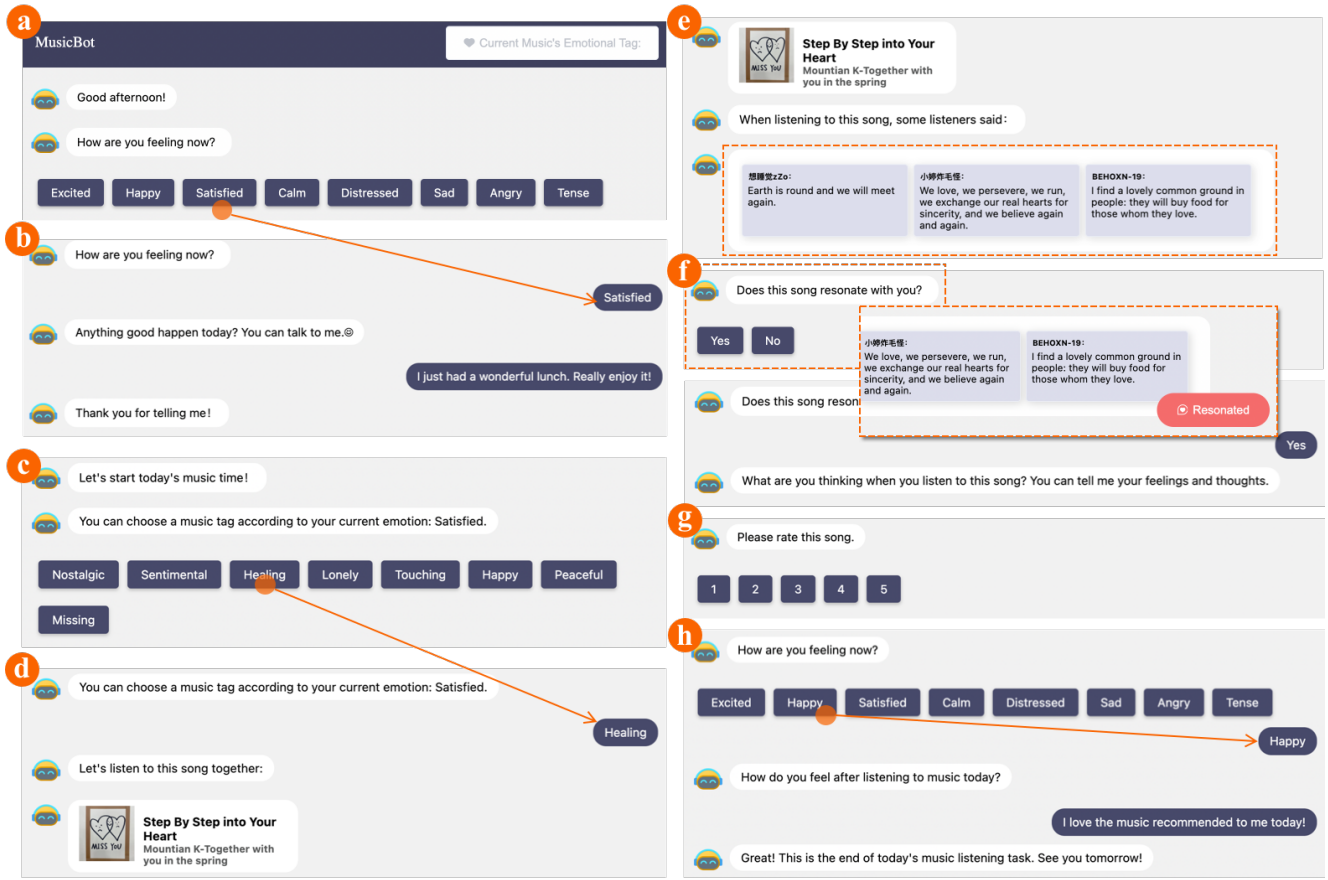


Figure 3: Screen shots of the dialog between a user and the conversational agent.

contain other listeners' feelings and thoughts. Figure 2 shows the user interface of the CA we designed, which consists of two main parts: *a dialog window* [Figure 2 (a)] and *a music player* [Figure 2 (b)]. The dialog window shows the conversations between the user and the CA, including eliciting the user's emotion, making music recommendations [Figure 2 (c)], and displaying music comments that reflect other listeners' feelings and thoughts [Figure 2 (d)]. In the following, we describe the detailed design and implementation of the CA used in our experiment.

3.1 Conversation Design

To design a CA that can guide people to listen to their feelings and express their inner thoughts while listening to music, we referred to the general procedure of receptive music therapy (i.e., the client listens to music and responds to the listening experience verbally, or non-verbally, e.g., by using art, dance [10, 35]) to design the conversation between users and the CA. Note that we used only the procedure of a receptive music therapy session to guide the conversation design of the CA, rather than integrating music therapy interventions in our CA. Specifically, the conversation procedure consists of three main parts (see Figure 3 that illustrates the dialog flow designed in the CA):

(1) *Preparation*: This part of the dialog is to make a brief assessment of users' current emotional status and to learn about their music preferences. In detail, the CA asks users first to indicate their current emotions by choosing one of the eight predefined emotions upon Russell's circumplex model [87] [Figure 3 (a)]. To establish rapport with users, the CA asks them to recall what happened that day made them happy [Figure 3 (b)]. Thereafter, the CA shows eight music tags and asks users to choose one [Figures 3 (c) and (d)]. This information was then used to generate a list of music recommendations (see next section).

(2) *Music Listening and Conversation*: During this part of the dialog, the CA presents users with five recommended songs and guides them to listen to their inner feelings and thoughts. If the system variant supports *social information*, the CA shows other listeners' comments on the song shortly after the music begins to play [Figure 3 (e)]. These comments were collected from NetEase Cloud Music, one of China's most popular social music communities (see more details in Section 3.3). Noted that our CA only presents social information (i.e., music comments) rather than simulating social interaction with users. Further, the CA also asks users to sense their resonance with the music: the CA with low *proactive guidance* shows a pink button, "Resonated" on top of the input field that allows users to indicate whether the song resonated with them,

while the CA with high *proactive guidance* asks users actively to indicate if the music resonates with them in dialogues [Figure 3 (f)]. If users indicate the emotional resonance with the listened music, the CA continues to guide users to express their feelings and thoughts. When it receives users' self-expression, the CA responds to acknowledge their sharing, such as saying "Thank you for sharing this with me!" or "Thank you for letting me know!" To control the effect of the quality of the recommendation, the CA also asks users to rate each song on a five-point scale [Figure 3 (g)].

(3) *Closing Dialogue*: This part of the dialog conducts a debriefing with users, during which the CA asks them to indicate their current emotions, and provide their feedback on that day's music listening experience [Figure 3 (h)].

3.2 Music Recommendation

Music psychology researchers have studied the relation between music and emotion extensively. One of the key functions of listening to music is to match users' current emotions and release their suppressed emotions [37]. Previous studies have also suggested that people tend to have different opinions on the choice of music to regulate mood and emotions, which depends upon the listener's preference for music, active engagement with music, and age [94]. Therefore, we employed a tag-based approach [26] to recommend music according to the user's desired emotions and their music preferences. Specifically, rather than matching the music with the emotion the user indicated, the CA asks the user to choose a music tag according to her/his current emotion [Figure 3 (b)]. In this way, our CA may recommend certain positive songs to those who prefer to listen to happy songs when they feel sad. The music tags were selected based on the NetEase Cloud Music tag system, which contains five categories of tags (i.e., languages, styles, activities, emotions, and themes). We selected eight emotion tags to collect music data to recommend: "Nostalgic", "Sentimental", "Healing", "Lonely", "Touching", "Happy", "Peaceful", and "Missing" [Figure 3 (c)]. We collected songs from playlists that contain one of the emotion tags above. To ensure the quality of music recommendations, we targeted the playlists NetEase Cloud Music has curated or the top recommended songs with the most likes. In total, we collected 200 songs for the eight emotion tags, each of which had 25 songs on average. Based on the emotion tag indicated by the user, the CA selects five songs as that day's music recommendations. In addition, users can change the music tag by clicking the button in the upper right corner if they feel that the songs do not represent their emotions.

3.3 Music Comments

The CA with *social information* presents music comments other listeners posted on the song playing currently. The music comment card displays three music comments with the author's nicknames to let users know what others felt and thought when listening to that song [Figure 2 (d)]. All of these comments were left by users of NetEase Cloud Music [14]. First, we collected the top 20 comments ranked by the number of likes they received. Then, we filtered out certain comments that were irrelevant to the context of listening to music, such as jokes, seeking help for mental health problems, and sharing bereavement and grief. We carried out this filtering process because, on the one hand, we expected that the comments might encourage users to express their feelings and thoughts; on the other hand, we wanted to avoid certain drastic comments due to ethical considerations (see Section 4.5). In the end, we kept three comments that disclosed other listeners' feelings and thoughts for each song.

4 USER EXPERIMENT

To investigate the effect of our CA design on users' self-awareness (at the *task* level), psychological need satisfaction as well as user acceptance (at the *technology* level), and mental well-being (at the *life* level), we conducted a 2 [*proactive guidance (PG)*: low vs. high] X 2 [*social information (SI)*: none vs. music comments] between-subjects user experiment, in which the participants were asked to use the CA we designed and implemented for five days within a week. Before we began our formal experiment, we conducted a pilot study with three people to ensure that the study procedure was complete and the system was functional. In this section, we will present the two system design factors (i.e., *PG* and *SI*) in our user experiment, and the operation of the experiment, including experimental procedure, participant recruitment, and measurements.

4.1 System Design Factors

This study investigated the effects of two system design factors in the CA, *proactive guidance* and *social information*, on users' self-awareness, psychological need satisfaction, acceptance, and mental well-being.

4.1.1 Proactive Guidance (PG). We manipulated the *PG* in the CA by offering either *low* or *high PG* while listening to music. The low *PG* provides a button "Resonated" for participants to indicate whether they have feelings or thoughts at any time about the song playing currently. In contrast, the high *PG* uses a nudging approach to support self-awareness through conversations [74]. For example, after playing a song for 30 seconds, the CA would ask the participants "*Does this song resonate with you?*" to guide participants to experience themselves and sense if the music resonates with them.

4.1.2 Social Information (SI). The *SI* presented in the CA refers to the music comments collected from NetEase Cloud Music where users with similar tastes in music gather and share their feelings about the music [18] (see Section 3.3). These music comments contain other listeners' emotional stories and personal thoughts [14, 22], which may help participants think about themselves and cultivate emotional resonance as well. To ensure exposure to the music comments, the CA that supports *SI* presents three well-selected comments in parallel immediately after the music plays (2 d).

4.2 Experimental Procedure

Figure 4 shows the flowchart of the five-day user study. On the first day, all participants were required to sign an online consent form before registering an account for using our CA (a web application). As when users use small-size screens (e.g., smartphones), they are likely to perform multiple activities simultaneously (e.g., listening to music and reading news and messages) [3], we required our participants to perform the experiment using personal computers (i.e., notebook or desktop) to interact with our CA, which

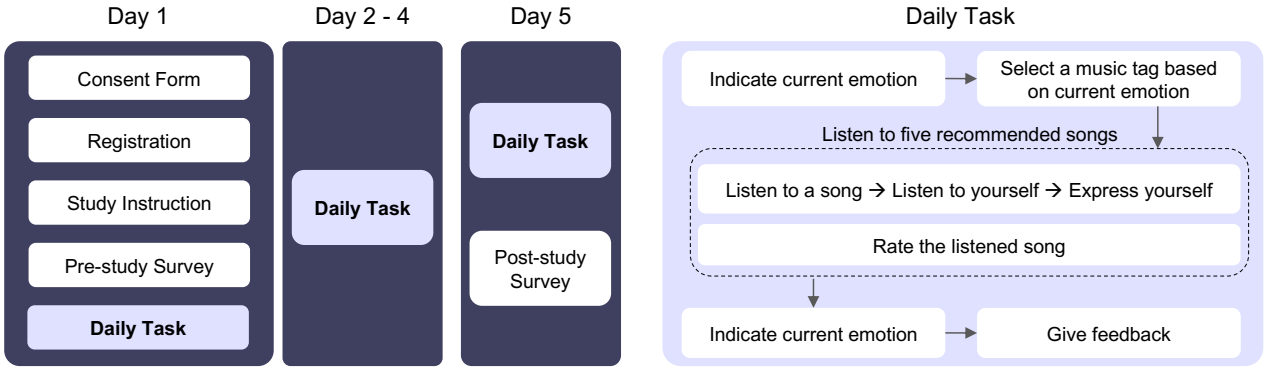


Figure 4: Experimental procedure of the five-day user study (left) and the flow of daily task (right).

guided users to increase their self-awareness (i.e., concentrate on their own feelings and thoughts) while listening to music. This was also because people tend to provide longer and higher-quality responses (e.g., conscientious and thoughtful answers) using personal computers than using mobile devices [3, 60]. After they registered, the participants received an introduction page that explained the study task and procedure, and they were instructed to finish the five-day study within a week. First, we determined whether the participants understood the introduction to the study by asking three multiple-choice questions, for example, "*How many songs should I listen to for each day?*" Thereafter, the participants who passed the study introduction test completed a pre-study survey that collected their demographic information (i.e., age, sex, and occupation) and measured their mental well-being. The participants then performed the *daily task*, including 1) listening to five recommended songs, 2) expressing their personal feelings and thoughts the music elicited if any, 3) rating each song on a five-point scale when it ended, 4) reporting their emotions before and after they finished listening to five songs, and 5) providing feedback on the day's music listening experience. In the end, the participants were asked to fill out a post-study survey after they finished the daily task on the last day (Day 5).

4.3 Participant Recruitment

We recruited the participants on WeChat, the most popular social media platform in China. We advertised our study primarily in several local university students' WeChat groups, and most members of these groups are recent graduate students. In total, 208 volunteers registered user accounts for our study (171 completed the pre-study survey), 132 of whom completed the first-day task, and 103 completed the five-day tasks. There are two possible reasons that the 39 participants did not finish the first-day task after they completed the pre-study survey. First, some participants signed up for the study probably simply to satisfy their curiosity about a CA. Second, they may have had difficulty participating in the study using personal computers (e.g., some people may not have personal computers in hand during non-office hours). Figure 5 shows user retention during the five-day study. Eight to ten participants dropped out of the study after they finished the daily task on the first four days. Two participants dropped out of the study on the last day. The dropout rate (29/132 = 21.97%) during the five-day study may be attributed primarily to the significant user effort that a longitudinal design usually entails [45].

In addition, we filtered out five participants who did not pass the attention checks on the post-study survey and eight who did not interact with the CA for more than 30 minutes between two songs in two daily tasks.[1] Finally, we obtained valid responses from 90 participants, among whom 57.78% were female (N = 52) and 42.22% were male (N = 38). All participants were young adults [19-22 (N = 29), 23-26 (N = 40), 27-30 (N = 15), > 30 (N = 6)], and more than half of the participants were students [students (N = 50), employees (N = 35), freelancers (N = 6)]. The participants who completed the study were compensated with a coupon for an e-commerce website worth 100 RMB (approximately 14.45 USD).

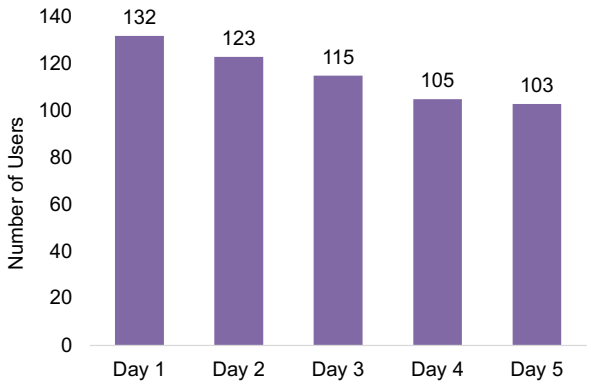


Figure 5: User retention during the five-day user study.

4.4 Measurements

During the experiment, we logged the participants' interaction with the CA to evaluate their actual self-awareness while listening to music. After the five-day study, we gauged the participants' perceived

[1]We determined a cut-off interval of 30 minutes because it exceeds the expected completion time (20-30 minutes) of the daily task. The eight participants were suspected of doing other things than listening to music.

Table 1: Explanation of Expression Depth and the Examples of Coded Responses

Depth	Explanation	Example
0	No expression of personal feelings and thoughts, incomplete and ambiguous expression	*"Yes"* *"Amazing!"*
1	Expression of superficial feelings and thoughts without ambiguity	*"This song makes me recall my younger life."* *"Sweet feeling of falling in love"*
2	Expression of relatively deep and personal feelings and thoughts beyond the comments on music	*"It makes me more eager to travel. I miss the time when I could travel around the world before the pandemic happened."* *"I recall my ex-girlfriend. Although she said she loved me I always complained that she did not love me enough. Maybe I was too suspicious and made her under too much pressure."*

satisfaction of the three basic psychological needs (i.e., autonomy, competence, and relatedness) suggested in Self-Determination Theory (SDT) [88] and their acceptance of the CA in a post-study survey. In addition, we assessed their mental well-being before and after the five-day study. All questions were presented in Chinese because all participants were Chinese. Therefore, we adopted the validated Chinese version of the mental well-being scale and translated the remaining questions into Chinese manually. Translation followed the parallel translation protocol [38] in which two authors finished the initial translation independently and addressed their differences by discussing with the third author. All authors who translated the questions are native Chinese speakers and have professional working proficiency in English.

4.4.1 Self-Awareness. In the context of our study, self-awareness refers to a mental state in which individuals pay attention to their feelings and thoughts [25]. Accordingly, we measured self-awareness by tracking the times that music resonated with the participants (emotional resonance), the number of words they used to express their feelings and thoughts about the music (expression length), and their degree of self-disclosure (expression depth).

Emotional resonance (self-reported). Being aware of inner feelings and thoughts the music aroused could be evidenced by clicking the "Resonated" button (in the system variant with low *PG*) or responding "yes" to the question "*Does this song resonate with you?*" that the CA asked (in the system variant with high *PG*). These two reporting methods were both intended to capture the emotional resonance that the participants consciously perceived while listening to music, and the difference between these two methods was associated with the system design factor (i.e., *PG*) in our CA. By counting the number of times that self-reported emotional resonance occurred, we know how many times the participant could be aware of their feelings and thoughts.

Self-expression. The evaluation of self-expression included the *length* and *depth* of expression. We measured the expression length by calculating the mean number of words that the participant used to express the feelings and thoughts for all recommended songs. The depth of expression was measured by coding manually according to the degree of self-disclosure that Barak and Gluck-Ofri defined [7]. Table 1 shows the explanations and two example responses for each of the three depth levels (0,1,2). In total, we recorded 871 responses to the self-expression questions, e.g., "*What does this song remind you of? You can tell me your feelings and thoughts.*" Two authors first coded 87 randomly selected responses independently and then discussed their differences in coding to ensure a consistent coding criterion. Finally, they finished coding the remainder of the responses in mutual agreement.

4.4.2 User Need Satisfaction and User Acceptance. Table 2 lists the questions we adopted to measure user satisfaction of the three psychological needs and user acceptance. Although previous studies have validated these questions, we still performed a confirmatory factor analysis (CFA) to ensure the validity of the constructs measured by the questions adopted. We retained a single item for two constructs: perceived autonomy and perceived competence, because the model failed to keep at least three items for one construct after the items with low factor loadings were removed.[2] The factor loadings of all of the question items were larger than 0.5, which met the accepted level.[3] Cronbach's alphas and average variance extracted (AVE) of each construct also indicated good internal consistency and convergent validity.[4]

User need satisfaction. People have three key psychological needs. With the support of these three needs, users are more likely to achieve their well-being goals [88]. We measured perceived autonomy by adopting three questions proposed in previous studies [97, 109]. The measures of perceived competence and relatedness were based upon the questions Xi and Hamari developed [109].

User acceptance. User acceptance was measured based upon the questions of three aspects developed in the Technology Acceptance Model (TAM) [105]: perceived usefulness, perceived ease of use, and intention to use.

The measure of all question items above was based on a seven-point Likert scale.

4.4.3 Mental Well-being. This was measured by using a short 7-item version of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) [101], which measures mental well-being based upon the respondents' experiences during the past week. For example,

[2]In CFA, each factor should retain at least three items; otherwise the factor should be treated as a single-item factor. A factor with two items cannot be identified with negative degrees of freedom [36]

[3]The cut-offs for factor loading: 0.32 (poor), 0.45 (fair), 0.55 (good), 0.63 (very good), or 0.71 (excellent) [98].

[4]The cut-offs for Cronbach's alpha: 0.5 (poor), 0.6 (questionable), 0.7 (acceptable), 0.8 (good), or 0.9 (excellent) [31]. The accepted value of AVE is greater than 0.5 [36].

Table 2: Post-Study Questionnaire to Measure Users' Perceived Need Satisfaction and User Acceptance

Construct	Items (each statement rated on a seven-point Likert scale)	Factor Loading
User Need Satisfaction		
Perceived Autonomy		
	When using this music chatbot, I have different options.	
	I feel free to express my ideas and opinions with this music chatbot.	
	I feel free from outside pressures when I use this music chatbot.	
Perceived Competence		
	I am satisfied with my performance when I use this music chatbot.	
	I feel like an expert using this music chatbot.	
	I feel like a competent person when I use this music chatbot.	
Perceived Relatedness *(Cronbach alpha: 0.800, AVE: 0.578)*		
	This music chatbot helps me to feel part of a larger community.	0.732
	This music chatbot makes me feel connected to other people.	0.734
	This music chatbot doesn't support meaningful connections to others.	0.809
User Acceptance		
Perceived Usefulness *(Cronbach alpha: 0.850, AVE: 0.665)*		
	I would find this music chatbot useful in my life.	0.877
	Using this chatbot would improve my satisfaction with life.	0.864
	Using this music chatbot would improve my awareness of my own feelings.	0.701
Perceived Ease of Use *(Cronbach alpha: 0.810, AVE: 0.604)*		
	My interaction with this music chatbot would be clear and understandable.	0.765
	I found this music chatbot confusing to use.	0.870
	Using this music chatbot was taxing.	0.697
Intention to Use *(Cronbach alpha: 0.941, AVE: 0.841)*		
	I intend to use this music chatbot in my life.	0.900
	I expect that I would use this music chatbot in the future.	0.921
	I intend to use this music chatbot regularly in the future.	0.931

"*I have been dealing with problems well.*" The seven statements are worded positively and measured on a five-point Likert scale.

4.4.4 Music Rating and Engagement.

Music rating. This refers to the extent to which the music matches the participants' preferences and emotions. Each recommended song is rated on a five-point scale. For each participant, the music rating overall is the mean rating of all recommended songs.

Engagement duration. The engagement duration refers to the mean amount of time the participants spent finishing the daily task. This duration includes the time spent in listening to music and viewing music comments, and the time the participants took to rate the music.

4.4.5 Open-Ended Questions. To obtain more detailed feedback on self-awareness that the CA supported, we asked three open-ended questions as follows: 1) *In what situation did the music or comments resonate with you (e.g., evoked a memory, sentiment, or nostalgia) while listening to music with our CA*; 2) *Were you willing to share your feelings and thoughts with the CA when the music or comments resonated with you? Why*; 3) *Did the CA help you express your feelings and thoughts? What do you think of the CA with respect to its support for self-awareness*?

4.5 Ethical Considerations

Conversational agents have become increasingly popular in daily use to improve mental well-being, and can be designed to provide various types of support (e.g., informational and emotional support) [57]. However, it is still unclear what should be expected from a CA or the appropriate levels of support that a CA should provide. When designing the CA to strengthen people's self-awareness, we considered each conversation element carefully and took a cautious step to design a CA capable of guiding people to listen to their feelings and express their inner thoughts. In particular, we took extra care to ensure that the music comments presented during the conversation did not contain any serious mental health problems or sensitive topics, such as suicide, depression, bereavement, and grief, which might be likely to trigger intense emotional responses [53]. Further, to allow more efficient communication and mitigate any potential harm during the experiment, all participants were asked to add our official WeChat account at the start of the experiment, so that they could contact the research team immediately to ask for help when they encountered any problems.

In addition to the ethical considerations mentioned above, we endeavored to conduct this study ethically. Before we conducted our experiment and collected any data, we got the approval of the

university's Research Ethics Committee (REC). Then, at the start of each study, the informed consent was obtained from the participant. Our participants were informed explicitly that they were free to choose to participate in, or withdraw from, the study at any point without any negative repercussions, and they were compensated at a reasonable hourly wage upon completion. Further, we ensured that all data collected were kept confidential and were only used for our research purpose.

5 ANALYSES & RESULTS

We analyzed both the quantitative and qualitative data collected in the study. Based upon the Self-Determination Theory (SDT) and the Technology Acceptance Model (TAM), we employed structural equation modeling (SEM) to build a path model to analyze all factors in the model simultaneously rather than separately. The resulting model allowed us to investigate the effects of the two system design factors, *proactive guidance (PG)* and *social information (SI)*, on self-awareness (RQ1 at the *task* level), the factors that influenced users' need satisfaction and their acceptance of CA (RQ2 at the *technology* level), and the CA's influence on mental well-being (RQ3 at the *life* level). For convenience, we introduced the notation, "PG ∗ SI", to denote different system variants. For example, High PG ∗ None denotes the variant that offers high *proactive guidance* but no *social information*, and Low PG ∗ MC denotes the variant that provides low *proactive guidance* and *social information*.

We performed an SEM analysis using the R package (version 0.6-12).[5] According to the rule that each observable variable needs to retain at least five subjects [11], our sample size (N = 90) met the minimum sample size (N = 60) required for 12 observable variables. In addition, we performed Shapiro's tests to determine all variables' normality and found that most of the variables were not distributed normally. Thus, we selected the maximum likelihood parameter (MLR) estimator with robust standard errors (Huber-White) rather than the default maximum likelihood (ML) estimator [28]. The constructs that the CFA validated were organized according to our hypothesized paths in SEM. Taking into account the potential effects of demographic information and music listening habits on listeners' emotional reaction to music [17, 86], we included age, gender, occupation, and users' previous music listening behavior[6] as control variables in the initial model. However, all of them were not related significantly to other variables, so we removed them in our subsequent analyses. After pruning the non-significant paths, the final model had a good model fit[7] ($\chi^2(173) = 223.987$, p<.001, Comparative Fit Index (CFI) = 0.96; Tucker-Lewis Index (TLI) = 0.95, RMSEA = 0.06; 90% CI: [0.02, 0.08]). Figure 6 presents the causal relations among the factors. The number in the arrow (A→B) between two factors is the estimate of β that indicates the degree of change in B for every unit of change in A. The number in parentheses is the standard error of this estimate. According to the convention, the paths with non-significant effects (p ≥ .05) were excluded from the model.

5.1 Self-Awareness in Music Listening (RQ1)

With the task to feel the resonance with music and express feelings and thoughts that the music elicited, we evaluated our CA based upon three self-awareness measures: emotional resonance, expression length, and expression depth.

5.1.1 Effects of Proactive Guidance (PG). Figure 6 shows that *PG* had a direct positive effect on *emotional resonance* (β = 0.542, p <.05), suggesting that high *PG* leads to more emotional resonance than low *PG*. Further, through mediation [*PG* → *emotional resonance* → *expression length* → *expression depth*], *PG* had an indirect positive effect on *expression length* (β = 0.380, p <.05) and *expression depth* (β = 0.272, p <.05). As shown in Table 3, with the same *SI* setting, low *PG* yielded less emotional resonance and shorter expressions than its counterpart; but it led to deeper disclosure than high *PG* when *SI* was provided. These effects of *PG* indicated that low *PG* may require participants to have a strong motivation to share their feelings and thoughts when the music has resonated with them. However, high *PG* encouraged participants to think of themselves and share their feelings and thoughts through more conversational guidance [53, 74], and thus induced more emotional resonance and longer expression.

5.1.2 Effects of Social Information (SI). *SI* had a direct positive effect on expression depth (β = 0.709, p <.001), indicating that presenting music comments (MC) helped participants express deeper feelings and thoughts. As shown in Table 3, the combination of high *PG* and *SI* (MC) achieved the highest score for expression depth (M = 1.25, SD = 0.34), while offering high *PG* without *SI* led to the lowest score (M = 0.77, SD = 0.43). It appears that the effect of high *PG* on expression depth depends upon SI, which was evidenced by a significant interaction effect of *PG* and *SI* on expression depth (see Figure 7), F (1, 86) = 4.71, p < .05, η_p^2 = 0.05. However, we did not find any direct or indirect effects of *SI* on emotional resonance and expression length.

5.1.3 Effects of Music Rating. As a situational factor, the music rating reflects the degree to which the recommended music matches the participants' preferences and emotions. System design factors did not affect the music rating significantly. However, the music rating had a direct positive effect on emotional resonance (β = 0.482, p <.001), while it had a direct negative effect on expression length (β = -0.207, p <.05). Still, the direct negative effect on expression length was counteracted by an indirect positive effect [*music rating* → *emotional resonance* → *expression length*]. The combined total effect was β = 0.131 (p =.330), indicating that songs that satisfied the participants were more likely to trigger emotional resonance and encourage them to write more words to express their feelings and thoughts.

5.1.4 Self-Awareness Over Time. In addition, we analyzed the trends of the three self-awareness measures by different variants of the system during the study period. The five-day study helped us reduce the novelty effect on the analysis of self-awareness. Despite

[5] https://lavaan.ugent.be/

[6] We measured users' previous music listening behavior by adopting two statements from Goldsmiths Musical Sophistication Index [65]: "*I spend a lot of my free time doing music-related activities.*", and "*I enjoy writing about music, for example on blogs and forums.*" Both of them were rated on a 7-point Likert scale.

[7] Hooper et al. [40] suggested cut-off values for several fit indices: CFI ≥ 0.95, TLI ≥ 0.95, and RMSEA < 0.08, with the upper limit of its 90% CI below 0.10.

Table 3: Descriptive Statistics of Self-Awareness Measures

	High PG ∗ None (N=23) Mean (SD)	High PG ∗ MC (N=22) Mean (SD)	Low PG ∗ None (N=22) Mean (SD)	Low PG ∗ MC (N=23) Mean (SD)
Emotional Resonance (times)	10.70 (5.91)	**11.27 (6.13)**	7.18 (6.59)	9.08 (7.06)
Expression Length (words)	18.64 (22.71)	**21.00 (15.01)**	13.61 (12.51)	17.17 (15.19)
Expression Depth	0.77 (0.43)	**1.25 (0.34)**	0.89 (0.42)	0.96 (0.50)
Music Rating	**3.64 (0.67)**	3.39 (0.48)	3.54 (0.65)	3.61 (0.68)
Engagement Duration (seconds)	260.21 (47.21)	**293.72 (82.42)**	270.16 (93.29)	275.72 (53.34)

Note: The highest value of each dependent variable is marked in bold.

Table 4: Descriptive Statistics of User Need Satisfaction and User Acceptance

	High PG ∗ None (N=23) Mean (SD)	High PG ∗ MC (N=22) Mean (SD)	Low PG ∗ None (N=22) Mean (SD)	Low PG ∗ MC (N=23) Mean (SD)
User Need Satisfaction				
Perceived Autonomy	5.61 (1.16)	5.22 (1.38)	5.77 (0.92)	**6.13 (1.01)**
Perceived Competence	5.22 (1.35)	4.95 (1.00)	5.09 (1.11)	**5.26 (1.39)**
Perceived Relatedness	4.16 (1.23)	3.89 (1.16)	3.88 (1.23)	**4.45 (1.48)**
User Acceptance				
Perceived Usefulness	4.74 (1.29)	4.35 (1.11)	4.98 (1.24)	**5.20 (1.12)**
Perceived Ease of Use	5.10 (1.08)	4.59 (1.26)	5.03 (0.97)	**5.39 (1.10)**
Intention to Use	4.29 (1.47)	4.08 (1.15)	**4.83 (1.35)**	4.80 (1.37)

Note: The highest value of each dependent variable is marked in bold.

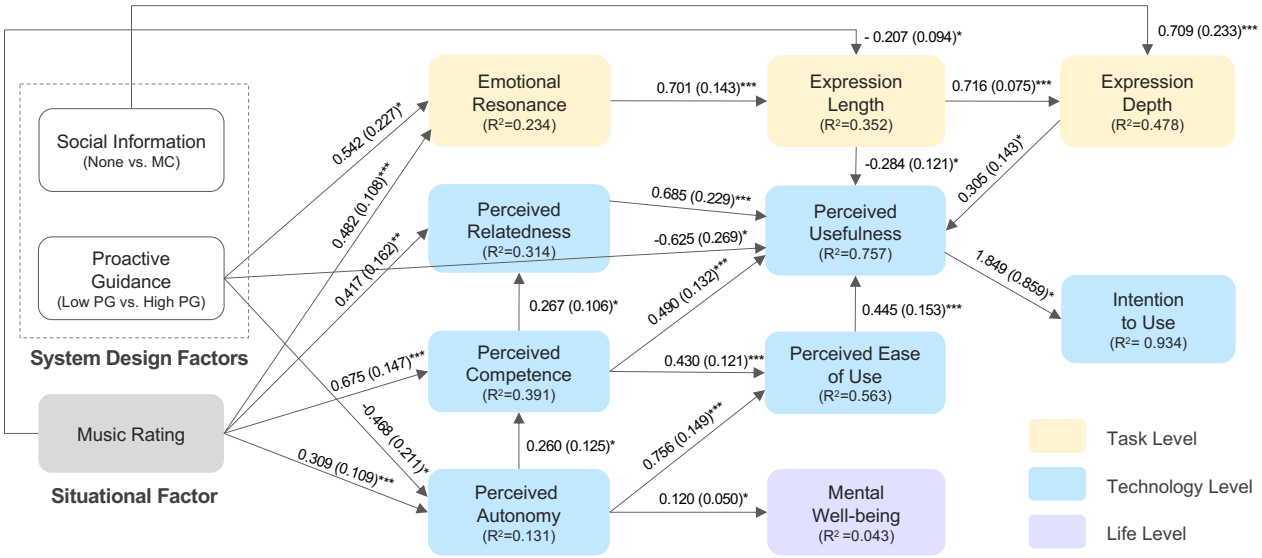


Figure 6: The structural equation model illustrates the relation among the factors of several aspects, including system design factors, self-awareness, user need satisfaction, user acceptance, and mental well-being. The numbers on the arrows represent the β coefficient and standard error (in parentheses) of the effect. Significance levels: *** $p < .001$, ** $p < .01$, * $p < .05$. R^2 is the proportion of variance the model explained.

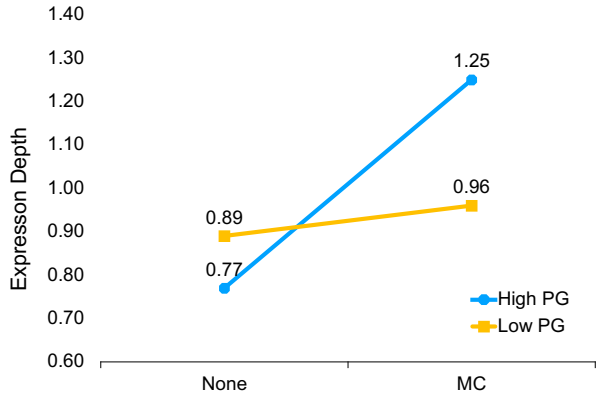


Figure 7: Interaction effect of *PG* and *SI* on expression depth. *SI* (MC) had a positive effect on expression depth, especially when combined with high *PG*.

the declining trend in self-awareness in the first four days, we found a slight increase in most self-awareness measures on Day 5.

Emotional resonance. Figure 8(a) shows that emotional resonance (times) continued to fall in Low PG ∗ None. Generally, the values on the first day were the highest in all system variants. Moreover, high *PG* always induced more self-reported emotional resonance than low *PG*, which is consistent with the significant positive effect of *PG* on emotional resonance.

Expression length. As shown in Figure 8(b), expression length fluctuated in all variants during the study. The value remained at the Day 1 level after the five-day task only for Low PG ∗ MC. For other variants, the initial values were the highest. Similarly, Low PG ∗ None was the lowest for the majority of the time.

Expression depth. Figure 8(c) shows that expression depth continued to decrease from Day 1 to Day 4 and then rose for most variants. High PG ∗ MC appeared to be the best on all five days, and Low PG ∗ None was still the lowest the majority of the time. In addition, offering *SI* often led to deeper expression, particularly when combined with high *PG*, which confirms the positive effect of *SI* on expression depth. However, in general, it demonstrated a decreasing trend during the study.

5.2 User Need Satisfaction and User Acceptance (RQ2)

As per Self-Determination Theory (SDT), satisfying the three basic psychological needs, i.e., autonomy, competence, and relatedness, determines the motivation to use the system [88]. Further, a previous study on a mobile assessment app showed the positive effect of user need satisfaction on factors of the technology acceptance model (TAM) [70]. Therefore, we analyzed the way the need satisfaction factors mediated the effect of the system design factors on user acceptance in our study.

5.2.1 Effects on User Need Satisfaction. There was only one significant effect of system design factors on user need satisfaction. Specifically, *PG* had a direct negative effect on *perceived autonomy* (β = -0.468, p <.05), suggesting that high *PG* led participants to perceive less autonomy. Moreover, the music rating had direct positive effects on all the need satisfaction factors, among which the effect of the music rating on perceived competence was the strongest (β = 0.675, p <.001).

5.2.2 Effects on User Acceptance. We observed a direct negative effect of *PG* on perceived usefulness (β = -0.625, p <.05), indicating that participants perceived that the CA, which guided users to think about their feelings proactively while listening to music, was less useful. In addition, *PG* had multiple indirect negative effects on perceived usefulness [*PG* → *mediators (self-awareness and need satisfaction)* → *perceived usefulness*]. Thus, the combined total negative effect was strong (β = -0.912, p <.01). In contrast, the path [*SI* → *expression depth* → *perceived usefulness*] demonstrated that *SI* had an indirect positive effect on perceived usefulness (β = 0.216, p =.07), suggesting that the presentation of music comments increased the participants' perceived usefulness of the CA. Similarly, the positive effects of music rating on perceived usefulness were mediated by all of the need satisfaction factors, suggesting that the higher the music rating, the more useful the participants' perception of the system.

Perceived autonomy and perceived competence mediated two indirect negative effects of *PG* on perceived ease of use. Combined, we obtained a total effect of -0.406 (p <.05). Through the same mediators, the music rating had two indirect positive effects on perceived ease of use, with a total effect of 0.558 (p <.001).

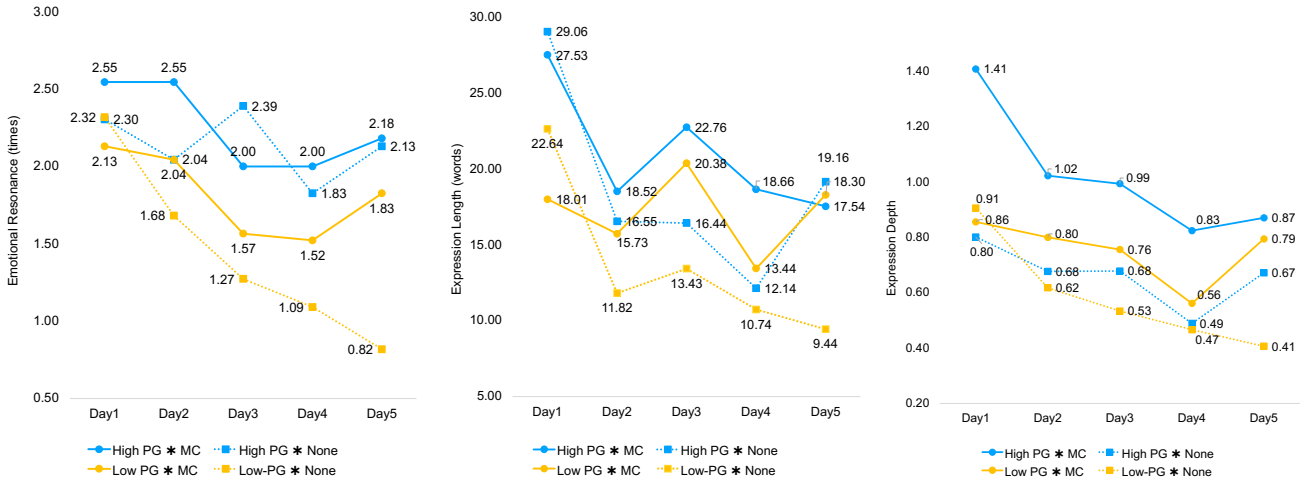
As perceived usefulness influenced intention to use strongly (β = 1.849, p <.05), all effects (direct and indirect) on perceived usefulness can be attributed to *intention to use*. Thus, both the system design factors and the situational factor (music rating) influenced user acceptance of the CA differently by satisfying the three basic psychological needs.

Table 4 shows an interesting phenomenon, in which presenting *SI* decreased the values for all variables when being combined with high *PG*, but increased the values for all variables (except intention to use) when combined with low *PG*. Further, the system variant Low PG ∗ MC was the top-ranked for all variables except intention to use. However, the interaction effects of the two design factors on these variables were insignificant.

5.3 CA for Mental Well-being (RQ3)

The difference between the mental well-being scores measured before and after the study reflected the CA's potential effect on mental well-being. In addition, we analyzed changes in participants' emotions before and after the daily task.

5.3.1 Effects on Mental Well-being. The model showed no direct effect of the two design factors on mental well-being, but mediation [*PG* → *perceived autonomy* → *mental well-being*] showed a weak indirect effect (β = -0.056, p =.100). Through the same mediator, perceived autonomy, the music rating also had a weak indirect effect on mental well-being. Thus, we can say that satisfying users' needs for autonomy may support human mental well-being. Overall, the participants' mental well-being showed a minor improvement (mean = 0.165, SD = 0.624) on a five-point scale. However, the group of participants with low initial mental well-being (below the median



(a) The trend of emotional resonance. (b) The trend of expression length. (c) The trend of expression depth.

Figure 8: The trends of the three self-awareness measures by different variants of the system during the study. (a) Emotional resonance (times) continued to fall in Low PG ∗ None during the five day. (b) Expression length fluctuated in all system variants and Low PG ∗ None was the lowest most of the time. (c) Expression depth decreased from Day 1 to Day 4 and then rose for most variants. High PG ∗ MC was the best on all five days.

of well-being values) had a better improvement (mean = 0.398, SD = 0.644).

5.3.2 Emotion Improvement. The eight emotions were categorized into four negative emotions (distressed, sad, angry, tense) and four positive emotions (excited, happy, satisfied, calm) according to Russell's circumplex model [87]. Figure 9 shows the percentage of the participants who reported positive emotions in two sessions: the pre-task (light purple) and the post-task (dark purple). The percentages measured before the task fluctuated throughout the study period. However, the percentages were always higher in the post-task (approximately 90%) than in the pre-task, suggesting that performing the daily task with the CA may improve the users' emotions.

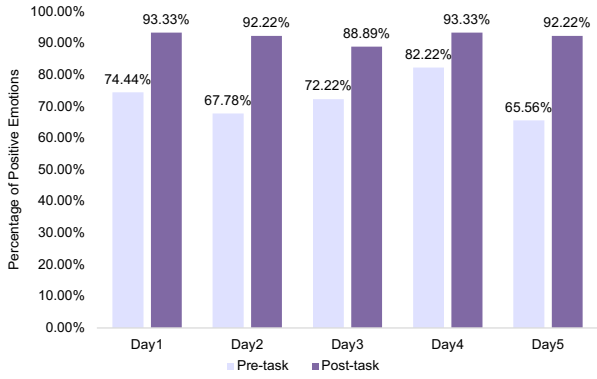


Figure 9: The percentage of positive emotions measured before and after the daily task.

5.4 Subjective User Feedback

Following an open-coding protocol, one author analyzed participants' responses to the three open-ended questions (see Section 4.4.5) in the post-study survey. The feedback referred to four main aspects: emotional resonance, self-expression, support for self-awareness, and benefits to mental well-being.

5.4.1 Emotional Resonance. Recommending music that matched users' preferences and emotions appeared to be crucial to trigger emotional resonance. Most of our participants (N=85) in our study mentioned that the listened music and the lyrics can resonate with them. For example, two participants said,

> *"The songs are more likely to resonate with me if they are familiar to me..."* (P6, Low PG ∗ MC)
> *"... the songs that resonate with me should be able to trigger my deepest memories and arouse my past emotions."* (P68, High PG ∗ None)

In addition, three participants explicitly mentioned that they were touched by seeing the same feelings and thoughts expressed in the music comments. For example, a participant noted,

> *"... the music tends to resonate with me if the comments for this music share some same experience with me..."* (P90, High PG ∗ MC)

5.4.2 Self-Expression. The users' acceptance of the CA influenced their willingness to express their feelings and thoughts to it, echoing the previous findings in human and chatbot communication [95]. According to the participants' feedback, 60% of the participants (N = 54) were willing to share their feelings and thoughts with the CA when emotional resonance occurred. As two participants mentioned,

> *"I would like to, because we don't know each other, and it's fun to tell a stranger what's on your mind."* (P59, Low PG ∗ MC)
> *"...this CA is not a real person after all, so I was more willing to talk about my feelings, without caring about audiences' responses and my wording."* (P83, Low PG ∗ MC)

However, some participants were reluctant to express their thoughts to a CA that demonstrated limited humanness. For example, one participant said,

> *"I was a bit reluctant to share because it is just a machine, and looks a little indifferent."* (P17, High PG ∗ None)

In addition, self-expression also depended on the current emotions the music aroused. Sometimes, it was difficult for participants to describe their feelings and emotions. For example, one participant noted,

> *"If I want to share my feelings, I will only say it when I can describe or express the feeling. Some emotions were subtle and difficult to describe accurately or made me feel ashamed. But if the emotion were very strong, I would like to share it."* (P69, Low PG ∗ None)

Users' privacy concerns about the CA may also be a reason that they did not share feelings and thoughts, as one participant mentioned,

> *"... not sure how well privacy protection is done in the backend."* (P69, Low PG ∗ None)

5.4.3 Support for Self-Awareness. The responses to the third question indicated that 68.9% of the participants (N = 62) perceived the CA's support in self-awareness. It appears that reading music comments helped them improve their self-expression skills.

> *"...At the beginning of using it, I had a lot of thoughts but no idea how to write them out clearly. Later, I know how to express my feelings and thoughts more concisely and clearly after reading other comments."* (P45, High PG ∗ MC)

Moreover, listening to music and expressing themselves also improved the participants' ability to sense their feelings and thoughts.

> *"... The task of listening to music and expressing myself trained my skills in expressing abstract feelings through language... Although the feeling is somewhat difficult, long-term training with this system should be more helpful"* (P76, Low PG ∗ None)

However, some participants believed that the CA would bring more benefits to people's self-expression only when it offers more interactive guidance for self-expression.

> *"Not really. I think the process of self-awareness is a step-by-step reflection and exploration. It is hard to get there if it is just one-sided thinking without interactive guidance"* (P36, Low PG ∗ MC)

5.4.4 Benefits to Mental Well-being. Despite the minor improvement in mental well-being measured by the WEMWBS questionnaire, some participants (N=16) acknowledged the positive influence of CA on mental well-being. We show several examples below.

> *"I often used this chatbot in the early morning. I felt like I had better emotions and was full of passion for life and work after finishing the task"* (P56, Low PG ∗ None)
> *"...chatbot encouraged me to face my emotions rather than avoiding them...I rarely had time to face my emotion and think about my inner thoughts. When expressing my emotion, even if I was not happy, I wanted to release negative emotions through self-expression... This chatbot made me feel relaxed. Many times when I felt nervous or sad, interacting with it allowed me to calm down..."* (P80, Low PG ∗ None)
> *"Expressing my emotions makes me happy. I enjoyed reading what other people shared in music comments. Sometimes when I found other people feeling the same way, I felt connected with them"* (P45, High PG ∗ MC)

6 DISCUSSION

In this research, we designed and implemented a novel conversational agent (CA) system to support self-awareness while listening to music. In particular, we considered the two system design factors, i.e., *proactive guidance (PG)* and *social information (SI)* in our CA design and performed a five-day longitudinal study to evaluate our system. As inspired by a model for well-being-supportive design [78], we investigated the effect of our CA design on user experience and well-being at three different levels: task (RQ1), technology (RQ2), and life (RQ3). This section briefly summarizes our main findings as follows.

- **Self-awareness was significantly influenced by the two system design factors.** High *proactive guidance (PG)* in the CA encouraged significantly more emotional resonance, and *social information (SI)* led to significantly deeper self-expression.
- **User acceptance was significantly influenced by the two system design factors, mediated by users' self-awareness and their need satisfaction.** With high *PG* the CA provided, participants tended to have a reduced feeling of autonomy and perceived that the CA was less useful.
- **User feedback revealed the CA's potential to improve emotions and support mental well-being.** Performing the daily task had a consistently positive effect on the participants' emotions over five days. Despite a minor increase in mental well-being measures, the participants' comments reflected the benefit they perceived of interacting with the CA to support self-awareness and mental well-being.
- **Music rating influenced emotional resonance and users' need satisfaction directly.** If the music matched the users' preferences and emotions, it was more likely to resonate with them and satisfy their needs for autonomy, competence, and relatedness.

In the following, we discuss our main findings and the implications derived for the future design of technology to foster self-awareness and support mental well-being.

6.1 Supporting Self-Awareness

Feelings and thoughts aroused by music have been shown to increase human self-awareness [2]. The two design factors' positive effects on self-awareness imply that *proactive guidance* can encourage users to actively think about their feelings and thoughts, while *social information* may serve as examples to provoke deep self-expression. Similar to chatbot guidance for self-reflection [74], high *PG* led users to get in touch with the feelings and thoughts associated with music consciously. Nevertheless, some participants thought that the current CA's *PG* reminded them only to think, but did not help them know how to think. For example, one participant said, *"...It is hard to get there if it is just one-sided thinking without interactive guidance."* (P36, Low PG $*$ MC). Therefore, **future design may offer example-based guidance to help users understand themselves and describe how they feel** [111]. Furthermore, in our system, all of the music comments were related to self-disclosure (e.g., feelings, experiences), which could connect the listeners with the music comment posters if they shared similar feelings or experiences [18]. Such *social information* can transform individual listening experiences into collaborative listening experiences and cultivate self-expression [58]. As per social exchange theory [20], online self-disclosure has a reciprocal effect that could promote others' self-disclosure [7], which may explain the positive effect of social information on expression depth.

In addition to the effects of the two system design factors, music per se influences human emotions profoundly [44, 102]. Thus, the music rating influenced the occurrence of emotional resonance positively. The subjective feedback indicated that the participants were more likely to be touched by the music with which they were familiar or that reflected their emotions. Despite the music rating's direct negative effect on expression length, when combined with the indirect positive effect, the total effect was positive. Hence, we speculate that **improving music recommendations can benefit self-awareness, for example, by personalizing music recommendations based upon emotions and mood** [1, 5].

Conversational systems have been demonstrated to have great potential in psychoeducation and self-adherence [103]. Several chatbots have been developed that are able to self-disclose to promote users' self-disclosure, thereby promoting self-awareness and mental health [53, 55]. Other chatbots have been designed to improve emotional self-awareness by incorporating different programs, such as mood reporting, therapeutic exercises [63], coaching intervention [30], and storytelling [90]. To the best of our knowledge, our work is the first attempt to explore design opportunities for a musical CA that is intended to support self-awareness. Compared to certain existing CAs for music recommendation and exploration [15, 42], our CA supports users' self-awareness by guiding them to listen to their inner feelings and thoughts and express them when listening to music.

6.2 User Acceptance of CA

Our study demonstrated that the level of *PG* in the CA strongly affected the users' satisfaction of psychological needs and their acceptance of the CA. As shown in our results (Figure 6), high *PG* influenced the users' perceived autonomy and usefulness negatively. This result is consistent with previous observations that a highly proactive system encourages users to make a decision and take actions, which may be perceived as intrusive and controlling and hence reduce the feeling of autonomy [75]. In our case, although the CA with high *PG* encouraged increased occurrences of resonance, the CA with low *PG* allowed users to listen to music and their inner thoughts at their own pace, which may have evoked the users' feeling of freedom and, thus, enhanced their perceptions of autonomy and usefulness. Moreover, consistent with the SDT literature [52, 70, 78], the satisfaction of the needs for autonomy, competence, and relatedness mediated the relation between the design factor and user acceptance. Users' perceived autonomy has been shown to affect their satisfaction of the need for competence and relatedness, which further demonstrates the importance of perceived autonomy in the CA design [69, 110]. These results suggest that, **when designing a CA to support self-awareness, designers should be cautious about increasing the level of *proactive guidance* and respect users' freedom to express their feelings and thoughts (e.g., when to express and what to express) while listening to music.**

With respect to the *SI* (e.g., music comments) the CA provided, it was found to encourage users to demonstrate a deeper level of thinking and expression, which increased users' perceived usefulness of CA further. From our results in Table 4, we can see that when the CA with low PG provided *SI* during the interaction, users tended to perceive more autonomy. This may be because such *SI* fosters users' connectedness with others; as one participant said, "*... I enjoy reading what other people shared. Sometimes when I see other people feeling the same way about the same song, I feel as if I know them.*" (P45, Low PG $*$ MC). These results are roughly in line with those in prior research [55, 82], suggesting that computer agents' information sharing (such as chatbot's self-disclosure) can increase users' perceived intimacy and trust of the agents. Our findings extend the previous literature by suggesting that the **social information the CA provides may also help satisfy users' psychological needs, and hence, improve their acceptance.**

Meanwhile, most of our participants were willing to share their feelings and thoughts with our CA. Similar to other studies of CAs for mental well-being [53, 73, 106], they felt less pressure to share their personal feelings and thoughts with the CA than a real person. However, they asked for more interactivity (e.g., empathetic responses) and personalized experiences (e.g., personalized music recommendations that are associated with their music tastes) in the CA. From the perspective of SDT [78, 110], more personalized CA can foster users' feelings of autonomy and competence better, as personalization may create a sense of ownership (that supports user autonomy) and help users accomplish the task in which they are engaged (that fulfills the need for competence). Therefore, **further research should be undertaken to investigate how to design better interactivity and a personalized experience in a CA that supports users' self-awareness**.

6.3 Music Technologies for Mental Well-being

A large body of evidence has demonstrated the positive effects of music on people's mental health and well-being [91, 104, 107]. In light of the benefits of music and the popular use of computer software to access it, various music technologies have been designed

to enhance health and well-being [12, 58, 66], e.g., tangible devices to support emotional regulation in musical activities [12]. This study extends the previous literature by designing a CA to support self-awareness in the context of everyday music listening, which has been demonstrated to help improve people's emotions with daily use. Consistent with the literature [99], our study demonstrated that personal music preferences (denoted by music ratings in our study) influenced users' satisfaction of psychological needs, which affected the benefits (e.g., self-awareness) they reaped from listening to music. Therefore, further studies that integrate personalization into music technologies to strengthen self-awareness are suggested.

Our findings also indicated the importance of supporting user autonomy to enhance mental well-being in the music context (as shown in Figure 6), suggesting that further research on autonomy-supportive design for music technologies should also be investigated to enhance well-being better [16, 78]. Moving forward, the CA we designed could be improved further with respect to its benefits to mental well-being by taking into account additional findings or approaches (such as active and receptive music therapy [2, 61]) from music psychology research. For example, in addition to listening to music passively, it would be more engaging if a CA can interact with people to encourage them to sing or play songs, or discuss music, such as song lyrics, which may also be beneficial in fostering self-awareness and supporting mental well-being [2].

7 LIMITATIONS AND FUTURE WORK

To interpret our research findings fairly, we discuss the limitations in our work and propose future work that could address them.

First, the majority of the participants in our study are Chinese young adults with a good educational background. However, young and older adults may experience different effects of listening to music on emotion and well-being [33, 34]. Also, culture may influence people's perception of and interactions with CAs [76, 112]. Thus, the findings of this study may not apply to other cohorts, e.g., the middle-aged and elderly populations. Future work may validate our findings by evaluating the design concept with other populations beyond Chinese young adults.

Second, we used a uniform recommendation pool that contained 200 songs for all users, which ignored the users' special requirements for other music attributes, such as genres, artists, and language. The CA recommended music from the most popular playlists that matched the emotion tags the users indicated. As these playlists were not tailored to the users' music attribute preferences, some unfamiliar songs scarcely resonated with the users, even if they matched their emotions. Moreover, the selection of music comments ignored the users' current emotions and identities; thus, some unmatched comments may have influenced users' self-awareness negatively. For example, a single student user is unlikely to be inspired by a comment about parental stress. In the future, we may build a user profile based on the user's history data on various online music platforms (e.g., Spotify and NetEase Cloud Music) and tailor the music comments to listeners' emotions, which may help further support self-awareness in the experience of listening to music.

Third, we chose a rule-based approach with predefined response options to limit the complexity and uncertainties in dialog management. The CA's utterances may thus appear to be dull and fail to show empathy for users' self-expression. In light of certain technical limitations in our system, future work may be undertaken to diversify the manner of proactive guidance and improve our CA's communication skills by testing some small-sized pre-trained language models used to build an end-to-end empathetic chatbot [57].

Fourth, the participants were required to test the CA on their personal computers in a quiet place. However, young people today are more likely to use smartphones to listen to music [48]. The smartphone allows users to access music with fewer restrictions on location and time and create a context-aware experience of listening to music [49]. In the future, we will adapt the CA's user interface to mobile devices. For example, to fit the small display on a smartphone, we could hide the current music player and allow the play control on the music card to be displayed in the dialog window. Further, we will conduct a study to validate the effect of the CA on self-awareness and well-being in a smartphone context.

8 CONCLUSION

Listening to music has long been shown to influence mental well-being positively. By sensing and expressing the feelings and thoughts the music arouses, users can understand their emotions and behavior better and achieve greater self-awareness. However, little research has investigated how to support self-awareness while listening to music. CAs have been shown to positively affect human mental well-being by encouraging self-disclosure [53, 55] or offering various intervention programs [30, 90], which inspired us to propose this novel system that combines conversational interaction and music listening to support self-awareness. By conducting a five-day longitudinal user study, our study investigated the design opportunities for such a system by examining the way the two prominent design factors, *proactive guidance* and *social information*, influenced user experience and well-being at three different levels (*task*, *technology*, and *life*). In general, *proactive guidance* appears to be a double-edged sword; designers need to make a trade-off between its positive effect on self-awareness and its negative side effect on user acceptance. Regarding *social information*, it can help foster users' self-awareness, which is likely to increase user acceptance. Users' subjective feedback also demonstrated the potential benefits of our CA on mental well-being. Based upon our findings, we offered several design suggestions to improve such a system to help people raise their self-awareness while listening to music. We hope our work will encourage more researchers to design more powerful CAs to support various music activities (e.g., lyric analysis and music composition) that could support mental well-being.

ACKNOWLEDGMENTS

This work was supported by Hong Kong Research Grants Council (RGC) GRF project (RGC/HKBU12201620), Hong Kong Baptist University IG-FNRA project (RC-FNRA-IG/21-22/SCI/01), and Hong Kong Baptist University Start-up Grant (RC-STARTUP/21-22/23).

REFERENCES

[1] Ashu Abdul, Jenhui Chen, Hua-Yuan Liao, and Shun-Hao Chang. 2018. An emotion-aware personalized music recommendation system using a convolutional neural networks approach. *Applied Sciences* 8, 7 (2018), 1103.

[2] Dorit Amir. 2004. Giving trauma a voice: The role of improvisational music therapy in exposing, dealing with and healing a traumatic experience of sexual abuse. *Music therapy perspectives* 22, 2 (2004), 96–103.

[3] Christopher Antoun, Mick P Couper, and Frederick G Conrad. 2017. Effects of mobile versus PC web on survey response quality: A crossover experiment in a probability web panel. *Public Opinion Quarterly* 81, S1 (2017), 280–306.

[4] Vahid Assadi, Khaled Hassanein, et al. 2017. Consumer adoption of personal health record systems: a self-determination theory perspective. *Journal of medical Internet research* 19, 7 (2017), e7721.

[5] Deger Ayata, Yusuf Yaslan, and Mustafa E Kamasak. 2018. Emotion based music recommendation system using wearable physiological sensors. *IEEE transactions on consumer electronics* 64, 2 (2018), 196–203.

[6] Nilufar Baghaei, Lehan Stemmet, Andrej Hlasnik, Konstantin Emanov, Sylvia Hach, John A Naslund, Mark Billinghurst, Imran Khaliq, and Hai-Ning Liang. 2020. Time to get personal: individualised virtual reality for mental health. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–9.

[7] Azy Barak and Orit Gluck-Ofri. 2007. Degree and reciprocity of self-disclosure in online forums. *CyberPsychology & Behavior* 10, 3 (2007), 407–417.

[8] Kari Bjerke Batt-Rawden, Tia DeNora, and Even Ruud. 2005. Music listening and empowerment in health promotion: A study of the role and significance of music in everyday life of the long-term ill. *Nordic journal of music therapy* 14, 2 (2005), 120–136.

[9] Paula Bitrián, Isabel Buil, and Sara Catalán. 2021. Enhancing user engagement: The role of gamification in mobile apps. *Journal of Business Research* 132 (2021), 170–185.

[10] Stine Camilla Blichfeldt-Ærø, Gro Trondalen, and Sigrun Halvorsen. 2019. Facilitated music listening: Music therapy in an invasive cardiac procedure. *British Journal of Music Therapy* 33, 1 (2019), 27–38.

[11] Kenneth A Bollen. 1989. *Structural equations with latent variables*. Vol. 210. John Wiley & Sons.

[12] Majed Bou Ghanem and JungKyoon Yoon. 2022. Variapsody: Creating Three Interactive Music Listening Experiences that Use Diversified Positive Emotion Regulation Strategies to Promote Subjective Well-being. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–7.

[13] Kirk Warren Brown and Richard M Ryan. 2003. The benefits of being present: mindfulness and its role in psychological well-being. *Journal of personality and social psychology* 84, 4 (2003), 822.

[14] Jie Cai, Ruiqi Shen, and Starr Roxanne Hiltz. 2021. Choice of Social Music Systems in China: A Study of NetEase Cloud Music. In *Adjunct Publication of the 23rd International Conference on Mobile Human-Computer Interaction*. 1–6.

[15] Wanling Cai, Yucheng Jin, and Li Chen. 2021. Critiquing for music exploration in conversational recommender systems. In *26th International Conference on Intelligent User Interfaces*. 480–490.

[16] Rafael A Calvo and Dorian Peters. 2014. *Positive computing: technology for wellbeing and human potential*. MIT press.

[17] São Luís Castro and César F Lima. 2014. Age and musical expertise influence emotion recognition in music. *Music Perception: An Interdisciplinary Journal* 32, 2 (2014), 125–142.

[18] Chunhui Cheng. 2021. Self-expression of user comments in music social scenes: taking NetEase Cloud Music as an example (in Chinese). *NEW MEDIA RESEARCH* (2021).

[19] Riccardo Chianella, Marco Mandolfo, Riccardo Lolatto, and Margherita Pillan. 2021. Designing for self-awareness: evidence-based explorations of multimodal stress-tracking wearables. In *International Conference on Human-Computer Interaction*. Springer, 357–371.

[20] Karen S Cook, Coye Cheshire, Eric RW Rice, and Sandra Nakagawa. 2013. Social exchange theory. In *Handbook of social psychology*. Springer, 61–88.

[21] David Coyle, Conor Linehan, Karen Tang, and Sian Lindley. 2012. Interaction design and emotional wellbeing. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*. 2775–2778.

[22] David Cwir, Priyanka B Carr, Gregory M Walton, and Steven J Spencer. 2011. Your heart makes my heart move: Cues of social connectedness cause shared emotions and physiological states among strangers. *Journal of Experimental Social Psychology* 47, 3 (2011), 661–664.

[23] Claudia Daudén Roquet, Nikki Theofanopoulou, Jaimie L Freeman, Jessica Schleider, James J Gross, Katie Davis, Ellen Townsend, and Petr Slovak. 2022. Exploring Situated & Embodied Support for Youth's Mental Health: Design Opportunities for Interactive Tangible Device. In *CHI Conference on Human Factors in Computing Systems*. 1–16.

[24] Edward L Deci and Richard M Ryan. 1980. Self-determination theory: When mind mediates behavior. *The Journal of mind and Behavior* (1980), 33–43.

[25] Shelley Duval and Robert A Wicklund. 1972. A theory of objective self awareness. (1972).

[26] Claudiu S Firan, Wolfgang Nejdl, and Raluca Paiu. 2007. The benefit of using tag-based profiles. In *2007 Latin American Web Conference (LA-WEB 2007)*. IEEE, 32–41.

[27] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 2017. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR mental health* 4, 2 (2017), e7785.

[28] David A Freedman. 2006. On the so-called "Huber sandwich estimator" and "robust standard errors". *The American Statistician* 60, 4 (2006), 299–302.

[29] William J Froming, G Rex Walker, and Kevin J Lopyan. 1982. Public and private self-awareness: When personal attitudes conflict with societal expectations. *Journal of Experimental Social Psychology* 18, 5 (1982), 476–487.

[30] Silvia Gabrielli, Silvia Rizzi, Sara Carbone, Valeria Donisi, et al. 2020. A chatbot-based coaching intervention for adolescents to promote life skills: pilot study. *JMIR Human Factors* 7, 1 (2020), e16762.

[31] Darren George and Paul Mallery. 2019. *IBM SPSS statistics 26 step by step: A simple guide and reference*. Routledge.

[32] Rúben Gouveia, Evangelos Karapanos, and Marc Hassenzahl. 2015. How do we engage with activity trackers? A longitudinal study of Habito. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. 1305–1316.

[33] Jenny M Groarke and Michael J Hogan. 2016. Enhancing wellbeing: An emerging model of the adaptive functions of music listening. *Psychology of Music* 44, 4 (2016), 769–791.

[34] Jenny M Groarke and Michael J Hogan. 2019. Listening to self-chosen music regulates induced negative affect for both younger and older adults. *PLoS One* 14, 6 (2019), e0218017.

[35] Denise Grocke and Tony Wigram. 2006. *Receptive methods in music therapy: Techniques and clinical applications for music therapy clinicians, educators and students*. Jessica Kingsley Publishers.

[36] et al. Hair, Joseph F. 2009. *Multivariate Data Analysis: A Global Perspective. 7th ed.* Upper Saddle River: Prentice Hall.

[37] Susan Hallam, Ian Cross, and Michael Thaut. 2011. *Oxford handbook of music psychology*. Oxford University Press.

[38] Janet Harkness, Beth-Ellen Pennell, and Alisú Schoua-Glusberg. 2004. Survey questionnaire translation and assessment. *Methods for testing and evaluating survey questionnaires* (2004), 453–473.

[39] Donald A Hodges and Robin W Wilkins. 2015. How and why does music move us? Answers from psychology and neuroscience. *Music Educators Journal* 101, 4 (2015), 41–47.

[40] Daire Hooper, Joseph Coughlan, and Michael R Mullen. 2008. Structural equation modelling: Guidelines for determining model fit. *Electronic journal of business research methods* 6, 1 (2008), pp53–60.

[41] Yucheng Jin, Wanling Cai, Li Chen, Yuwan Dai, and Tonlin Jiang. 2023. Understanding Disclosure and Support for Youth Mental Health in Social Music Communities. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–32.

[42] Yucheng Jin, Wanling Cai, Li Chen, Nyi Nyi Htun, and Katrien Verbert. 2019. MusicBot: Evaluating critiquing-based music recommenders with conversational interaction. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 951–960.

[43] Yucheng Jin, Li Chen, Wanling Cai, and Pearl Pu. 2021. Key qualities of conversational recommender systems: From users' perspective. In *Proceedings of the 9th International Conference on Human-Agent Interaction*. 93–102.

[44] Patrik N Juslin, John A Sloboda, et al. 2001. Music and emotion. *Theory and research* (2001).

[45] Evangelos Karapanos, Jens Gerken, Jesper Kjeldskov, and Mikael B Skov. 2021. Introduction to "Advances in Longitudinal HCI Research". In *Advances in Longitudinal HCI Research*. Springer, 1–7.

[46] Christina Kelley, Bongshin Lee, and Lauren Wilcox. 2017. Self-tracking for mental wellness: understanding expert perspectives and student experiences. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 629–641.

[47] Rafal Kocielnik, Lillian Xiao, Daniel Avrahami, and Gary Hsieh. 2018. Reflection companion: a conversational system for engaging users in reflection on physical activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 1–26.

[48] Amanda E Krause, Adrian C North, and Lauren Y Hewitt. 2015. Music-listening in everyday life: Devices and choice. *Psychology of music* 43, 2 (2015), 155–170.

[49] Amanda E Krause, Adrian C North, and Lauren Y Hewitt. 2016. The role of location in everyday experiences of music. *Psychology of Popular Media Culture* 5, 3 (2016), 232.

[50] Kwangyoung Lee and Hwajung Hong. 2017. Designing for self-tracking of emotion and experience with tangible modality. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 465–475.

[51] Minha Lee, Sander Ackermans, Nena Van As, Hanwen Chang, Enzo Lucas, and Wijnand IJsselsteijn. 2019. Caring for Vincent: a chatbot for self-compassion. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[52] Younghwa Lee, Jintae Lee, and Yujong Hwang. 2015. Relating motivation to information and communication technology acceptance: Self-determination theory perspective. *Computers in Human Behavior* 51 (2015), 418–428.

[53] Yi-Chieh Lee, Naomi Yamashita, and Yun Huang. 2020. Designing a chatbot as a mediator for promoting deep self-disclosure to a real mental health professional. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–27.

[54] Yi-Chieh Lee, Naomi Yamashita, and Yun Huang. 2021. Exploring the Effects of Incorporating Human Experts to Deliver Journaling Guidance through a Chatbot. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–27.

[55] Yi-Chieh Lee, Naomi Yamashita, Yun Huang, and Wai Fu. 2020. " I Hear You, I Feel You": encouraging deep self-disclosure through a chatbot. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–12.

[56] Xue Li and Junlong Dong. 2021. User Demand Awareness and Analysis of Online Music——Take NetEase Cloud Music Platform as an Example. In *2021 5th International Conference on E-Education, E-Business and E-Technology*. 82–88.

[57] Zhaojiang Lin, Peng Xu, Genta Indra Winata, Farhad Bin Siddique, Zihan Liu, Jamin Shin, and Pascale Fung. 2020. Caire: An end-to-end empathetic chatbot. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 13622–13623.

[58] KuanTing Liu and Roger Andersson Reimer. 2008. Social playlist: enabling touch points and enriching ongoing relationships through collaborative mobile music listening. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. 403–406.

[59] Kai Lukoff, Taoxi Li, Yuan Zhuang, and Brian Y Lim. 2018. TableChat: mobile food journaling to facilitate family support for healthy eating. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–28.

[60] Aigul Mavletova. 2013. Data quality in PC and mobile web surveys. *Social Science Computer Review* 31, 6 (2013), 725–743.

[61] Kimberly Sena Moore. 2013. A systematic review on the neural effects of music on emotion regulation: implications for music therapy practice. *Journal of music therapy* 50, 3 (2013), 198–242.

[62] Alain Morin. 2011. Self-awareness part 1: Definition, measures, effects, functions, and antecedents. *Social and personality psychology compass* 5, 10 (2011), 807–823.

[63] Margaret E Morris, Qusai Kathawala, Todd K Leen, Ethan E Gorenstein, Farzin Guilak, William DeLeeuw, and Michael Labhard. 2010. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. *Journal of medical Internet research* 12, 2 (2010), e1371.

[64] Nika Mozafari, Welf H Weiger, and Maik Hammerschmidt. 2021. Resolving the chatbot disclosure dilemma: leveraging selective self-presentation to mitigate the negative effect of chatbot disclosure. In *Proceedings of the 54th Hawaii International Conference on System Sciences*. 2916.

[65] Daniel Müllensiefen, Bruno Gingras, Jason Musil, and Lauren Stewart. 2014. The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PloS one* 9, 2 (2014), e89642.

[66] Amrita Nair, Smriti Pillai, Ganga S Nair, and T Anjali. 2021. Emotion based music playlist recommendation system using interactive chatbot. In *2021 6th International Conference on Communication and Electronics Systems (ICCES)*. IEEE, 1767–1772.

[67] Carla Nave. 2017. Designing self-monitoring technologies for emotional self-awareness and wellbeing. *Electronic Visualisation and the Arts (EVA 2017)* (2017), 1–6.

[68] Johan YY Ng, Nikos Ntoumanis, Cecilie Thøgersen-Ntoumani, Edward L Deci, Richard M Ryan, Joan L Duda, and Geoffrey C Williams. 2012. Self-determination theory applied to health contexts: A meta-analysis. *Perspectives on Psychological Science* 7, 4 (2012), 325–340.

[69] Quynh N Nguyen and Anna Sidorova. 2018. Understanding user interactions with a chatbot: A self-determination theory approach. (2018).

[70] Stavros A Nikou and Anastasios A Economides. 2017. Mobile-Based Assessment: Integrating acceptance and motivational factors into a combined model of Self-Determination Theory and Technology Acceptance. *Computers in Human Behavior* 68 (2017), 83–95.

[71] Kathleen O'Leary, Stephen M Schueller, Jacob O Wobbrock, and Wanda Pratt. 2018. "Suddenly, we got to become therapists for each other" Designing Peer Support Chats for Mental Health. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.

[72] Zoe E Papinczak, Genevieve A Dingle, Stoyan R Stoyanov, Leanne Hides, and Oksana Zelenko. 2015. Young people's uses of music for well-being. *Journal of Youth Studies* 18, 9 (2015), 1119–1134.

[73] SoHyun Park, Jeewon Choi, Sungwoo Lee, Changhoon Oh, Changdai Kim, Soohyun La, Joonhwan Lee, Bongwon Suh, et al. 2019. Designing a chatbot for a brief motivational interview on stress management: Qualitative case study. *Journal of medical Internet research* 21, 4 (2019), e12231.

[74] SoHyun Park, Anja Thieme, Jeongyun Han, Sungwoo Lee, Wonjong Rhee, and Bongwon Suh. 2021. "I wrote as if I were telling a story to someone I knew.": Designing Chatbot Interactions for Expressive Writing in Mental Health. In *Designing Interactive Systems Conference 2021*. 926–941.

[75] Zhenhui Peng, Yunhwan Kwon, Jiaan Lu, Ziming Wu, and Xiaojuan Ma. 2019. Design and evaluation of service robot's proactivity in decision-making support process. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[76] Joonas A Pesonen. 2021. 'Are You OK?'Students' Trust in a Chatbot Providing Support Opportunities. In *International Conference on Human-Computer Interaction*. Springer, 199–215.

[77] Dorian Peters. 2022. Wellbeing supportive design–Research-based guidelines for supporting psychological wellbeing in user experience. *International Journal of Human–Computer Interaction* (2022), 1–13.

[78] Dorian Peters, Rafael A Calvo, and Richard M Ryan. 2018. Designing for motivation, engagement and wellbeing in digital experience. *Frontiers in psychology* 9 (2018), 797.

[79] Rafikatiwi Nur Pujiarti, Bumho Lee, and Mun Yong Yi. 2022. Enhancing User's Self-Disclosure through Chatbot's Co-Activity and Conversation Atmosphere Visualization. *International Journal of Human–Computer Interaction* (2022), 1–18.

[80] Bahlol Rahimi, Hamed Nadri, Hadi Lotfnezhad Afshar, and Toomas Timpka. 2018. A systematic review of the technology acceptance model in health informatics. *Applied clinical informatics* 9, 03 (2018), 604–634.

[81] Amon Rapp and Maurizio Tirassa. 2017. Know thyself: a theory of the self for personal informatics. *Human–Computer Interaction* 32, 5-6 (2017), 335–380.

[82] Abhilasha Ravichander and Alan W Black. 2018. An empirical study of self-disclosure in spoken dialogue systems. In *Proceedings of the 19th annual SIGdial meeting on discourse and dialogue*. 253–263.

[83] Faustine Régnier, Louis Chauvel, et al. 2018. Digital inequalities in the use of self-tracking diet and fitness apps: interview study on the influence of social, economic, and cultural factors. *JMIR mHealth and uHealth* 6, 4 (2018), e9189.

[84] Kelly Richards, C Campenni, and Janet Muse-Burke. 2010. Self-care and well-being in mental health professionals: The mediating effects of self-awareness and mindfulness. *Journal of Mental Health Counseling* 32, 3 (2010), 247–264.

[85] Nikki Rickard, Hussain-Abdulah Arjmand, David Bakker, Elizabeth Seabrook, et al. 2016. Development of a mobile phone app to support self-monitoring of emotional well-being: a mental health digital innovation. *JMIR mental health* 3, 4 (2016), e6202.

[86] Claudio Robazza, Cristina Macaluso, and Valentina D'Urso. 1994. Emotional reactions to music by gender, age, and expertise. *Perceptual and Motor skills* 79, 2 (1994), 939–944.

[87] James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.

[88] Richard M Ryan and Edward L Deci. 2017. *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publications.

[89] Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, et al. 2019. HCI and Affective Health: Taking stock of a decade of studies and charting future research directions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–17.

[90] Kyle-Althea Santos, Ethel Ong, and Ron Resurreccion. 2020. Therapist vibe: children's expressions of their emotions through storytelling with a chatbot. In *Proceedings of the interaction design and children conference*. 483–494.

[91] Thomas Schäfer, Peter Sedlmeier, Christine Städtler, and David Huron. 2013. The psychological functions of music listening. *Frontiers in psychology* 4 (2013), 511.

[92] Michael F Scheier. 1980. Effects of public and private self-consciousness on the public expression of personal beliefs. *Journal of personality and social psychology* 39, 3 (1980), 514.

[93] Jessica Schroeder, Chelsey Wilkes, Kael Rowan, Arturo Toledo, Ann Paradiso, Mary Czerwinski, Gloria Mark, and Marsha M Linehan. 2018. Pocket skills: A conversational mobile web app to support dialectical behavioral therapy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–15.

[94] Roni Shifriss, Ehud Bodner, and Yuval Palgi. 2015. When you're down and troubled: Views on the regulatory power of music. *Psychology of Music* 43, 6 (2015), 793–807.

[95] Marita Skjuve, Asbjørn Følstad, Knut Inge Fostervold, and Petter Bae Brandtzaeg. 2021. My chatbot companion-a study of human-chatbot relationships. *International Journal of Human-Computer Studies* 149 (2021), 102601.

[96] Wally Smith, Greg Wadley, Sarah Webber, Benjamin Tag, Vassilis Kostakos, Peter Koval, and James J Gross. 2022. Digital Emotion Regulation in Everyday Life. In *CHI Conference on Human Factors in Computing Systems*. 1–15.

[97] Martyn Standage, Joan L Duda, and Nikos Ntoumanis. 2005. A test of self-determination theory in school physical education. *British journal of educational psychology* 75, 3 (2005), 411–433.

[98] James P Stevens. 2012. *Applied multivariate statistics for the social sciences*. Routledge.

[99] Joanna Stewart, Sandra Garrido, Cherry Hense, and Katrina McFerran. 2019. Music use for mood regulation: Self-awareness and conscious listening choices in young people with tendencies to depression. *Frontiers in psychology* 10 (2019), 1199.

[100] Anna Sutton. 2016. Measuring the effects of self-awareness: Construction of the self-awareness outcomes questionnaire. *Europe's journal of psychology* 12, 4 (2016), 645.

[101] Ruth Tennant, Louise Hiller, Ruth Fishwick, Stephen Platt, Stephen Joseph, Scott Weich, Jane Parkinson, Jenny Secker, and Sarah Stewart-Brown. 2007. The Warwick-Edinburgh mental well-being scale (WEMWBS): development and UK

validation. *Health and Quality of life Outcomes* 5, 1 (2007), 1–13.

[102] Myriam V Thoma, Stefan Ryf, Changiz Mohiyeddini, Ulrike Ehlert, and Urs M Nater. 2012. Emotion regulation through listening to music in everyday situations. *Cognition & emotion* 26, 3 (2012), 550–560.

[103] Aditya Nrusimha Vaidyam, Hannah Wisniewski, John David Halamka, Matcheri S Kashavan, and John Blake Torous. 2019. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *The Canadian Journal of Psychiatry* 64, 7 (2019), 456–464.

[104] Daniel Västfjäll, Patrik N Juslin, and Terry Hartig. 2012. Music, subjective wellbeing, and health: The role of everyday emotions. *Music, health, and wellbeing* (2012), 405–423.

[105] Viswanath Venkatesh, Michael G Morris, Gordon B Davis, and Fred D Davis. 2003. User acceptance of information technology: Toward a unified view. *MIS quarterly* (2003), 425–478.

[106] Charles Welch, Allison Lahnala, Veronica Perez-Rosas, Siqi Shen, Sarah Seraj, Larry An, Kenneth Resnicow, James Pennebaker, and Rada Mihalcea. 2020. Expressive interviewing: A conversational system for coping with COVID-19. *arXiv preprint arXiv:2007.03819* (2020).

[107] Graham F Welch, Michele Biasutti, Jennifer MacRitchie, Gary E McPherson, and Evangelos Himonides. 2020. The impact of music on human development and well-being. , 1246 pages.

[108] Robin W Wilkins, Donald A Hodges, Paul J Laurienti, Matthew Steen, and Jonathan H Burdette. 2014. Network science and the effects of music preference on functional brain connectivity: from Beethoven to Eminem. *Scientific reports* 4, 1 (2014), 1–8.

[109] Nannan Xi and Juho Hamari. 2019. Does gamification satisfy needs? A study on the relationship between gamification features and intrinsic need satisfaction. *International Journal of Information Management* 46 (2019), 210–221.

[110] Xi Yang and Marco Aurisicchio. 2021. Designing conversational agents: A self-determination theory approach. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–16.

[111] Su-Fang Yeh, Meng-Hsin Wu, Tze-Yu Chen, Yen-Chun Lin, XiJing Chang, You-Hsuan Chiang, and Yung-Ju Chang. 2022. How to Guide Task-oriented Chatbot Users, and When: A Mixed-methods Study of Combinations of Chatbot Guidance Types and Timings. In *CHI Conference on Human Factors in Computing Systems*. 1–16.

[112] Jennifer Zamora. 2017. Rise of the chatbots: Finding a place for artificial intelligence in India and US. In *Proceedings of the 22nd international conference on intelligent user interfaces companion*. 109–112.