Changing Your Tune: Lessons for Using Music to Encourage Physical Activity

MATTHEW CLARK, University of Virginia, USA
AFSANEH DORYAB, University of Virginia, USA

Our research investigated whether music can communicate physical activity levels in daily life. Past studies have shown that simple musical tunes can provide wellness information, but no study has examined whether musical feedback can affect daily behavior or lead to healthier habits. We conducted a within-subject study with 62 participants over a period of 76 days, providing either musical or text-based feedback on their daily physical activity. The music was built and personalized based on participants' step counts and baseline wellness perceptions. Results showed that participants were marginally more active during the music feedback compared to their baseline period, and significantly more active compared to the text-based feedback ($p = 0.000$). We also find that the participant's average activity may influence the musical features they find most inspiration within a song. Finally, context influenced how musical feedback was interpreted, and specific musical features correlated with higher activity levels regardless of baseline perceptions. We discuss lessons learned for designing music-based feedback systems for health communication.

CCS Concepts: • **Human-centered computing → Auditory feedback**; **Ubiquitous and mobile computing design and evaluation methods**.

Additional Key Words and Phrases: Music, Sonification, Ambient Display, Health & Wellness, Physical Activity

ACM Reference Format:
Matthew Clark and Afsaneh Doryab. 2024. Changing Your Tune: Lessons for Using Music to Encourage Physical Activity. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 2, Article 79 (June 2024), 24 pages. https://doi.org/10.1145/3659611

1 INTRODUCTION

Personalized wellness technology aims to create distinctive yet individualized and efficient means of conveying wellness-related information. Various approaches offer insights into wellness using tangible mediums, such as illuminating elements [41], adaptive textiles [8], or garments that change colors [24]. Conversely, alternative methods harness existing user devices to present wellness data through ambient displays [12, 13, 19, 35, 37, 43], or gamification techniques [2, 20, 27, 57, 62], all geared towards promoting healthier lifestyles. It is important to note that, although these delivery methods have proven effective, they predominantly rely on the user's visual perception.

Research has shown that music has the potential to enhance user wellness significantly. Music therapy is a widely accepted approach for addressing complex psychological conditions like depression and anxiety and alleviating symptoms of challenging diseases such as Parkinson's disease, multiple sclerosis, and even cancer [46]. In addition to traditional music therapy, scientists have leveraged sonification technology to translate real-time

Authors' Contact Information: Matthew Clark, University of Virginia, Charlottesville, Virginia, USA, wdd6ts@virginia.edu; Afsaneh Doryab, ad4ks@virginia.edu, University of Virginia, Charlottesville, Virginia, USA.

ACM 2474-9567/2024/6-ART79
https://doi.org/10.1145/3659611

data into auditory cues. Sonification converts data into acoustic signals and has been used in various well-being contexts such as enhancing mindfulness [56] and alleviating pain in therapeutic settings [47]. Broader acoustic signals are frequently used to alter behavior, such as an alarm clock waking a user up. However, these interventions are typically designed for immediate relief and rely on real-time, relatively straightforward data. Nevertheless, given the success of previous wellness interventions, the effectiveness of music therapy, and the promising field of sonification, it is plausible that we can promote healthier lifestyles by conveying wellness information through music.

A recent survey-based investigation demonstrated that people can discern feelings of wellness through simple musical melodies [9]. In this study, participants assessed the perceived healthiness of customized musical compositions. Although this research offers compelling evidence regarding music's ability to convey information about wellbeing, there remains a need to explore how music could serve as a feedback mechanism for individuals to gauge their personal health status based on daily behavioral data.

In this paper, we take the first step toward this goal by exploring the feasibility of using musical feedback to convey one's daily physical activity status. We define musical feedback as the incorporation of an individual's data into a song. We conducted a 76-day within-subject study wherein participants were tasked to share their Fitbit data and fill out daily surveys that offered musical or textual feedback on their daily step count. Our system created tailored models based on each participant's initial perception of musical wellness, infusing wellness indicators within musical tunes to help participants evaluate their level of physical activity. We evaluate the step count recorded for each participant after getting their daily feedback and measure how their interpretation of the musical feedback aligns with the intended message. Additionally, we provide a qualitative assessment of participants' insights on receiving feedback through music. This paper's primary contribution lies in analyzing the effectiveness and challenges of using music as a feedback tool for daily physical activity.

Our analysis follows HCI guidelines for evaluating emerging behavior change strategies, focusing on understanding how the approach works in practice rather than solely demonstrating efficacy [31]. Based on our findings, musical feedback systems can promote physical activity, especially compared to text-based approaches. While personalization helps to communicate, basic musical models have the capacity to convey wellness information effectively, thus reducing the necessity for excessive customization. In these systems, as users become more familiar with musical feedback, their understanding and interpretation of the melodies will likely be refined. Yet, this learning period can potentially be reduced if systems are be mindful of how a user's current context might influence their interpretation of the music. Finally, future systems must consider the pervasiveness of music, and ensure the feedback is delivered in helpful and private ways. These insights suggest ways to enhance future musical feedback systems, making them more effective tools in behavior change technology.

2 RELATED WORK

In this work, we often use the terms "wellness" and "physical activity" interchangeably, but we recognize that they are two distinct concepts that are related. The term "wellness" is used in reference to previous literature [9], where the music models were created based on participants' perceptions of musical wellness. However, when real-world behavioral data is involved in music or equations, we use the term "physical activity" to make it clear that our study focuses on wellness in that context.

2.1 Wellness Technology

Technology can serve as a highly effective tool for encouraging healthier behaviors in users. These interventions can vary in complexity, with some systems deploy straightforward communication methods, such as sending text reminders [11], which are effective in addressing various health-related issues. Text reminders have helped participants with weight loss through constantly supplying information and promoting self-monitoring [50].

These reminders have also been used to provide distractions, support, and information about smoking, helping more participants to quit after 6 weeks of intervention [54]. Text systems have helped participants to better manage their diabetes, by promoting communication between patients and physicians [7]. Information sent via texts as infrequent as once a week has also helped participants to better control their asthma [49]. Some text-based reminders attempt to enhance engagement by incorporating elements like jokes and emojis [10]. Text feedback has been further developed, creating agents to actively engage users in a dialogue about their wellbeing. Reflection Companion [32] sent users prompts and follow-up questions encouraging users to reflect on their behavior. The study found this reflection encouraged users to remain engaged in the system, and them feel more motivated and empowered to be active. Other systems take a similarly straightforward approach of providing direct recommendations to encourage healthier behavior. MyBehavior [53] provided users with personalized physical activity and dietary recommendations. While receiving these recommendations, the studies found participants spent significantly more time walking. Although many of these systems offer simple reminders and recommendations, they have proven effective in altering behavior, even after the novelty has worn off [45].

Indirect intervention and persuasion strategies have also been developed to encourage healthier behavior. Gamificationseeks to enhance compliance by transforming behavior change into an engaging game [51]. Through gamification, these approaches seek to make behavior change enjoyable and something users eagerly anticipate, thus increasing the likelihood of sustaining the desired behavior modifications. Gamified approaches have also explored ways to promote healthier behaviors across various domains, including physical activity, dietary habits, personal hygiene, hand washing, and managing chronic illness [51]. One notable example of a gamified wellness application is Pokémon Go [2, 29], estimated to have added 144 billion steps to the United States over a single month [2]. Previous work has investigated gamified approaches specifically for children, taking characters from animated children's movies and altering the avatar's actions based on the child's detected activity level [57]. Other gamified approaches harness the power of competition among users, encouraging participants to vie with each other, allowing this competitive element to drive individuals toward healthier behaviors [27]. While participants report enjoying competing with one another through simple comparisons such as leaderboards [3], competition is most effective if participants compete with similar users [20] indicating even simple competitions benefit from a level of underlying personalization.

On the other hand, ambient displays promote healthier behavior by incorporating health data into metaphors that reward users for making healthier choices. These ambient displays have creatively portrayed health information in various ways, such as transforming phone home screens into flourishing gardens [12, 13], simulating aquatic ecosystems [35, 37], and utilizing text bubbles [19]. More recently, studies explored ambient displays that expand beyond one simple scene to tell a story. One such narrative-based ambient display is WhoIsZuki [42, 43], which visually tells a story of a cute alien adventuring to find his brother. The story progresses as users meet their physical activity goals, encouraging physical activity so participants can learn what happens to the characters. Other narrative-based approaches integrate real-world stories. StoryMap [55] allows families to post their own physical-activity-related stories, and read the stories posted by other users. While not perfectly an ambient display due to their textual nature, the results indicated these stories cause users to reflect on their own experiences, driving a sense of community that encourages users to be more active. In all, these ambient displays and narrative-based feedback systems show creative communication of wellness information allow the data to be more interesting, and possibly even foster emotional connections to the message.

In addition to these digital representations, tangible ambient displays have been developed, including lights and color-changing clothing to encourage breaks for physical activity [24, 41] and shape-changing fabric to symbolize breathing patterns [8]. In a multi-week long study, 3D printed artifacts have even been used to communicate physical activity data [30]. Users were presented with 5 printable objects ranging from a graph of their daily heart rate to a frog which changes size as the users are more active. The 3D printed artifacts served as useful conversation starters around the participant's activity, and even sparked a competition between participants to create the largest frog. Employing these metaphors allows technology to seamlessly blend into the user's environment, offering constant information that promotes healthier lifestyles.

2.2 Music & Sonification for Wellness

The methods mentioned earlier primarily rely users' visual perception. While vision is the most common mode for conveying health information, sounds and music have also demonstrated their effectiveness in communicating information and influencing behavior. The utilization of music and sonification in healthcare often leverages music's capacity to stimulate various brain functions, including cognitive and emotional processing [33]. This can be particularly beneficial for users, potentially offering advantages beyond traditional therapeutic approaches. However, nearly all existing health data sonification techniques are geared towards providing real-time feedback on immediate stimuli, while our approach constructs musical models of an individual's healthy behavior that can be played at a later time without relying on sonifying data in real time.

Research into sonification for wellness has been used to inspire users. Studies have investigated using musical scales, comprised of ascending and descending notes, to enhance confidence during physical therapy [59]. They have also utilized repeated recordings of waves and birds for mindfulness during meditation [56], ambient noise for mindfulness during walking [6], and sounds like wind, gears, and beeps to foster a sense of agency during exercise [36]. Music's ability to influence users' mental state is also evident in the effectiveness of music therapy, a field that employs music for patients' needs, through listening, performing, and composing [46]. Music therapy has demonstrated effectiveness in the treatment and mitigation of several diseases. Reviews of music therapy research found listening to music can help reduce anxiety regardless of context [60], individuals with autism spectrum disorder use music to overcome social barriers by better communicating their thoughts and feelings [21], and depression as a supplement to traditional treatment [40], among other challenging health concerns. Given these applications and its success in complex treatments, it is notable that sonification's applications have remained relatively limited.

Indeed, sonification's applications in promoting wellbeing often revolve around representing real-time physiological changes. In the context of physical therapy, these changes frequently emanate from muscle or body movements during rehabilitation [22, 48]. These types of sonification can communicate data through different ways, including deviating from the expected conclusion of a song representing the squats a participant is doing [47] or changing the music tempo while the participant is walking and lifting objects [23]. Studies have also sonified physiological signals, such as EEG [26, 52], pulse [28], and respiration [61]. These studies use synthesized and symphonic music, which effectively engages users despite the inherent complexity of translating muscle movements and internal physiological cues into auditory experiences.

It is somewhat surprising, given music's proven ability to encourage both physical movement and improved internal physiological states, that sonification has not been extensively explored as a means to promote everyday physical wellbeing. One possible explanation for this limited exploration could be the abstract nature of physical wellness compared to the concrete aspects of muscle movements and physiological signals. However, a recent study has demonstrated that individuals can interpret wellness information conveyed through simple musical melodies [9], albeit without conducting real-world testing of these melodies:the central focus of our current work.

3 STUDY DESIGN

To explore the feasibility of communicating wellness data through music, we conducted a 76-day within-subject study, during which participants received one month of musical feedback followed by one month of textual feedback on their activity levels or vice versa. Subsequently, we analyzed participants' behavior using data from Fitbit activity trackers and self-reported surveys to assess the effectiveness, shortcomings, and potential avenues for improvement in the practical application of musical feedback.

3.1 Study Procedure

Participants were recruited through email and word of mouth. All participants attended a 1.5-hour onboarding meeting with the research team upon enrollment. During this session, they received detailed information about the study, completed the music modeling surveys, and were provided with a Fitbit Sense device. All Fitbits were linked to an account managed by the research lab, ensuring anonymous data access. Following the onboarding process, participants were requested to submit a week's worth of Fitbit data without receiving any feedback. This initial data collection week established a baseline for each participant, supplying the essential background data necessary for the operation of our activity level calculation. Participants were given a two-week window to provide this data, allowing for potential technical issues and the adaptation period required for syncing their Fitbit devices. Failure to comply within this timeframe resulted in removal from the study.

After completing the initial onboarding period, participants received two months of feedback on their activity levels. In a randomly chosen order, the feedback was provided to them through 31 consecutive days of surveys. These surveys presented either a musical or textual representation of their activity level. Our approach for embedding activity level data into the music is explained in the following subsection (3.2).

Textual feedback was chosen as our control condition due to its ease of use and role as a foil to our experimental music condition. Feedback delivered by text is very clear, presenting one message with little to no interpretation. It is also widely used, with a literature review of mHealth behavior change techniques reporting that 47% of reviewed systems provide feedback through SMS or email [14]. Other systems for promoting physical activity are slightly less interpretable, such as ambient displays conveying information through metaphors. While less direct than blatantly stating the wellness information, users can still verify their impression by visually analyzing the display. Musical feedback, however, is even less interpretable than ambient displays. The musical feedback was intended to deliver an emotional message about the users' level of physical activity. The musical feedback is then intended to encourage the user to reflect on why their physical activity made the music feel a certain way, ideally creating a sense of intrinsic motivation. Thus, text was used as a baseline because it is simple for users to understand, and its blunt and direct nature contrasts with the music.

The feedback surveys were delivered to participants daily at 6 p.m. local time through email and were hosted on the Qualtrics platform. Within each daily survey, participants were first asked to assess their valence and arousal using a 5-point Likert scale, which aimed to gauge their emotional state. Following this, they received information about their physical activity level. In the case of the music surveys, participants were prompted to identify which activity level they believed the accompanying song correlated with, while in the textual representation surveys, participants were simply informed of their activity level. Upon receiving and completing 62 days of these daily surveys, participants were invited to participate in a final survey. This concluding survey asked participants to reflect on their overall experience and share their thoughts, opinions, and insights gained throughout the study.

3.2 Music Generation

To explore music as a method to communicate wellness status, we developed a procedure to convert a participant's data into music. Our approach consists of two primary steps. First, we generate personalized music models for individual users according to their perceived wellness of musical melodies at baseline, allowing the system to convey five distinct wellness levels. Then, we collect Fitbit data from users and apply an activity-evaluation algorithm to convert their steps into one of these five physical activity levels. The resulting physical activity level is then translated into music and emailed to the participants. This approach is visually illustrated in figure 1 and further explained in the subsequent sections.

3.2.1 Music Model Generation. The initial step in our approach involves creating music models for which we adopt a modified procedure as outlined in [9]. In this approach, a song is generated using Bach CoCoCo [38], and musical features are adjusted to communicate wellness levels. The adjusted musical features include tempo

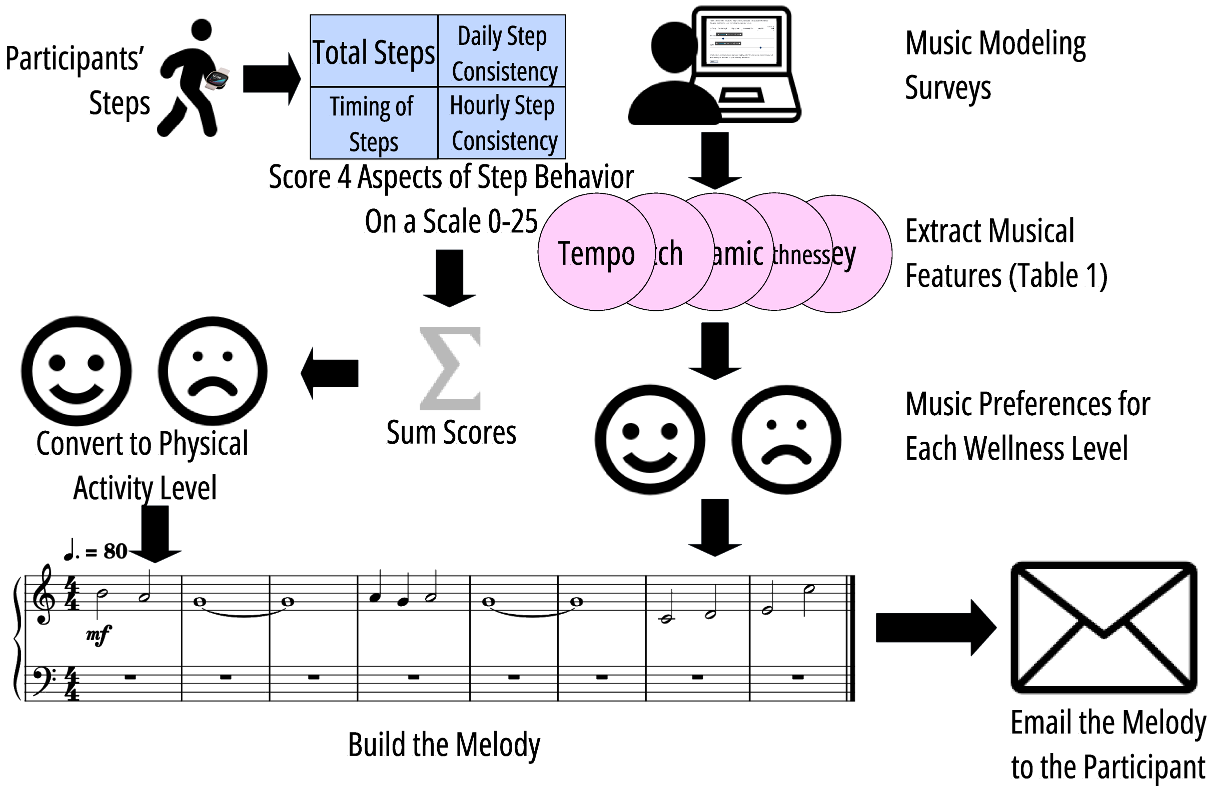

Fig. 1. Our approach to converting physiological data into sonified wellness information requires modeling participant's perceptions of wellness in music, and converting physiological data into physical activity levels. Creating the musical models follows the approach outlined in [9], where participants listen to music with altered features, and state how 'well' they sound. This is used to generate models for 5 distinct physical activity levels. We then collect Fitbit step data from participants, and score it to one of 5 physical activity levels. The assigned physical activity level is then converted to music, and sent to the participant.

(speed), pitch (note frequency), dynamic (volume), smoothness (interval between notes), and key (note variations). Musical and real-world explanations of each musical feature are provided in table 1. To generate the musical models, participants complete two surveys. In both surveys, participants listen to the same song repeatedly, with each iteration altering a single musical feature. In the initial survey, participants indicate their preferences for each musical feature concerning whether it sounds healthy or unhealthy to them (e.g., very fast tempo sounding healthy). In the second survey, participants can listen to all 32 combinations of these healthy and unhealthy features from the first survey and specify which combinations the model should employ to convey each of the five wellness levels.

3.2.2 Activity Level Calculation. The next step in our approach involves establishing a means to transform behavioral data into a corresponding physical activity level. We use Fitbit Senses [16] to gather data from our participants. Our algorithm and study are designed around participants' step data, specifically focusing on healthy daily and hourly step count thresholds recommended by literature and also calculated from an individual's historical data. The combination provides four numeric scores each ranging from 0 to 25 which are subsequently

Table 1. Musical feature values explained in musical and scientific terms. Tempo indicates the speed of the song, which we measure in beats per minute (BPM). Pitch denotes range of frequencies the song is in, and can be measured in hertz. Dynamics indicates the volume of the music. Key is a highly musical concept, which changes a few notes by a half-step. Smoothness denotes the amount of space between notes, and can be quantified as the duration of a note (in seconds). Table adapted from [9] with permission.

Feature Value	Scientific Measurement	Musical Notation
	Tempo	
Very Slow	40 BPM	Eighth Note = 80
Slow	60 BPM	Dotted Eighth Note = 80
Moderately Slow	80 BPM	Quarter Note = 80
Moderately Fast	120 BPM	Dotted Quarter Note = 80
Fast	160 BPM	Half Note = 80
Very Fast	240 BPM	Dotted Half Note = 80
	Pitch	
Very Low	65.41 - 123.47 Hz	C2
Low	130.81 - 146.94 Hz	C3
High	261.62 - 493.88 Hz	C4
Very High	523.25 - 987.77 Hz	C5
	Dynamics	
Quiet	63.4 dB	*Pianissimo*
Middle	71.5 dB	*Mezzo-Forte*
Loud	85.5 dB	*Fortissimo*
	Key	
Major	329.63, 440.00, & 493.88 Hz	C Major
Minor	311.13, 415.30 & 466.16 Hz	C Minor
	Smoothness	
Staccato	1.11 seconds	Staccato
Legato	2.22 seconds	Not Staccato

aggregated to generate a unified score between 0 and 100. Once designed, we needed to convert our 0-100 activity score to a 1-5 scale for the 5 musical wellness levels. To make the algorithm more accurate than just linearly mapping scores, we decided to base this conversion on the distribution of previously collected step data. During the previous academic year, we collected naturalistic Fitbit data from 158 college students, and used our algorithm to calculate 13,667 distinct activity scores. We then computed the quartiles of these scores, to serve as the bounds for each of the five physical activity levels.

The initial segment of our algorithm compared a participant's daily step count to the recommendations provided by the CDC [1]. A score of 0 was assigned if the participant took fewer than 4,000 steps, while a total of 25 points were granted for exceeding 8,000 steps. For step counts falling between 4,000 and 8,000, a proportional score between 0 and 25 was assigned.

The subsequent phase of our algorithm allotted points to participants based on the timing of their steps. Existing research suggests that intermittent walking breaks have favorable effects on both mental and physical health [4, 5]. However, the literature offers limited guidance regarding the optimal timing and duration of these breaks. We patterned this portion of the algorithm after Fitbit's recommendation, which suggests taking 250 steps for 9 hours during the day [17]. Participants earned 2.77 points (calculated as 25 divided by 9) for each hour they achieved this goal, up to a maximum of 25 points.

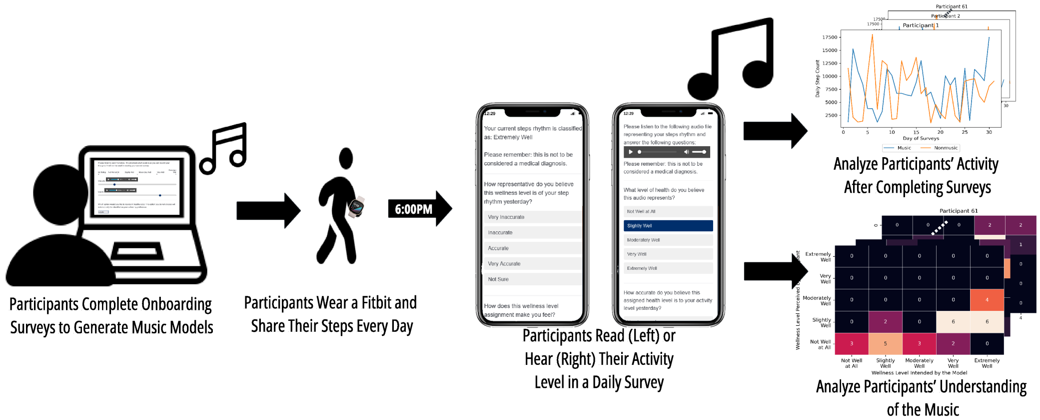

Fig. 2. Our study and approach to quantitative analysis. Participants began by completing onboarding surveys to generate their musical models. They then proceeded to wear a provided Fitbit for upwards of 76 days and complete daily surveys containing feedback on their physical activity. For analysis, we first seek to identify the relationships between music and physical activity. This analysis seeks to identify whether participants were more active after hearing their melodies, what musical features had the largest impact on activity, and whether baseline activity levels influenced what made music seem healthy. We then analyze how the melodies were interpreted. We analyze how frequently participants interpreted music to be the same physical activity level as the model intended to portray and identify contextual factors that influence these interpretations.

The third component of the algorithm rewarded participants for surpassing their usual daily step counts. It computed the average and standard deviation of their daily step counts over the past week. Participants received 0 points for falling one standard deviation below the average and a full 25 points for exceeding it.

The final component of our algorithm analyzed the timing of participants' steps, promoting consistent schedules that have been linked to improved stress levels [25] and depression [34, 44]. We evaluated each day using a sliding window of hourly segments, calculating the standard deviation for each hour of the day over the past week. Participants were awarded points for hours with a standard deviation under 500 steps, indicating adherence to a consistent physical activity schedule.

3.3 Participants

We enrolled 62 students from an American university for our research. Among them, 31 participants identified as female, 30 as male, and 1 as non-binary or third gender. The participants' average age was 22.7 years (σ = 3.8 years). Nearly all participants completed the study, with only three exceptions. One participant failed to meet the initial compliance criteria and was consequently excluded, while two participants withdrew during the study – one due to illness and the other for undisclosed reasons. As compensation, participants were allowed to retain their Fitbit devices or receive compensation up to $77. Our study was approved by the university's institutional review board.

4 ANALYSIS AND FINDINGS

The 61 participants who passed the compliance period completed 3,441 surveys during the study, including 1,728 music and 1,713 textual surveys. On average, each participant completed approximately 56.41 (σ = 10.23) surveys, out of which 28.33 (σ = 5.84) were music surveys, and 28.08 (σ = 5.90) were textual surveys. We concentrate on quantitative data and qualitative participant feedback to conduct our analysis. Our quantitative analysis, shown in figure 2, primarily centers on daily step counts recorded by participants' Fitbits and their interpretations of the music in relation to the intended message. Unless stated otherwise, we use linear mixed models from python's statsmodels package [58] for our statistical analysis. Linear mixed models were chosen because most tests include multiple data samples per participant, violating the random sample assumption of most common statistical tests. All linear mixed models in our analysis account for per-participant differences. On the qualitative front, we delve into participants' responses gathered through our final reflection survey. The analysis provides insight into how participants perceived the approach of conveying wellness information through music.

4.1 Physical Activity and Music

Participants averaged 8,700 ($\sigma = 3,635.74$) steps a day during the baseline period, 8,704 ($\sigma = 3,376.85$) steps after completing a textual survey, and 9,260 ($\sigma = 3,623.52$) daily steps after finishing a music survey. We further categorized participants into groups based on their physical activity level rather than evaluating the entire population to avoid overlooking individual variations. We segmented participants into three activity-level groups based on the CDC's recommended daily step counts [1]. Participants averaging fewer than 4,000 steps a day during the baseline period were classified as low activity (7 participants), those exceeding 8,000 steps were classified as high activity (31 participants), and individuals falling between 4,000 and 8,000 steps were considered middle activity (23 participants).

4.1.1 Participants Were Generally More Active After Musical Feedback Compared to Baseline, and Significantly More Active Compared to Textual Feedback. We first sought to investigate how participants' daily step counts changed after receiving each type of feedback. The mean and standard deviation of participants' steps are shown in figure 3A, and table 2. As can be seen, when ignoring per-participant differences in behavior, the average steps taken by all participants were highest under the musical condition. Looking more closely at the groups, while the combined low- and middle-activity participants also averaged the most steps while receiving musical feedback, the high-activity participants averaged the most under the baseline condition. Middle-activity participants showed increased steps compared to the baseline during both conditions and the most steps during the musical intervention. Low-activity participants took, on average, about a thousand more daily steps during both feedback conditions than the baseline period. However, they took the most steps during the textual feedback. As such, when analyzing the average steps participants took in each activity level group, we observed participants were more active after hearing musical feedback.

While we observe that middle-activity participants appeared to benefit the most from the musical feedback, it remains unclear why this group in particular would benefit the most. Intuitively, one would expect low-activity participants to most benefit, as they can improve the most. Analyzing the differences between the groups, the musical models for the middle-activity participants varied the intended wellness level more frequently than the models for low and high-activity participants. In other words, changes in a middle-activity participants' level of physical activity were more likely to be reflected by the musical models. For low-activity participants, the music communicated the same two wellness levels ("slightly well" and "moderately well") in 71.49% of melodies, indicating while the musical feedback did make these participants more active, the melodies were largely consistent throughout the entire study. Meanwhile, the high-activity participants received the wellness levels of "very well" and "extremely well" in 79.66% of melodies.These participants were already physically active and were likely to appreciate a well-composed melody without any major changes to their daily routine. However, the participants with moderate activity levels frequently heard melodies with different wellness levels, with each level appearing between 16.79% and 31.03% of melodies. Because the daily routines of the moderate activity participants were neither exceptionally healthy nor unhealthy to begin with, the changes in their daily behavior

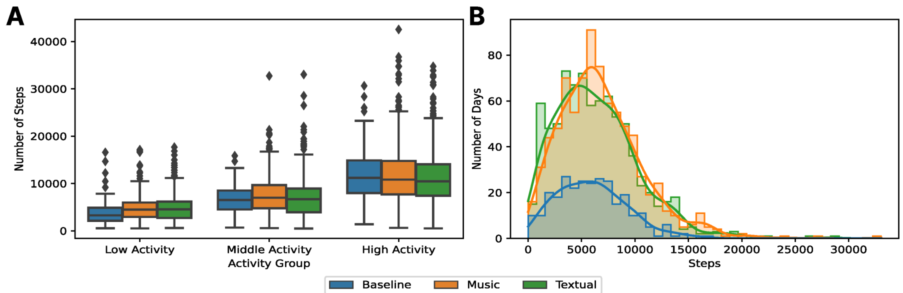

Fig. 3. The daily steps taken by participants during our study. The baseline was collected during the week where participants received no feedback. Music and text contain the steps taken the day after a participant completed each type of survey. A) The number of steps each activity group took for each condition. B) The frequency of daily step counts taken by low and middle-activity participants during each condition.

Table 2. The average number and standard deviation of daily steps taken by each activity group during each condition in the study.

	Baseline		Musical Feedback		Textual Feedback	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
High Activity	**11,702.88**	5306.72	11,686.48	5,939.21	11,160.03	5,528.69
Middle Activity	6,657.23	2,950.89	**7,415.99**	3,940.29	6,929.96	4,102.71
Low Activity	3,950.41	3,216.30	4,907.41	2,930.33	**5,127.54**	3510.77
All 3	8,700.17	3,635.74	**9,260.09**	3,623.52	8,704.38	3,376.85
Low & Middle	5,945.44	3,211.96	**6,786.88**	3,926.16	6,440.39	4,025.55

were better highlighted in the melodies. This potentially encouraged them to become more active, as they could hear the differences their behavior made in the music.

We also analyzed how participants' steps varied during the different conditions when accounting for personal differences in behavior. Using linear mixed models, we first compared each participant's steps during the baseline to those taken during their first feedback method. Given that participants' activity during the second condition may be influenced by their first month of feedback, we only compare a participant's activity during the first condition to their baseline steps. Although non-significant, we found participants took more steps after hearing musical feedback than the baseline, regardless of their baseline activity level ($p_{\text{All}} = 0.091$, $p_{\text{High}} = 0.077$, $p_{\text{Middle}} = 0.309$, $p_{\text{Low}} = 0.106$). Similarly, high-activity participants took more steps during the textual feedback condition than the baseline ($p = 0.811$). Interestingly, however, our models indicate many participants were actually less active after receiving textual feedback than during their baseline period ($p_{\text{All}} = 0.859$, $p_{\text{Middle}} = 0.672$, $p_{\text{Low}} = 0.758$). Although non-significant, these models further support our initial finding that participants were more active while receiving musical feedback compared to their baseline activity.

Finally, we compared the two feedback types to each other. We first identified that participants were significantly more active during the second feedback condition ($p = 0.047$), regardless of the order of conditions. This could indicate that the first condition influenced the second, or, since our study was run during the spring semester, participants were more active due to the improved weather. Due to this effect, we separated our models to consider participants based on the feedback they received first, generating linear mixed models for just the participants who received the music first, and another for those who received the text first. In line with the

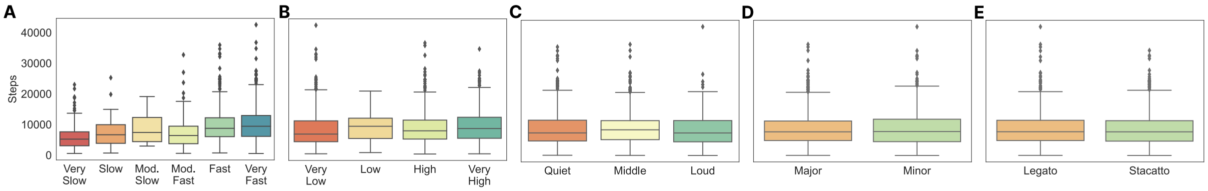

Fig. 4. The distribution of steps participants took the day after hearing every value for each musical feature. A) Tempo (speed) of the music. B) Pitch (frequency) of the music. C) Dynamics (volume) of the music. D) Key (exact notes) of the music. E) Smoothness (space between notes) of the music.

aforementioned effect, participants who started with the musical feedback showed a *non-significant* increase during the textual feedback ($p_{\text{All}} = 0.361$, $p_{\text{High}} = 0.861$, $p_{\text{Middle}} = 0.173$, $p_{\text{Low}} = 0.505$), indicating they may have marginally benefited from the textual feedback over the musical feedback. The participants who received the textual feedback first, however, showed a *significant* improvement during the musical feedback ($p_{\text{All}} = 0.000$, $p_{\text{High}} = 0.015$, $p_{\text{Middle}} = 0.002$, $p_{\text{Low}} = 0.000$). When considering that the models indicating textual feedback worked better were non-significant while those supporting musical feedback were significant, we can conclude that participants were generally more active during the musical feedback compared to textual.

4.1.2 Musical Features Correlate with Activity Level. Next, we wanted to determine whether the musical features correlated with participants' physical activity. For more information on the musical features, we refer the reader to table 1. We began by investigating a potential relationship between a participant's average activity and their baseline perceptions of musical wellness. This analysis could provide insights allowing for a small level of personalization without burdening the user to create the musical models explicitly. We collected each participant's well and unwell choices for each musical feature from onboarding and used ordinal regression to analyze whether the participant's average steps during the baseline week could indicate the participant's choices. We find participants' baseline activity appears to relate to one musical feature. While more active participants perceived faster melodies to be healthier, participants who took fewer steps during the baseline condition were more likely to perceive slower tempos as healthier ($p = 0.040$). This indicates a small level of personalization could be inferred from participant behavior, rather than explicitly their preference.

Since participants' perceptions of musical wellness correlated to their baseline activity, we further looked into identifying relationships between musical feature values by analyzing the number of steps participants took after hearing melodies with the various features. The distribution of steps taken after hearing each musical feature is shown in figure 4. Although non-significant, the models indicated potential trends. Specifically, participants were more active after hearing faster tempos ($p = 0.096$), higher pitches ($p = 0.547$), quieter dynamics ($p = 0.597$), minor keys ($p = 0.747$), and smoother melodies ($p = 0.224$). We then investigated the different activity level groups. Interestingly, all three groups appeared to be most active after hearing a different type of feature. High-activity participants' steps most correlated with the song's key ($p = 0.083$) and low-activity with pitch ($p = 0.205$). Meanwhile, middle-activity participants significantly correlated with the song's tempo ($p = 0.027$).

4.2 Interpreting the Music

We further analyze how participants interpreted the melodies' messages built from their activity data, and how in-the-wild contexts influenced their interpretation. Throughout this analysis, we focus on a phenomenon we call *alignment.* We define alignment as the relationship between the physical activity level interpreted by the participant compared to the activity level the model intended to communicate. If the participant interpreted the melody to contain the same activity level the model was trying to communicate, then we consider the user

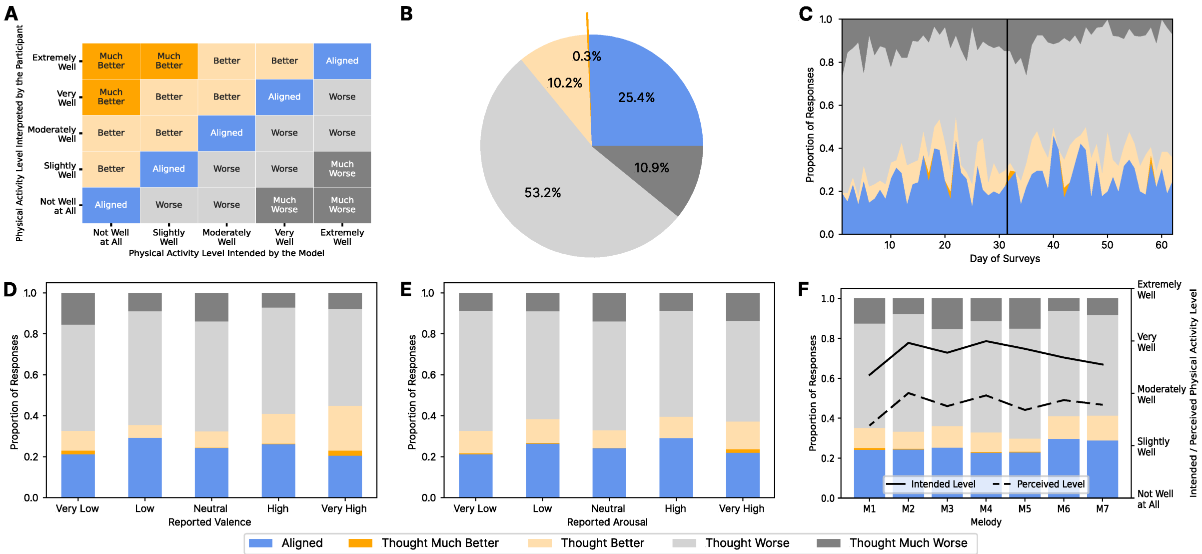

Fig. 5. The alignment findings in our study. A) A summary of how alignment terms are defined. B) The overall proportion of music survey responses falling into each alignment category. C) The proportion of responses falling into each type alignment grouping for each day of the study. D) Alignment sorted by self-reported valence. E) Alignment sorted by self-reported arousal. F) Alignment for each of the 7 songs that feedback was embedded in.

and model to be 'aligned.' The term 'thought better' indicates the participant interpreted the music sounded healthier than the model intended, while 'thought worse' means the opposite. Lastly, we use the terms 'much better' and 'much worse' to indicate this misalignment was off by at least two physical activity levels. These terms are visually summarized in figure 5A.

4.2.1 Activity Levels Encoded in Melodies Were Interpreted Worse Than Intended. We first analyzed the distribution of survey responses across different alignment categories, as illustrated in figure 5B. Participants' interpretations and the model's intended message were aligned in just 25.41% of survey responses, slightly above the chance alignment rate of 20%. Two-thirds of responses (64.06%) interpreted the melody as conveying a lower activity level than intended. Conversely, only 10.53% of responses reflected an interpretation of the music as better than intended. We will delve into this negative skew in participants' interpretations of musical feedback later as we discuss the potential influence of our algorithm design on music melody construction.

4.2.2 Music Interpretations Became More Aligned With the Model Over Time. Given the subjective nature of musical feedback, participants' interpretations can evolve over time, necessitating a 'recalibration' of the musical models. To assess this temporal evolution, we tracked the distribution of musical survey responses across alignment categories for each study day, as illustrated in Figure 5C. Using linear regression, we examined changes in the proportion of responses for each of the five alignment categories. Our analysis reveals that, as the study progressed, participants became more inclined to interpret the music as better than the model intended ($\beta = 0.001$, $p = 0.014$), while their likelihood of perceiving the music as much worse than intended decreased ($\beta = -0.002$, $p < 0.001$). In contrast, the remaining three alignment categories exhibited no significant changes over time ($p_{\text{Aligned}} = 0.105$, $p_{\text{Much Better}} = 0.688$, $p_{\text{Worse}} = 0.752$).

4.2.3 *Mood Effects How the Music is Interpreted.*

We next investigated how participants' emotional state affected how they interpreted the melodies. We analyzed the alignment between participants' activity interpretation of melodies and the model's intended activity level while considering participants' self-reported valence and arousal from each survey (figure 5D & E). We first analyzed the impact of valence on alignment using a linear mixed model, yielding a significant result ($p = 0.002$). As shown in figure 5D, participants tended to interpret the melody as better than intended when their valence was higher. In other words, participants experiencing positive emotions were more inclined to interpret the presented melodies as containing positive feedback than those experiencing negative emotions. Next, we delved into the influence of arousal on how the melodies were interpreted. Our model revealed that participants' interpretations did not significantly vary with arousal ($p = 0.428$). Consequently, as reflected in arousal, the intensity of emotions did not significantly affect how the melodies were perceived.

Our findings suggest that a user's positive or negative mood can alter their interpretation of melodies, with positive emotions leading to more favorable interpretations. However, the intensity of the emotion, as measured by arousal, does not seem to play a significant role in these interpretations. Moving forward, future musical feedback systems should either mitigate the influence of valence on users' interpretations or incorporate the user's current valence when generating music.

4.2.4 *The Song Itself & Day of Week May Effect How The Music is Interpreted.*

To maintain long-term engagement with the system and enhance compliance, it is beneficial to introduce variability into the underlying songs of our melodies. This approach may keep the system interesting for users over an extended period. To address this, we structured our study so that each day of the week featured a different song while maintaining consistent alterations in musical features to convey physical activity levels. This section investigates how our method of embedding wellness data in music can apply uniformly across different songs. The alignment for each musical composition is shown in figure 5F.

Our linear mixed model revealed no significant difference among the seven melodic tunes ($p = 0.073$). However, as illustrated in Figure 5F, M6, and M7 appeared to have higher correct alignment than the other five ($p = 0.000$). Interestingly, these songs were played on Saturday and Sunday, respectively. There were no differences between weekday melodies ($p = 0.074$) and weekend melodies ($p = 0.926$). While the difference in interpretation might be attributed to the choice of melodies, an alternative explanation could be that participants might have been in different environments or exhibited distinct behaviors during the weekend. We initially hypothesized that participants' weekend moods could differ, yet linear mixed models on self-reported valence ($p = 0.122$) and arousal ($p = 0.703$) showed no significant disparities between weekdays and weekends.

Our subsequent hypothesis considered that participants' interpretations remained consistent, but the activity levels intended by the model varied on weekends. We calculated each day's average intended and interpreted messages, as shown in Figure 5F. Linear mixed models indicated a correlation between weekends and lower intended activity levels ($p = 0.003$), yet no significant difference in the activity levels participants interpreted ($p = 0.218$). These tests support our final hypothesis that the shift in alignment during the weekend corresponds to a change in participants' physical activity rather than their interpretations.

4.3 Impressions of Musical Feedback

Upon concluding the study, each participant was invited to provide feedback through a final reflection survey. 55 of the 59 invited participants completed the survey. This survey posed a series of questions to capture participants' experiences during the study. Unlike the previous sections, which primarily centered on behavioral data and daily survey responses, the analysis in this section will exclusively concentrate on the insights gleaned from the responses to this concluding survey. The survey commenced by prompting participants to express their level of agreement with a series of statements using a Likert scale.

Subsequently, participants were presented with three open-ended questions that delved into their experiences with the melodies, such as how the melodies contributed to their motivation, what aspects they found appealing, and any aspects they found less appealing about this approach. Due to the small sample size, the qualitative responses were analyzed with manual encoding. To do so, all responses were placed in a spreadsheet, reviewed by the research team, summarized, and then grouped based on similar themes. After all groupings were made, the exact text of each review was reanalyzed to ensure it was placed in the appropriate group. This last step was repeated until an iteration when no responses were recategorized. The classifications were then assessed and confirmed by an independent member of the research team.

4.3.1 Participants Report Being Motivated by the Music but Not Much More than the Textual Condition. Participants responded to whether hearing the melody helped motivate them to be more conscious of their physical activity than reading their activity level using a 5-point Likert scale. In total, 58% of participants either agreed or strongly agreed that the melodies played a role in motivating them. However, only 22% indicated that the melodies were more motivating than simply reading their activity levels.

4.3.2 Participants Wanted to Receive Better-Sounding Music. Of the 55 participants, 30 (54%) reflected on how the melodies contributed to their motivation to be more conscious of their physical activity. Among the responses, 13 (24%) participants mentioned that the melodies motivated them by either avoiding ‘bad’ sounding melodies or seeking out ‘good’ melodies. For example, P21 stated, "Hearing a more sad melody motivated me to move around more. Over the study, I began walking more to get happier results." Five (9%) participants directly referenced specific musical features within the melodies as motivational factors. Some commented on the tempo, with three (5%) mentioning it, while one (2%) each mentioned pitch and key. These findings gather support for our findings in section 4.1.2, where we found weak, non-significant correlations between the musical features, with the strongest difference being present after hearing fast tempos. Between the qualitative and quantitative analysis, our results appear to indicate tempo had the strongest relationship with how participants’ activity would differ after hearing a melody. Meanwhile, two (4%) participants indicated that the open-ended nature of the melodies made them contemplate the intended activity level, keeping them engaged and curious. For instance, P10 said, "It kept me guessing what my result actually was." These views lend credence to our underlying theory that musical feedback’s open-ended nature would encourage the participants to reflect on their activity.

On the other hand, four participants (7%) expressed that the melodies did not inspire them. P41 stated, "Honestly, the melodies were highly confusing, and if you’re having a bad day, playing bad music to it just makes your mood worse." Some responses didn’t fit into the mentioned categories. P30 responded about the study, indicating that it made them more mindful of their sleep and activity levels. P47, dealing with chronic health issues during the study, expressed uncertainty about whether the melodies helped them.

While the melodies motivated many participants, future approaches may consider combining music and textual representations to cater to a broader range of preferences and motivations.

4.3.3 Participants Enjoyed the Musical Feedback. According to a 5-point Likert scale, 78% of participants reported enjoying hearing their activity levels through music melodies, while 14% disagreed, and 5% neither agreed nor disagreed. Only one participant (2%) strongly disagreed. These results indicate strong overall support for the concept of musical feedback for physical activity. Given that more than three-quarters of participants enjoyed it, and the majority had positive responses, these findings suggest that musical feedback for physical activity could be further developed and considered a viable option for promoting healthier behaviors.

4.3.4 Melodies Were Seen as Fun, Holistic, and Novel. Participants were also asked to explain what they liked about hearing melodies that encode their activity level information. All 55 responses answered this question. 21% of respondents (12 participants) expressed their enjoyment of the novelty surrounding the melodies and the overall concept. They articulated sentiments such as "Music wellness interpretations provide a holistic approach to health and well-being" (P7). Similarly, 12 participants (21%) highlighted their appreciation for the engaging and creative qualities of the melodies. They expressed that the music-based activity interpretations were more captivating and enjoyable than text-based representations. P28 remarked, "I felt like the music wellness interpretations were more interesting and fun than the text representation." Approximately 14% of participants commented on enjoying the melodies themselves. They relished the opportunity to listen to a brief melody each day, especially when they were aware of their high level of physical activity from the previous day. For example, P25 mentioned, "I liked getting to listen to a little melody every day, especially when I knew that I was very active the day before." It is important to note that if participants find the melodies to be fun and engaging, it could mean that musical feedback has the potential to keep users interested over time. This is particularly important for systems that aim to promote behavior change. However, it is important to conduct further qualitative analysis to determine how participants, who are not being compensated for each survey completed, interact with the system in order to fully prove this hypothesis. Another 8 (14%) noted that they derived pleasure from the emotional impact of the melodies. They appreciated that the melodies sounded cheerful on days when they were active and healthy, finding it to be a somewhat rewarding and mood-enhancing experience. P52 stated, "I liked that when I was being active and healthy, they sounded cheery. It was somewhat rewarding and made me happy." Participants noting the emotionality of the melodies is an additional sign that the musical feedback may encourage intrinsic motivation, as the participants will be driven to receive melodies that make them feel better about the feedback. Lastly, 5 participants (9%) expressed satisfaction with the melodies' personalization: "I liked how personalized they were. It was nice" (P34).

The remaining responses were categorized into four less common themes. Two participants (3%) expressed their satisfaction with the simplicity of understanding the melodies, with comments like, "I could instantly recognize the wellness level without having to think back on my day" (P4). Another 2 participants (3%) discussed the privacy aspect of the melodies, though their viewpoints contradicted each other. P41 noted, "They were very creative from a privacy perspective: Other onlookers likely had no idea what information was being communicated," while P44 mentioned, "I feel like if I would have shared the melodies with a friend, they would be able to understand what my health levels were." An additional 2 participants (3%) provided negative feedback about the melodies, offering concise comments such as, "I did not like them" (P5). The remaining four responses (7%) were considered off-topic. Among these, 3 participants (5%) commented positively on the wellness algorithm, with P8 stating, "I liked the measurement of my wellness." The final respondent, P1, (2%) expressed enjoyment in the reflective nature of the daily surveys, noting their pleasure in "daily [evaluation] of thyself while [filling out] the survey end of the day."

4.3.5 Participants were aware they were misinterpreting the melodies. A significant number, comprising 51% of the participants, found it challenging to interpret the melodies, often struggling to discern the intended physical activity level conveyed by the music: "It was sometimes challenging to figure out what the wellness level represented by the music was intended to be" (P55). Some participants preferred a more varied and fuller sound than the basic piano tones used in the melodies, constituting 16% of the responses. Another 9% mentioned difficulties in remembering the specific musical features they had associated with each activity level, which occasionally led to guesswork. This awareness matches the results of our quantitative analysis (section 4.2), participants struggled to interpret the exact message from the music. These reports, however, highlight that participants knew they were misaligned with the system. Future implementations could utilize this knowledge to receive feedback and correct the models until participants and the model can correctly communicate.

4.3.6 Participants expressed some concerns about the melodies. Some (11% of) participants reported not enjoying the emotional impact of the melodies, particularly when slower tempos were used, such as P13 saying, "It upset me to have slow tempos." Other participants, constituting 5% of the responses, expressed dissatisfaction with the music's open-ended nature. P41 summarizes this viewpoint, "I have to actively listen and interpret the music, and for health situations I have no desire to interpret what a doctor is telling me. I want them to just tell me so I can fix it as soon as possible." Two participants (3%) compared the musical interpretations and a textual format, with P24 commenting "It felt too vague and open to interpretation, I was never completely sure if I did well or not. I prefer the written feedback, as it is inarguable and quick and simple to interpret." P10, however, suggested the simultaneous use of both musical and textual formats. P16 (2%) voiced privacy concerns, stating discomfort with the idea that other people could hear the melodies. The feedback received suggests that although musical feedback could motivate some users to be more active, it should not be mandatory. As everyone has different preferences, some may prefer more straightforward approaches, while others may prefer the more traditional vision-based feedback.

Finally, some participants expressed off-topic concerns about the study, rather than the music itself. 3 participants (5%) focused on the algorithm's accuracy concerns, such as receiving unsettling tones after achieving certain activity levels or that the melodies were too finely categorized, conveying a sense of granularity. P14 (2%) indicated their sole complaint was related to the timing of the surveys, expressing dissatisfaction with the the time of day the surveys were sent.

5 DISCUSSION

5.1 Creativity in Systems Promoting Physical Activity

The most novel aspect of our system is the use of musical feedback to encourage users to be more active using their actual physical activity data. While the models have been previously proposed [9], applying them to encourage real-world behavior demonstrates the possibility of using music as a ubiquitous tool for delivering wellness feedback. Such ubiquitous delivery means musical systems could be used to communicate with the user without requiring them to be actively engaged with a device, potentially creating brand new ways and times to communicate with the user that might be meaningful, such as an alarm clock. While our system is not the first to consider using sounds to encourage physical activity [9, 36], it is the first to experiment with delivering personalized behavioral feedback rather than in-the-moment sonification of a live signal.

Many systems exist to encourage physical activity through creative approaches. We will briefly discuss several such systems and highlight how our findings relate to and differ from these other studies. A challenge with these comparisons, however, is the inconsistency in metrics to measure physical activity. Our study used a longitudinal approach and equipped every participant with their own Fitbit, providing a level of resolution in participants' activity that not every study reports. Other studies evaluate their results using metrics such as number of activity goals met [42, 43] or qualitative reports [30]. As such, direct comparisons of efficacy may be challenging, but it is still possible to compare the extraneous findings from each study, particularly a common trend of using creative feedback to foster intrinsic and extrinsic motivation for the purported benefits of each system.

Our study joins existing literature in demonstrating the benefits of creatively presenting feedback on physical activity through intrinsic motivation, fostering reflection and emotionality. Reflection Companion [32] encourages reflection through dialogue with a conversational agent. In section 4.3.2 of our qualitative analysis, we found that musical feedback has the potential to encourage reflection in users. This is because the message conveyed through music requires active interpretation by the user, as opposed to being told directly. Similarly, these systems also promote reflection by requiring users to reflect on their actions, which in turn encourages them to be more active. Some approaches motivate participants by eliciting specific emotions, like in WhoIsZuki [42, 43], where only those who meet their physical activity goals can help the main character find his brother. According to feedback from some of our participants, listening to certain melodies affected their emotions. This suggests that musical feedback could encourage engagement as people tend to be more active when they hear happier melodies. In this way, musical feedback is in line with existing research that has shown how creative communication can effectively promote wellness information and stimulate reflection and emotional engagement.

Another trend within creative approaches to physical activity feedback is generating a sense of community for extrinsic motivation. Gamified approaches often encourage users to compete with one another [20]. StoryMap encourages community members to share their stories [55], and 3D printed wellness artifacts sparked conversation and competition between participants [30]. Each of these systems, either intentionally or incidentally, fosters a sense of community around the feedback. Communities can serve as support networks, connecting individuals with similar goals to encourage physical activity. Our study does not capture this aspect of creative feedback. To avoid any potential issues in our study, we made sure that our participants did not communicate with each other. Although some participants may have recognized each other, we did not specifically investigate this aspect of our data. However, it is worth noting that one participant in the qualitative analysis (section 4.3.4) mentioned that they felt comfortable sharing the melodies with their friends. This could indicate, much like the 3D printed wellness artifacts, that with proper configuration, the artistic nature of the melodies could potentially foster community. If participants enjoyed listening enough, they could share the tunes with friends and family, creating a sense of community around the feedback. This aspect of creative feedback can be demonstrated in future studies.

Although creative wellness feedback may foster community, engagement and reflection, more traditional approaches still provide some benefits over creative feedback. Recommendation systems, such as MyBehavior [53], and text-based approaches [11] offer direct and clear suggestions and feedback on users' physical activity. As seen in section 4.3.6, some users may prefer this straightforward communication, not being asked to reflect on their well-being but simply being told how to be more active. Sometimes, direct interventions can be helpful for people who do not have enough knowledge on how to be physically active. In such cases, these interventions can provide necessary information to help users make a change, rather than relying on self-reflection to feel motivated and engaged. The creative and direct forms of feedback need not be mutually exclusive. Musical feedback can be used in conjunction with more straightforward reflection to provide both direct feedback and emotional data, thus combining the benefits of both types of systems.

5.2 Design Considerations

Our research marks an initial effort to explore the utilization of music as a means of providing users with feedback based on real-world activity data. The results of our study offer insights into the feasibility, constraints, and necessary directions for the development of musical feedback systems.

5.2.1 Musical Feedback for Behavior Change. Our findings in section 4.1.1 imply that music can serve as an effective feedback mechanism to motivate some participants to increase their physical activity. This could be integrated into health apps or systems for a more engaging user experience. The study showed that musical feedback improved behavior slightly over the baseline, and significantly over textual feedback. However, depending on their average activity, participants responded differently to different musical features. As supported by previous literature allowing participants to set their own fitness goals [42, 43] or indicating the importance of personalized communication [57], this suggests algorithms should factor in user-specific behaviors and preferences. Besides, while there is a clear preference for rewarding users with more pleasant melodies as a response to increased activity, less pleasant-sounding melodies can also serve as effective motivators. These less pleasant melodies can act as warning alarms, drawing users' attention to their inactivity and encouraging them to take action. Future systems can account for the type of feedback the developers want to use. Emphasizing negative-sounding melodies may encourage users to push themselves harder, while positive-sounding melodies may be more rewarding.

5.2.2 Personalization and Customization of Music Melodies. **??** Enhancing the melodies for users' enjoyment and engagement could be achieved by considering individual music tastes rather than individual perceptions of musical wellness. Currently, the approach generates all feedback in the classical musical style of Bach. However, given that our participants' average age indicates that less than 3% of the music they listen to is classical [15], expanding the modeling to encompass a broader range of musical styles may result in melodies that are more pleasant to users, potentially making them sound healthier. Future systems could automatically determine the style of music to deliver by analyzing users' music streaming history, such as on Spotify. This expansion into various musical genres can add appeal to the songs, rendering them more captivating, relatable, and personal to users. The importance of providing multiple communication methods is supported in previous literature [30], where participants used multiple types of 3D printed artifacts to visualize their wellness. Furthermore, factoring in participants' preferred musical styles might even obviate the necessity for specific musical features. For instance, a model could utilize songs aligned with an individual's musical taste to convey healthy behavior while employing songs the user dislikes to represent unhealthy behavior.

Moreover, our study delves into musical feedback as a distinct form of physical activity feedback. Nevertheless, in practical implementation, the musical delivery of feedback could be seamlessly integrated with existing, more traditional vision-based approaches such as ambient displays, textual and graphical representations, or gamified systems. In fact, a combination of approaches is suggested in previous literature [42]. This musical and visual feedback integration could create a synergistic relationship where each form complements the other's strengths. Incorporating musical feedback could aid visual feedback in conveying the emotional aspect of the message, enhancing overall communication. Simultaneously, visual feedback can mitigate the risk of misalignment in open-ended musical delivery, ensuring that users receive a more coherent and effective physical activity message.

5.2.3 Musical Models Can Be Simplified. As discussed in section 5.1.2, only tempo showed any significant correlation with the steps participants took after hearing a melody. While this indicates tempo may be able to directly influence the inspiration middle-activity users take from the music, it indicates the other features may not. Therefore, it is conceivable that participants do not consider the musical features at all, and future approaches may not need to rely upon them. Music inherently conveys emotions that transcend socio-cultural boundaries [18]. Embedding wellness data in the emotional quality of a song could make the feedback easier to interpret. Incorporating wellness data into the emotional elements of a song could make it easier to understand the feedback. Additionally, enhancing the emotional connection could improve the effectiveness of musical interventions. The success of WhoIsZuki is a testament to the power of emotional connection between the audience and the character [42, 43]. Instead of needing to analyze musical features, users would simply reflect on the emotion of the song. While this alteration might introduce a degree of subjectivity into the approach, it could also simplify the interpretation and internalization of the wellness message.

Additionally, in light of the overall low alignment, it is also plausible that the utilization of multiple musical features unnecessarily complicates the model. Future models might explore the possibility of relying on a single feature to convey wellness feedback effectively. For instance, a melody could embed wellness data using solely tempo, , which, in addition to encouraging users to be more active, has been found to improve the perceived wellness of music [9]. In this scenario, faster melodies could signify the participant performing well as positive reinforcement and encouraging continued physical activity. With this approach, users would only need to interpret the speed of the music rather than consider all five features. This streamlined approach will likely simplify comprehending the melodies and potentially enhance overall alignment. Designing systems around a single factor, such as emotionality or tempo, will make the music easier for users to understand while seemingly remaining effective in promoting physical activity.

5.2.4 Musical Feedback Interpretations May Require Time to Calibrate. In section 5.2.2, we identified that participants' interpretations of the music became more aligned over time. This may signify a growing familiarity with the model. Initially, at the outset of the study, participants may hear the melodies and perceive them as unpleasant, immediately associating this with unwell data. However, as they listen to the music over several days, participants better understand how the different levels of physical activity sound in the melodies. Consequently, this gradual acclimation refines their perceptions, bringing them closer to the intended message. Thus, future systems may need to account for this ’learning curve’ in interpreting melodies or devise strategies to expedite the calibration period. This calibration period presents a limitation with our proposed feedback method, as other more direct approaches do not require time for users to adjust to a ‘learning curve.’ Future work should consider investigating whether model simplification or better personalization can reduce the user’s burden in learning how to interpret the feedback. This calibration period can be eliminated by blending music and textual feedback to convey the message clearly while retaining emotional impact.

Future systems should be designed to account for potential calibration changes and lowered usage rates with prolonged use. As users become more familiar with the melodies, they may become less novel. This may result in weaker compliance from the users. However, our study cannot properly analyze this effect because our participants’ compensation was directly influenced by their compliance with using the system. It is possible that this effect could be mitigated by changing the underlying melody while consistently applying the same musical models to embed the message in the music. However, future systems should be aware that the time required to calibrate may simultaneously make users less interested in the musical feedback as the novelty effect wears off.

5.2.5 Musical Feedback Should be Context-Aware. Our findings show that the way melodies are perceived can be influenced by various contextual factors (sections 4.2.3 & 4.2.4). Our linear mixed models indicated that the time elapsed since generating the models, the participants’ valence and the day of the week significantly relates to how the melody is interpreted. Allowing the model the flexibility to adapt its behavior based on these contextual factors should enhance the consistency of alignment between musical feedback and users’ interpretations, ultimately leading to the improved overall performance of the musical feedback system. Context-aware computing has previously been used to improve music recommendation systems [39], indicating building models to account for contextual factors may improve the performance of our system. Context awareness is also essential for addressing privacy concerns, especially if musical feedback is audible to others. As observed in sections 4.3.4 & 4.3.6, some users were conscious of the music’s potential lack of privacy, and may hesitate to use the system if their wellness data is not adequately protected. Improved context-awareness can facilitate just-in-time active interventions by learning the user’s preferred time and setting for receiving musical feedback, such as when they return home from work, and identifying the most receptive time for delivering a message.

5.2.6 Musical Feedback Should Balance Accessibility, Usability, and Privacy. Using musical feedback for conveying health information can enhance the accessibility of personal wellness technology by expanding the means of communication beyond traditional visual displays. Developing these systems with simplicity in mind is essential to ensure they are accessible. Our analysis in section 4.1.2 suggests that one approach to ease the generation of musical models is to generalize musical features. We observed baseline activity correlates with participants’ perception of wellness from a melody’s tempo, which could be exploited to drastically reduce the burden of generating individualized musical models. Through this reduction of mental burden, the ease of using our system is improved, ideally making it more accessible to all users.

However, simplifying these models may raise privacy concerns. Given that auditory output is pervasive and can be heard by anyone nearby (if used without headsets), there is a risk that others nearby may perceive and interpret the wellness messages intended for the user. Therefore, exploring secure yet meaningful methods of delivering melodies is important. For instance, one potential approach could involve playing wellness melodies as the user’s morning alarm clock in a setting with minimal unintended listeners. Although previous research found that users may wish to share their data, and doing so can encourage them to be even more active [30, 55]. Applying this finding to musical feedback, it is possible that some users may not be worried about privacy, but rather want their musical feedback to be heard by friends and loved ones. Future systems must consider the appropriate deployment settings to ensure the melodies remain accessible, beneficial, and private while avoiding unintended disclosure of personal information.

5.3 Limitations and Future Directions

The largest limitation and perhaps most important future work for our approach of providing physical activity feedback through music is our inability to show behavior change at this time. At the current stage of musical feedback development, it is not possible to demonstrate behavior change within a sufficiently long duration. It is important to track how users adhere to and use the music system over time to study the long-term impact of the music models. However, our study did not have provisions for observing this metric as participants were paid based on the number of daily surveys they completed. Future studies exploring new musical approaches should consider different compensation plans that allow measuring the usage and compliance with the system over time.

Another potential limitation of our study is the possible influence of the participants' Fitbits. We chose Fitbits as an easy-to-use method to gather detailed step data, but these devices give users precise measurements of their daily step counts. It is possible that having this exact measurement could affect how participants perceive the music. However, we believe this is unlikely since the device only provides a number of steps, and participants probably already have an internal awareness of their activity level from the previous day. Although we cannot completely dismiss this as a possible confounding factor in our study.

Our study only contained musical and textual feedback. As stated in section 3.1, the reason for selecting textual feedback as a baseline condition was due to its directness, which is the opposite of the open-ended interpretations of musical feedback. However, we did not implement other physical activity encouragement methods developed by other studies in our work to compare their effectiveness against our musical system. Including more than two systems in a within-participant study structure would have required shortening the duration of each feedback type. This would have lessened our ability to observe the prolonged effects of musical feedback. Alternatively, the study could have been extended beyond one semester. However, this would change the quality of our data collection as participants may exhibit different activity levels during academic breaks. As such, we cannot directly compare the effectiveness of musical feedback to other approaches, such as narrative-based systems, using exact measures. To better understand the benefits and drawbacks of each approach, future work can investigate multiple novel approaches through direct comparison.

We have identified three potential future directions based on the findings of our study. Firstly, we discovered that middle activity participants benefited the most from the musical intervention. However, we could not find any clear demographic or background indicators that could explain why they were more receptive to the musical feedback. Therefore, future research can focus on defining the target population that can benefit the most from musical feedback. Secondly, our study hinted that musical feedback could be an effective tool for accessibility purposes, especially for visually impaired users. The auditory nature of musical feedback can communicate better with visually impaired users than vision-based approaches. This application of musical models can be explored further by accessibility researchers to optimize the models for visually impaired users. Lastly, our study found that, over time, participants' interpretations became better aligned with the model, which may indicate that they learned how to understand the feedback. Our implementation of musical models remained consistent, allowing participants to learn how to interpret the feedback. However, future studies could consider using an adaptive algorithm to implement musical feedback. This approach will allow the system to learn how to communicate with users while they simultaneously learn how to interpret the feedback.

6 CONCLUSION

We conducted an experiment to investigate the feasibility of using sonification to convey users' physical-activity levels through music. Our approach involved generating music based on the analysis of participants' daily step counts and then adjusting musical elements such as tempo, pitch, dynamics, key, and smoothness to communicate the intended physical-activity level. Our findings highlight opportunities to simplify the models, primarily by reducing or streamlining the musical features to align with general trends observed among participants. Despite the frequent misinterpretation of the model, we found that individuals were generally the most active while receiving musical feedback. This underscores the potential of music as a tool for encouraging healthier behaviors. We also found that the current melodies are highly influenced by contextual factors, such as the participant's mood. Our exploration into using music for physical-activity feedback provides valuable insights and design considerations for future iterations of these systems. It underscores the need for refining the music modeling approach, potentially by simplifying musical features and accounting for contextual factors that influence users' interpretations.

7 ACKNOWLEDGEMENTS

This work was funded by the National Science Foundation (NSF-IIS-1816687). We thank Andrew Balch, Runze Yan, and Matt Landers for their help with participant onboarding and Fitbit collection.

REFERENCES

[1] 2020. Higher Daily Step Count Linked with Lower All-cause Mortality | CDC Online Newsroom |. https://www.cdc.gov/media/releases/2020/p0324-daily-step-count.html

[2] Tim Althoff, Ryen W. White, and Eric Horvitz. 2016. Influence of Pokémon Go on Physical Activity: Study and Implications. *Journal of Medical Internet Research* 18, 12 (Dec 2016), e6759. https://doi.org/10.2196/jmir.6759

[3] Paul Barratt. 2017. Healthy competition: A qualitative study investigating persuasive technologies and the gamification of cycling. *Health Place* 46 (July 2017), 328–336. https://doi.org/10.1016/j.healthplace.2016.09.009

[4] Paolo Riccardo Brustio, Paolo Moisè, Danilo Marasso, Franco Miglio, Alberto Rainoldi, and Gennaro Boccia. 2018. Feasibility of implementing an outdoor walking break in Italian middle schools. *PLOS ONE* 13, 8 (08 2018), 1–13. https://doi.org/10.1371/journal.pone.0202091

[5] Sophie E. Carter, Richard Draijer, Sophie M. Holder, Louise Brown, Dick H. J. Thijssen, and Nicola D. Hopkins. 2018. Regular walking breaks prevent the decline in cerebral blood flow associated with prolonged sitting. *Journal of Applied Physiology* 125, 3 (Sep 2018), 790–798. https://doi.org/10.1152/japplphysiol.00310.2018

[6] Sixian Chen, John Bowers, and Abigail Durrant. 2015. "Ambient walk": a mobile application for mindful walking with sonification of biophysical data. In *Proceedings of the 2015 British HCI Conference*. ACM, Lincoln Lincolnshire United Kingdom, 315–315. https://doi.org/10.1145/2783446.2783630

[7] Jae-Hyoung Cho, Hye-Chung Lee, Dong-Jun Lim, Hyuk-Sang Kwon, and Kun-Ho Yoon. 2009. Mobile communication using a mobile phone with a glucometer for glucose control in Type 2 patients with diabetes: as effective as an Internet-based glucose monitoring system. *Journal of Telemedicine and Telecare* 15, 2 (Mar 2009), 77–82. https://doi.org/10.1258/jtt.2008.080412

[8] Kyung Yun Choi, Valentina Sumini, and Hiroshi Ishii. 2019. reSpire: Self-awareness and Interpersonal Connectedness through Shape-changing Fabric Display. In *Proceedings of the 2019 on Creativity and Cognition*. ACM, San Diego CA USA, 449–454. https://doi.org/10.1145/3325480.3329176

[9] Matthew Clark and Afsaneh Doryab. 2022. Sounds of Health: Using Personalized Sonification Models to Communicate Health Information. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (December 2022), 206:1–206:31. https://doi.org/10.1145/3570346

[10] Mihail Cocosila, Norm Archer, R. Brian Haynes, and Yufei Yuan. 2009. Can wireless text messaging improve adherence to preventive activities? Results of a randomised controlled trial. *International Journal of Medical Informatics* 78, 4 (Apr 2009), 230–238. https://doi.org/10.1016/j.ijmedinf.2008.07.011

[11] H. Cole-Lewis and T. Kershaw. 2010. Text Messaging as a Tool for Behavior Change in Disease Prevention and Management. *Epidemiologic Reviews* 32, 1 (Apr 2010), 56–69. https://doi.org/10.1093/epirev/mxq004

[12] Sunny Consolvo, Predrag Klasnja, David W. McDonald, Daniel Avrahami, Jon Froehlich, Louis LeGrand, Ryan Libby, Keith Mosher, and James A. Landay. 2008. Flowers or a robot army?: encouraging awareness activity with personal, mobile displays. In *Proceedings of the 10th international conference on Ubiquitous computing - UbiComp '08*. ACM Press, Seoul, Korea, 54. https://doi.org/10.1145/1409635.1409644

[13] Sunny Consolvo, Ryan Libby, Ian Smith, James A. Landay, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, and Louis LeGrand. 2008. Activity sensing in the wild: a field trial of ubifit garden. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. ACM Press, Florence, Italy, 1797. https://doi.org/10.1145/1357054.1357335

[14] Michelle Dugas, Guodong Gao, and Ritu Agarwal. 2020. Unpacking mHealth interventions: a systematic review of behavior change techniques used in randomized controlled trials assessing mHealth effectiveness. *Digital health* 6 (2020), 2055207620905411.

[15] Bruce Ferwerda, Marko Tkalcic, and Markus Schedl. 2017. Personality Traits and Music Genre Preferences: How Music Taste Varies Over Age Groups. *11th ACM Conference on Recommender Systems* (2017).
[16] Fitbit. 2023. Advanced Health Smartwatch | Fitbit Sense. https://www.fitbit.com/global/us/products/smartwatches/sense
[17] Fitbit. 2023. What are the reminders I see on my Fitbit device? https://help.fitbit.com/articles/en_US/Help_article/1878.htm
[18] Thomas Fritz, Sebastian Jentschke, Nathalie Gosselin, Daniela Sammler, Isabelle Peretz, Robert Turner, Angela D. Friederici, and Stefan Koelsch. 2009. Universal Recognition of Three Basic Emotions in Music. *Current Biology* 19, 7 (Apr 2009), 573–576. https://doi.org/10.1016/j.cub.2009.02.058
[19] Mads Frost, Afsaneh Doryab, Maria Faurholt-Jepsen, Lars Vedel Kessing, and Jakob E. Bardram. 2013. Supporting disease insight through data analysis: refinements of the monarca self-assessment system. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, Zurich Switzerland, 133–142. https://doi.org/10.1145/2493432.2493507
[20] Kyosuke Futami, Tsutomu Terada, and Masahiko Tsukamoto. 2021. A Method for Behavior Change Support by Controlling Psychological Effects on Walking Motivation Caused by Step Count Log Competition System. *Sensors* 21, 2323 (Jan 2021), 8016. https://doi.org/10.3390/s21238016
[21] Christian Gold, Tony Wigram, and Cochavit Elefant. 2006. Music therapy for autistic spectrum disorder. *Cochrane Database of Systematic Reviews* 2 (2006). https://doi.org/10.1002/14651858.CD004381.pub2
[22] J. Guerra, L. Smith, D. Vicinanza, B. Stubbs, N. Veronese, and G. Williams. 2020. The use of sonification for physiotherapy in human movement tasks: A scoping review. *Science Sports* 35, 3 (Jun 2020), 119–129. https://doi.org/10.1016/j.scispo.2019.12.004
[23] Jenny Hansdotter. 2017. Audio feedback in music - a study of experience of audio feedback in mu- sic for rehabilitation treatment for fear avoidance. (2017), 38.
[24] Emmi Harjuniemi, Ashley Colley, Piia Rytilahti, Hong Li, Jesse Forest, and Jonna Häkkilä. 2018. Idle stripes shirt: ambient wearable display for activity tracking. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers*. ACM, Singapore Singapore, 254–259. https://doi.org/10.1145/3267242.3267303
[25] Wai Kai Hou, Francisco TT Lai, Menachem Ben-Ezra, and Robin Goodwin. 2020. Regularizing daily routines for mental health during and after the COVID-19 pandemic. *Journal of Global Health* 10, 2 (Dec 2020), 020315. https://doi.org/10.7189/jogh.10.020315
[26] Vilelmini Kalampratsidou and Elizabeth B. Torres. 2019. Bodily Signals Entrainment in the Presence of Music. In *Proceedings of the 6th International Conference on Movement and Computing*. ACM, Tempe AZ USA, 1–8. https://doi.org/10.1145/3347122.3347125
[27] Dennis L. Kappen and Rita Orji. 2017. Gamified and persuasive systems as behavior change agents for health and wellness. *XRDS: Crossroads, The ACM Magazine for Students* 24, 1 (Sep 2017), 52–55. https://doi.org/10.1145/3123750
[28] Loic Kessous, Christian Jacquemin, and Jean-Julien Filatriau. 2008. REAL-TIME SONIFICATION OF PHYSIOLOGICAL DATA IN AN ARTISTIC PERFORMANCE CONTEXT. (2008).
[29] Madina Khamzina, Kaustubh V. Parab, Ruopeng An, Tiffany Bullard, and Diana S. Grigsby-Toussaint. 2020. Impact of Pokémon Go on Physical Activity: A Systematic Review and Meta-Analysis | Elsevier Enhanced Reader. *American Journal of Preventative Medicine* (2020), 13. https://doi.org/10.1016/j.amepre.2019.09.005
[30] Rohit Ashok Khot, Larissa Hjorth, and Florian "Floyd" Mueller. 2014. Understanding physical activity through 3D printed material artifacts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 3835–3844. https://doi.org/10.1145/2556288.2557144
[31] Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. 2011. How to evaluate technologies for health behavior change in HCI research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Vancouver BC Canada, 3063–3072. https://doi.org/10.1145/1978942.1979396
[32] 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 (July 2018), 70:1–70:26. https://doi.org/10.1145/3214273
[33] Stefan Koelsch. 2009. A Neuroscientific Perspective on Music Therapy. *Annals of the New York Academy of Sciences* 1169, 1 (Jul 2009), 374–384. https://doi.org/10.1111/j.1749-6632.2009.04592.x
[34] Francisco Tsz Tsun Lai, Vivien Kin Yi Chan, Tsz Wai Li, Xue Li, Stevan E Hobfoll, Tatia Mei-Chun Lee, and Wai Kai Hou. 2021. Disrupted daily routines mediate the socioeconomic gradient of depression amid public health crises: A repeated cross-sectional study. *Australian New Zealand Journal of Psychiatry* (Oct 2021), 000486742110512. https://doi.org/10.1177/00048674211051271
[35] Nicholas Lane, Mashfiqui Mohammod, Mu Lin, Xiaochao Yang, Hong Lu, Shahid Ali, Afsaneh Doryab, Ethan Berke, Tanzeem Choudhury, and Andrew Campbell. 2011. BeWell: A Smartphone Application to Monitor, Model and Promote Wellbeing. In *Proceedings of the 5th International ICST Conference on Pervasive Computing Technologies for Healthcare*. IEEE, Dublin, Republic of Ireland. https://doi.org/10.4108/icst.pervasivehealth.2011.246161
[36] Judith Ley-Flores, Laia Turmo Vidal, Nadia Berthouze, Aneesha Singh, Frédéric Bevilacqua, and Ana Tajadura-Jiménez. 2021. SoniBand: Understanding the Effects of Metaphorical Movement Sonifications on Body Perception and Physical Activity. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–16. https://doi.org/10.1145/3411764.3445558

[37] James J. Lin, Lena Mamykina, Silvia Lindtner, Gregory Delajoux, and Henry B. Strub. 2006. *Fish'n'Steps: Encouraging Physical Activity with an Interactive Computer Game*. Lecture Notes in Computer Science, Vol. 4206. Springer Berlin Heidelberg, Berlin, Heidelberg, 261–278. https://doi.org/10.1007/11853565_16

[38] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. 2020. Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 1–13. https://doi.org/10.1145/3313831.3376739

[39] Álvaro Lozano Murciego, Diego M. Jiménez-Bravo, Adrián Valera Román, Juan F. De Paz Santana, and María N. Moreno-García. 2021. Context-Aware Recommender Systems in the Music Domain: A Systematic Literature Review. *Electronics* 10, 13 (2021). https://doi.org/10.3390/electronics10131555

[40] Anna Maratos, Christian Gold, Xu Wang, and Mike Crawford. 2008. Music therapy for depression. *Cochrane Database of Systematic Reviews* 1 (2008). https://doi.org/10.1002/14651858.CD004517.pub2

[41] Victor Mateevitsi, Khairi Reda, Jason Leigh, and Andrew Johnson. 2014. The health bar: a persuasive ambient display to improve the office worker's well being. In *Proceedings of the 5th Augmented Human International Conference*. ACM, 1–2. https://doi.org/10.1145/2582051.2582072

[42] Elizabeth L. Murnane, Yekaterina S. Glazko, Jean Costa, Raymond Yao, Grace Zhao, Paula M. L. Moya, and James A. Landay. 2022. Narrative-Based Visual Feedback to Encourage Sustained Physical Activity: A Field Trial of the WhoIsZuki Mobile Health Platform. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (March 2022), 1–36. https://doi.org/10.1145/3580786

[43] Elizabeth L. Murnane, Xin Jiang, Anna Kong, Michelle Park, Weili Shi, Connor Soohoo, Luke Vink, Iris Xia, Xin Yu, John Yang-Sammataro, Grace Young, Jenny Zhi, Paula Moya, and James A. Landay. 2020. Designing Ambient Narrative-Based Interfaces to Reflect and Motivate Physical Activity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–14. https://doi.org/10.1145/3313831.3376478

[44] Greg Murray, John Gottlieb, and Holly A Swartz. 2021. Maintaining Daily Routines to Stabilize Mood: Theory, Data, and Potential Intervention for Circadian Consequences of COVID-19. *The Canadian Journal of Psychiatry* 66, 1 (Jan 2021), 9–13. https://doi.org/10.1177/0706743720957825

[45] Adity U. Mutsuddi and Kay Connelly. 2012. Text messages for encouraging physical activity Are they effective after the novelty effect wears off?. In *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*. 33–40. https://doi.org/10.4108/icst.pervasivehealth.2012.248715

[46] NCCIH. 2022. Music and Health: What You Need To Know. https://www.nccih.nih.gov/health/music-and-health-what-you-need-to-know

[47] Joseph W. Newbold, Nadia Bianchi-Berthouze, and Nicolas E. Gold. 2017. Musical Expectancy in Squat Sonification for People Who Struggle with Physical Activity. In *Proceedings of the 23rd International Conference on Auditory Display - ICAD 2017*. The International Community for Auditory Display, University Park Campus, 65–72. https://doi.org/10.21785/icad2017.008

[48] Thomas H. Nown, Priti Upadhyay, Andrew Kerr, Ivan Andonovic, Christos Tachtatzis, and Madeleine A. Grealy. 2022. A Mapping Review of Real-Time Movement Sonification Systems for Movement Rehabilitation. *IEEE Reviews in Biomedical Engineering* (2022), 1–16. https://doi.org/10.1109/RBME.2022.3187840

[49] Vedran Ostojic, Branimir Cvoriscec, Sanja Barsic Ostojic, Dimitry Reznikoff, Asja Stipic-Markovic, and Zdenko Tudjman. 2005. Improving Asthma Control Through Telemedicine: A Study of Short-Message Service. *Telemedicine and e-Health* 11, 1 (Feb 2005), 28–35. https://doi.org/10.1089/tmj.2005.11.28

[50] Kevin Patrick, Fred Raab, Marc Adams, Lindsay Dillon, Marion Zabinski, Cheryl Rock, William Griswold, and Gregory Norman. 2009. A Text Message-Based Intervention for Weight Loss: Randomized Controlled Trial. *Journal of Medical Internet Research* 11, 1 (Jan 2009), e1100. https://doi.org/10.2196/jmir.1100

[51] Pedro Pereira, Emília Duarte, Francisco Rebelo, and Paulo Noriega. 2014. *A Review of Gamification for Health-Related Contexts*. Lecture Notes in Computer Science, Vol. 8518. Springer International Publishing, Cham, 742–753. https://doi.org/10.1007/978-3-319-07626-3_70

[52] Daniël Ploeger. 2011. Sounds Like Superman? On the Representation of Bodies in Biosignal Performance. (2011), 13.

[53] Mashfiqui Rabbi, Min Hane Aung, Mi Zhang, and Tanzeem Choudhury. 2015. MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, Osaka Japan, 707–718. https://doi.org/10.1145/2750858.2805840

[54] A Rodgers. 2005. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. *Tobacco Control* 14, 4 (Aug 2005), 255–261. https://doi.org/10.1136/tc.2005.011577

[55] Herman Saksono, Carmen Castaneda-Sceppa, Jessica A. Hoffman, Magy Seif El-Nasr, and Andrea Parker. 2021. StoryMap: Using Social Modeling and Self-Modeling to Support Physical Activity Among Families of Low-SES Backgrounds. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–14. https://doi.org/10.1145/3411764.3445087

[56] Kavous Salehzadeh Niksirat, Chaklam Silpasuwanchai, Mahmoud Mohamed Hussien Ahmed, Peng Cheng, and Xiangshi Ren. 2017. A Framework for Interactive Mindfulness Meditation Using Attention-Regulation Process. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 2672–2684. https://doi.org/10.1145/3025453.3025914

[57] Hanna Schäfer, Joachim Bachner, Sebastian Pretscher, Georg Groh, and Yolanda Demetriou. 2018. Study on Motivating Physical Activity in Children with Personalized Gamified Feedback. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*. ACM, Singapore Singapore, 221–226. https://doi.org/10.1145/3213586.3225227

[58] Skipper Seabold and Josef Perktold. 2010. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*.

[59] Aneesha Singh, Stefano Piana, Davide Pollarolo, Gualtiero Volpe, Giovanna Varni, Ana Tajadura-Jiménez, Amanda CdeC Williams, Antonio Camurri, and Nadia Bianchi-Berthouze. 2016. Go-with-the-Flow: Tracking, Analysis and Sonification of Movement and Breathing to Build Confidence in Activity Despite Chronic Pain. *Human–Computer Interaction* 31, 3–4 (Jul 2016), 335–383. https://doi.org/10.1080/07370024.2015.1085310

[60] Judith H. Wakim, Stephanie Smith, and Cherry Guinn. 2010. The Efficacy of Music Therapy. *Journal of PeriAnesthesia Nursing* 25, 4 (Aug 2010), 226–232. https://doi.org/10.1016/j.jopan.2010.05.009

[61] Marcus Watson and Penelope Sanderson. 2004. Sonification Supports Eyes-Free Respiratory Monitoring and Task Time-Sharing. *Human Factors* (2004).

[62] Johann Weber, Mojdeh Azad, William Riggs, and Christopher R. Cherry. 2018. The convergence of smartphone apps, gamification and competition to increase cycling | Elsevier Enhanced Reader. *Transportation Research* (2018), 11. https://doi.org/10.1016/j.trf.2018.04.025