Unlock the Therapeutic Potential of Music: The BEATS Dataset

Benjamin Gutierrez Serafin*, Tanaya Guha*, and Fahim Kawsar*†
* University of Glasgow, † OmniBuds Ltd.

I. Introduction

Music holds broad therapeutic potential across diverse health and well-being applications, including mood and stress regulation, pain reduction, memory support, and improvements in neurological and mental health conditions [1]. Wearable sensing offers a promising route to scale and further enhance music-based interventions through biofeedback, allowing real-time adaptation to individual physiological states. However, the mechanisms underlying music's ability to modulate our physiological responses remain insufficiently understood, making it difficult to achieve reliable therapeutic outcomes.

A key factor constraining progress is the scarcity of large-scale datasets coupling physiological responses with auditory input [2], hindering the identification of consistent relationships and long-term physiological effects of music. Moreover, most existing datasets have been collected in controlled laboratory settings, limiting ecological validity and potential transferability to everyday listening scenarios.

To address these challenges, we developed a wearable sensing and music streaming system for in-the-wild data collection [3], laying the foundation for a unique resource. Building on this effort, we introduce the BEATS dataset, a longitudinal collection of raw physiological signals from music listeners in free-moving environments, enabling the study of music–physiology relationships in real-world settings.

II. Dataset Overview & Potential Applications

As illustrated in Figure 1, the BEATS dataset offers in-the-wild data from 30 participants collected over 21 days, combining continuous wearable sensing with self-paced listening to curated therapeutic playlists, self-reports of perceived effects, individual differences, and pre/post-experiment well-being measures. It includes 10,725 hours of raw physiological recordings from PPG, EDA, skin temperature, and triaxial accelerometry. Notably, a subset of 262.2 hours is aligned with music playback, corresponding to 844 listening sessions across 3 music listening contexts involving 117 curated tracks.

Beyond enabling the study of music as an affective stimulus, the comprehensive nature of the dataset supports the investigation of how temporally structured interventions can modulate autonomic and behavioural physiology in everyday settings. By integrating continuous physiological streams with temporally aligned acoustic descriptors, listening context, and self-reported outcomes, new avenues are opened for analysing patterns ranging from short-term physiological responses during individual sessions to longer-term regulation across repeated exposure over extended periods. Crucially, the inclusion of distinct listening contexts (affect regulation, eudaimonic and goal-attainment) in our study design enables music responses to be examined as context-dependent processes shaped by user intent. This contrasts with laboratory-based settings, where responses are typically treated as uniform reactions to externally imposed stimuli. Such a distinction is critical for therapeutic design, as physiological regulation is unlikely to depend solely on whether music is played, but on the interaction between the listener's current state, the acoustic structure of the stimulus, the intended listening goal, and the resulting physiological trajectory.

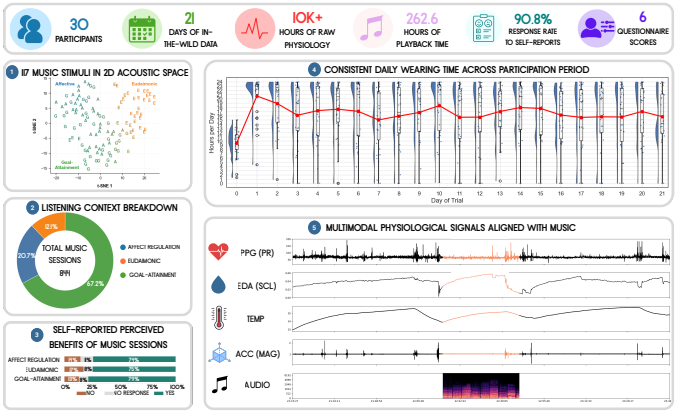

Fig. 1. Overview of the BEATS dataset, highlighting consistent wearable use, listening behaviour, and audio–physiology alignment in naturalistic settings.

Accordingly, the BEATS dataset is expected to provide the basis to address questions that were previously constrained by data limitations, including the identification of optimal moments for intervention delivery, the quantification of physiological regulation in response to music, and the identification of acoustic features associated with beneficial outcomes. While the experimental framework does not constitute a validated therapeutic intervention, it provides the empirical substrate required to model music–physiology dynamics and to inform the development of adaptive biofeedback-driven interventions for physiological regulation in real-world conditions.

III. Discussion & Future Work

As one of the largest datasets of its kind, the BEATS dataset represents a substantial contribution to research on the physiological effects of music. The dataset and associated codebase will be publicly available at *https://beats-dataset.github.io/*. Technical validation is underway, examining data quality through analyses that identify distinctive acoustic features across therapeutic contexts, alongside physiological trends indicating shifts towards individual baseline levels during music listening. Preliminary findings also suggest associations between specific sessions and physiological responses. Future work will leverage self-supervised learning for condition modelling (e.g., stress and mental fatigue detection) and for learning music–physiology dynamics via cross-modal alignment.

References

[1] D. Levitin, *Music as Medicine: How We Can Harness Its Therapeutic Power*. Random House, 2025.

[2] J. Kang and D. Herremans, "Are we there yet? a brief survey of music emotion prediction datasets, models and outstanding challenges," *IEEE Transactions on Affective Computing*, 2025.

[3] B. Gutierrez Serafin, T. Guha, and F. Kawsar, "Towards capturing the therapeutic effects of music in everyday life: Building the beats dataset," in *Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2025, pp. 1665–1669.