



# NeuralPMG: A Neural Polyphonic Music Generation System Based on Machine Learning Algorithms

Tommaso Colafiglio<sup>1,3</sup> · Carmelo Ardito<sup>2</sup> · Paolo Sorino<sup>1</sup> · Domenico Lofù<sup>1</sup> · Fabrizio Festa<sup>4</sup> · Tommaso Di Noia<sup>1</sup> · Eugenio Di Sciascio<sup>1</sup>

Received: 30 June 2023 / Accepted: 26 March 2024  
© The Author(s) 2024

## Abstract

The realm of music composition, augmented by technological advancements such as computers and related equipment, has undergone significant evolution since the 1970s. In the field algorithmic composition, however, the incorporation of artificial intelligence (AI) in sound generation and combination has been limited. Existing approaches predominantly emphasize sound synthesis techniques, with no music composition systems currently employing Nicolas Slonimsky's theoretical framework. This article introduces *NeuralPMG*, a computer-assisted polyphonic music generation framework based on a Leap Motion (LM) device, machine learning (ML) algorithms, and brain-computer interface (BCI). ML algorithms are employed to classify user's mental states into two categories: focused and relaxed. Interaction with the LM device allows users to define a melodic pattern, which is elaborated in conjunction with the user's mental state as detected by the BCI to generate polyphonic music. *NeuralPMG* was evaluated through a user study that involved 19 students of Electronic Music Laboratory at a music conservatory, all of whom are active in the music composition field. The study encompassed a comprehensive analysis of participant interaction with *NeuralPMG*. The compositions they created during the study were also evaluated by two domain experts who addressed their aesthetics, innovativeness, elaboration level, practical applicability, and emotional impact. The findings indicate that *NeuralPMG* represents a promising tool, offering a simplified and expedited approach to music composition, and thus represents a valuable contribution to the field of algorithmic music composition.

**Keywords** Machine learning · Human–machine interaction · Brain-computer interface · Leap motion · Music generation · Slonimsky's theory

## Introduction

In recent years, there has been a significant rise in the usage of technology in music composition, markedly advancing the field of algorithmic music composition (AMC). The roots of AMC can be traced back to the works of composers such as F. Bach Haydn and W. A. Mozart, who pioneered the creation of polyphonic structures through probabilistic games [1–3]. Nowadays, AMC demonstrates vast potentials in various areas, ranging from the application of cycle theories and permutations of musical intervals, to the use of

✉ Domenico Lofù  
domenico.lofu@poliba.it

Tommaso Colafiglio  
tommaso.colafiglio@poliba.it

Carmelo Ardito  
ardito@lum.it

Paolo Sorino  
paolo.sorino@poliba.it

Fabrizio Festa  
fabrizio.festa@conservatoriomatera.it

Tommaso Di Noia  
tommaso.dinoia@poliba.it

Eugenio Di Sciascio  
eugenio.disciascio@poliba.it

<sup>1</sup> Dipartimento di Ingegneria Elettrica e dell'Informazione (DEI), Politecnico di Bari, Bari, BA 70125, Italy

<sup>2</sup> Dipartimento di Ingegneria, Università LUM Giuseppe Degennaro, Casamassima, BA 70010, Italy

<sup>3</sup> Dipartimento di Ingegneria Informatica Automatica e Gestionale (DIAG), Sapienza Università di Roma, Rome, RM 00185, Italy

<sup>4</sup> Conservatorio di Musica E.R. Duni, Matera, MT 75100, Italy

stochastic processes in the generation of music [4]. Within the realms of harmony and composition, AMC techniques have proven invaluable for generating extensive collections of sound material [5].

Two notable technologies supporting AMC are brain-computer interfaces (BCIs) and Leap Motion (LM). BCIs are computer-based systems capable to acquire, analyze, and translate brain signals into real-time outputs, allowing users to communicate and control functions independently by muscle activity. The user's intentions are discerned from activity recorded by electrodes on the scalp or cortical surface. The term "BCI" was first used by Jacques Vidal, who in the 1970s developed a BCI system utilizing visual evoked potentials [6]. Since then, significant advancements in computer technology, artificial intelligence (AI), and neuroscience have facilitated the development and adoption, in particular in the research field [7] of diverse low-cost BCI systems, characterized by their affordability and quality of signal acquisition [8]. LM, on the other hand, is a sensor-based input device adept at tracking hand and finger movements in a three-dimensional space. It employs a blend of sophisticated optical sensors and algorithms, enabling users to interact with digital environments through intuitive gestures, thereby eliminating the need for traditional input devices like a mouse or keyboard [9]. Moreover, integrating machine learning (ML) techniques enhances the precision and effectiveness of brain-controlled applications. Its efficacy is evident in various BCI applications, with rapid advancements in ML prompting its extensive use in monitoring, detection, classification, and other tasks. These algorithms are vital for substantial progress toward human-level AI, especially in BCIs, where ML techniques play a critical role in data analysis, extracting valuable insights for targeted tasks [10].

Given the current limitations of BCI systems in comprehending complex brain activities, the integration of ML and BCI technologies promises a revolution in understanding of intricate brain signals and enhancing action recognition. BCI and ML technologies hold large potential for electronic and electroacoustic music composers, aiding in the exploration of vast sound textures and composition techniques.

Managing music materials remains a challenge for composers, for which N. Slonimsky's treatise "Thesaurus of Patterns and Melodic Scales" offers a promising solution. Published in 1947, it outlines computing permutations of notes around a reference interval, a concept explored by many composers throughout music theory and harmony history, with the primary goal of managing all the possible relationships between notes in a simple and autonomous manner [11]. This complex endeavor involves various music aspects, including note organization and rhythmic divisions. According to a large body of literature on AMC, it is widely recognized that music functions as a language interacting

with human emotions, yet it demands a solid foundation in music composition methods by the composer (see for example [12–16]).

This paper presents a system utilizing diverse information sources, such as electroencephalogram (EEG) signals and hand movement data. It encodes this information using ML algorithms for EEG signals and algorithmic processes based on Slonimsky's grammar for melodic profile and polyphony generation. Our work aims to support composers to create music polyphonies, assisting them in exploring all possible melodic patterns using Slonimsky's grammar, and artificial intelligence-based methodologies and tools, as well as LM and BCI devices.

The LM has been included in the system since each melodic profile can be represented as a succession of points in the two-dimensional space of the musical staff, with finger positions detected by LM corresponding to these points. Joining these points mimics the graphical representation of a musical melody.

For rhythmic value selection based on detected mental states, the BCI is used. Furthermore, the dynamics of each note in the polyphony are controlled by the power of Theta, Alpha, and Beta frequency bands, reflecting the user's mental conditions like activation, relaxation, and concentration [17].

Additionally, the proposed solution includes a graphical user interface (GUI) providing continuous visual feedback on the music generation process, including finger positions detected by the LM device and the user's mental state captured by the BCI. The GUI also displays the produced polyphony in real-time and allow parameters adjustments.

The overall framework, named *NeuralPMG*, extends beyond Slonimsky's treatise by incorporating polyphony generation techniques, thus contributing significantly to the state of the art about computer-assisted music composition.

An evaluation study was conducted addressing various perspective, including user experience (using the AttrakD-iff questionnaire), creativity support (Creative Support Index (CSI)), workload (Nasa-TLX), user engagement (User Engagement Scale (UES)), emotional response (Self-Assessment Manikin), participants' self-assessment, and appreciation of the generated polyphony. Additionally, two Maestros analyzed the aesthetics, innovativeness, elaboration, usage potential, and emotional capability of the polyphonies produced by participants. This article pioneers authors' research in using BCI, LM, and AI for AMC based on Slonimsky's theory, and, to the best of their knowledge, represents a novel contribution to the AMC field.

The article is organized as follows. Section "Example Scenario" describes typical usage scenario of the system. Section "Related Work" discusses the related work. Section "NeuralPMG Framework Architecture" describes the overall NeuralPMG framework architecture. Section "Polyphony Generation Process" illustrates the polyphony

generation process, and describes in detail: (i) mental state generation, (ii) melodic pattern generation, (iii) polyphony generation, and (iv) NeuralPMG Engine, including the implemented ML algorithms. Section “[Evaluation Study](#)” describes the evaluation study. Section “[Polyphony Assessment by Music Experts](#)” reports the assessment, performed by two Maestros, of the polyphonies produced by study participants involved in the study. Results of both the evaluation study and the Maestros’ assessment are discussed in “[Discussion](#)” section. Concluding remarks are provided in “[Conclusions and Future Work](#)” section. A “[Glossary](#)” explaining some terms specific of the music domain is available before the References section.

## Example Scenario

This section proposes a scenario depicting a key usage situation of NeuralPMG. Scenarios are commonly used in human-computer interaction for “bringing requirements to life” [18] since they provide, in a concrete narrative style, a description of specific people performing work activities in a specific work situation within a specific work context [19].

Tom is an expert music composer, who creates soundtracks for movies, advertisements, music videos, and audiovisual setups, as well as more traditional twentieth-century classical music. He frequently manages multiple composition projects in parallel and in different genres and styles, each taking from a week to several months to complete. He aspires to provide more complex compositions in the form of an audio and musical score, suitable for performance with virtual instruments or real music groups.

Tom needs to streamline the production process. He aims to obtain semifinished pieces that he can later refine, modify, or complete as desired. He is familiar with Nicolas Slonimsky’s theory, which enables him to experiment with various generations of scales and melodic patterns. He can construct polyphonies using sound organization grammar drawn from contemporary composition methods, such as the permutation of notes, inversions, transpositions of musical phrases, and variations in rhythmic values. Furthermore, Tom is well acquainted with the rules of polyphonic and counterpoint organization, which he uses to create tonal and post-tonal musical compositions.

Tom does not wish to rely solely on an automated music generation system, even if it can be programmed with specific parameters, as he wants to retain control over his own artistic decisions in each composition. Thus, Tom decides to use the NeuralPMG, whose GUI is shown in Fig. 1.

As first step, Tom wears a Emotiv headset, i.e., the BCI, and trains the system to recognize his mental state, so that it can determine if he is focused or relaxed. To accomplish this task, he selects the “Focused” and then the “Relaxed”

buttons in the “Mental states training” panel of the system interface. Once the system has been trained on Tom’s mental states, he can carry on using it. Tom activates the Leap Motion device to produce a melodic pattern: he moves his fingers on the device that detects the coordinates of Tom’s fingers and translates them into musical notes that are displayed on the musical staff. The finger position and the musical staff are both displayed in the “Melodic pattern creation” panel, on the left and central sides, respectively. The melodic pattern can be played and stopped. Once he has found a satisfactory melodic pattern, he can change and explore different versions of melodic patterns through the box “selecting Interval Axes in Semitones.” With this button, he can transpose and refit the melodic pattern on different interval axes. When Tom thinks he has found the right melodic pattern, he saves it as a MIDI or XML file. Using this melodic pattern as a seed, the system exploits the Slonimsky’s grammar to generate all the melodic patterns related to that seed. Tom then goes on with “Polyphony Generation.” The trained classifier identifies Tom’s mental state and generates a polyphonic composition. Tom knows that when he is focused, the polyphony has rhythmic values of quarter note, eighth note, and sixteenth note, while when he is relaxed, the rhythmic values are whole note, half note, and quarter note. Thus, depending on the polyphony he wants to generate, Tom tries to be focused or relaxed. Tom can modify the generation intervals, the “performance tempo” expressed in bits per minute (*bpm*) and the interval of each polyphonic voice according to what he feels for each composition. When he finds the generated polyphony useful, he saves and exports it in a MIDI or XML file format.

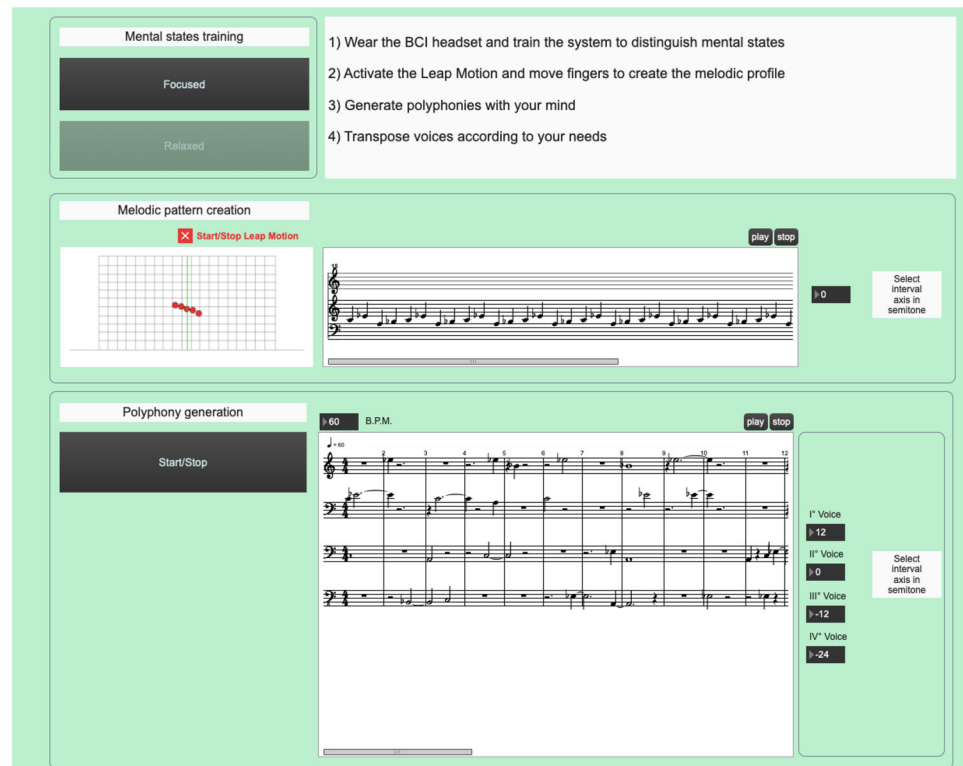
Later, having an archive of polyphonic and melodic compositions at his disposal, he can edit, arrange, and manipulate them as desired using either a digital audio workstation (DAW) or a music notation software.

## Related Work

In the current literature, the use of BCI for music generation is a relatively unexplored area. This section addresses the related work that informed our research.

The first recording of the EEG signal was made by the German psychiatrist Hans Berger in 1929 and was a historical breakthrough providing a new neurologic and psychiatric diagnostic tool [20]. The EEG process involves measuring the electrical activity of active neurons in the brain. This data must be filtered to discern various frequencies that serve distinct functions. Since the late 1960s, EEG signals have found applications beyond the medical field. Notably, BCIs have expanded their use to monitor brain activity (i.e., EEG) and control computers across various sectors beyond neuroscience and medicine. These include areas like video games,

**Fig. 1** GUI of the NeuralPMG framework: three panels outline the main steps in the process of creating the final polyphony, i.e., *Mental states training*, *Melodic pattern creation*, *Polyphony generation*. Instructions for using the system can be found in the panel at the top



media art, music, and several other fields [21]. In the preliminary stages of biofeedback research, musicians and scientists have used EEG technology to make sound, but mostly “passively generated sound” based on brain activity. One of the primary limitations was the believing that the conscious state of relaxation was directly linked to the value of Alpha band waves. Consequently, to induce enhancement of the Alpha band, a certain level of mastery over one’s consciousness was necessary to attain and sustain the Alpha state. Regrettably, early on, it was discovered that the Alpha state could be easily disrupted by even minor motor, visual, or mental exertions. As a result, music created using biofeedback techniques often turned out to be passive, drone-like, and repetitive. This led to the short-lived popularity of biofeedback music [22]. Experiments such as Cage’s ones, where an EEG electrode was directly connected to an audio amplifier, did not yield appreciable results from a compositional standpoint, despite their inherent interest [23].

David Rosenboom in his book *Biofeedback and the Arts, Results of Early Experiments* describes several attempts of using EEG in artistic performances [24]. Lloyd Gilden’s 1968 article, titled “Instrumental Control of EEG Alpha Activity with Sensory Feedback,” provides descriptions of the experience of participants involved in various alpha-wave experiments [25]. For many individuals, achieving relaxation state required dedicated practice and self-discipline, much like the mental state one enters during Zen meditation (as exemplified by Pauline Oliveros’ “Sonic Meditations” [26]).

Sanyal et al. [27] shows an effective way to correlate users’ emotional states with auditory stimulation related to a specific type of Indian music. The authors highlight that by exploiting Multifractal Detrend Cross-Correlation Analysis (MFDXA), it is possible to enhance the electrical activity’s sound more explicitly in response to a specific external auditory stimulus, resulting in an aroused brain state.

Quite notable are the many pieces that Teitelbaum realized between 1966 and 1974 using psychophysical feedback mixed with a variety of East spiritual disciplines and rituals. Most pieces described by Teitelbaum required the performer to strap on various EEG/ECG sensors and contact microphones, which were used to amplify the performer’s heartbeat and breathing. Many of the participants in these pieces trained in consciousness and physical awareness disciplines such as yoga and Zen meditation as a means of honing their performance skills (see [28]).

Alvin Lucier was a pioneer in using EEG as a source for music generation [29]. As described in [30], in Lucier’s “Music for Solo Performer” seminal work the composer employed the amplitude modulation of Alpha waves, precisely captured using two electrodes positioned frontal, to interact with percussion instruments. This approach involved harnessing the amplitude variations of Alpha waves to facilitate the manipulation of the percussion instrument’s surface through the utilization of loudspeaker energy. Direct cognitive control over the amplitude of these waves looks to be exercised by the composer, thereby assuming an active role

in steering the real-time generative course of the compositional process. In the analysis of Lucier's performances, as evidenced in available videos<sup>1</sup> a notable behavior is observed: Lucier frequently closes his eyes. This action is relevant because closing one's eyes is known to physiologically increase Alpha band power, a process independent of cognitive intention. Additionally, Lucier's habit of touching his eyes introduces artifacts into the EEG signal, further influencing the output. Moreover, he actively modulates the dynamics of the performance by adjusting the signal gain via a hand-operated control on the EEG amplifier, which is in turn connected to the speaker amplifier. This observation leads to the conclusion that Lucier's system represents an early example of the sonification of natural alpha wave trains, which naturally occur in the human brain upon eye closure.

In 1990, Knapp et al. developed a system called BioMuse, a musical instrument based on physiological biosensing technology [31]. A series of electrodes detect cerebral activity (EEG), cardioid activity (ECG), and muscular activity (EMG), which are sensed and digitized, and become human interface data for commanding specific computer operations.

Only later it emerged that to control the state of relaxation it is necessary to consider the ratio between Alpha and Delta waves, and for concentration that between Beta and Theta: "EEG-controlled musical instruments" have appeared.

Brouse and Miranda pioneered the field of algorithmic musical composition by reprocessing EEG signals through sophisticated programs. These programs are designed to extract pertinent information from the EEG data, such as fluctuations and spikes, employing a range of mathematical and algorithmic operations to interpret these signals musically. A notable example of this innovative approach is Brouse's "InterHarmonium" introduced in 2001. This groundbreaking system functioned as a "network of brains," where multiple composers cooperatively engaged in the creation of a polyphonic musical piece. In this setup, the BCI captures EEG signals from each participant. These signals are then transmitted over the Internet using the UDP protocol to a central server. On this server, the incoming EEG data from each interconnected composer is processed and integrated to produce a cohesive polyphonic composition. This method effectively combines the neural activities of various individuals into a unified musical expression, showcasing a novel intersection of neuroscience, technology, and art [32]. As described in [33], the authors developed the BCMI-Piano, i.e., Brain-Computer Music Interfaces Piano, a music generation system through an EEG signal processing chain. In BCMI-Piano, some specific features are extracted from the raw signal and used to drive music generation algorithms. The creation of Tempo and Dynamics is related to the characteristics of Hjorth Activity, Mobility, and Complexity. Pitch

organization is related to the frequency band domain. The Midi protocol enables communication between various electronic musical instruments based on these characteristics.

Hamadicharef et al. introduced *I2R NeuroComm*, a BCI system for music composition [34]. This system allows users to compose short and simple melodies by inserting or deleting notes in a musical partition, and playing them, all controlled through brain waves based on the P300 paradigm [35]. The focus, and primary contribution of their work, was on the GUI and its components, which enables users to create musical compositions by displaying the melody as text and as a real musical partition and includes some functions for modifying it on the fly.

Folgieri and Zichella proposed a low-cost BCI approach that enables users to consciously produce specific musical notes using their brainwaves [36]. This method integrates audio, gesture, and visual stimuli. Specifically, the application collects information with the following characteristics: (i) users listen to each note only once; (ii) before reproducing a note, the program instructs the user to associate a simple gesture with it, suggesting different gestures for different notes; (iii) while listening to each note, the software displays an associated image—the name of the note on a unique, note-specific colored background—aiding users in playing the note using EEG signals.

Pinegger et al. proposed Brain Composing, a BCI-controlled music composing software [37]. They evaluated this system with five volunteers, noting enhanced usability due to a tap water-based electrode bio-signal amplifier. Remarkably, three of the five subjects achieved accuracies above 77% and successfully copied and composed a given melody. The positive questionnaire results indicate that the Brain Composing system offers an attractive and user-friendly approach to music composition via BCI. This system employs a P300-based approach with a shrinkage linear discriminant analysis (sLDA) classifier.

Deuel et al. in 2017 presented the Encephalophone, a BCI with the double-edged purpose of exploring new frontiers in music technology and as a possible therapeutic tool for people who had suffered from strokes or neurological problems like ALS [38].

Ehrlich et al. [39] presented a functional prototype algorithm integrated into a BCI architecture, and able to generate real-time continuous and controllable patterns of affective music synthesis. The evaluation comprised two distinct studies: the first addressed the affective quality of synthesized musical patterns produced by the automated music generation system. The second one explored affective closed-loop interactions across multiple participants. Results highlighted that participants demonstrated the ability to deliberately modulate musical feedback by self-inducing emotions.

The Leap Motion (LM) controller plays a crucial role in the NeuralPMG architecture. This device has gained

<sup>1</sup> see for example: <https://www.youtube.com/watch?v=bIPU2ynqy2Y>.



recognition in various studies for its utility in managing sound parameters through hand movements. One of its notable applications is in controlling virtual instrument players, such as Native Instrument's Kontakt Player4, where it has yielded satisfactory results. The LM, in conjunction with Kontakt Player, enables users to manipulate aspects like velocity (dynamics) and key switches (changing articulation for virtual instruments). For pitch control, more standard input methods like keyboards or Midi guitars are typically employed.

In the context of commercial music production or performance, the LM controller is instrumental in adjusting elements such as filters, equalizers, instrumental loops, and performance pitch. These functionalities are often executed through the Gecko Midi application,<sup>2</sup> which facilitates the mapping of various hand gestures to Midi parameters, allowing for a high degree of customization. A notable application of the LM controller is found in the work of Croassacipto et al. [40], who developed a height recognition system using the LM and the K-nearest neighbor classifier. This system is designed to classify hand gestures based on the Zoltán Kodály method, which is instrumental in teaching interval intonation and sung "solfeggio." Kodály's approach creates a system of signs that visually represent musical intervals, making them easier to recognize and understand. The Leap Motion Controller aids this process by capturing Pitch, Roll, and Yaw values, drawing inspiration from basic aircraft principles.

LM is now widely used by musicians to control basic sound parameters, emphasizing both advantageous and disadvantageous aspects. The advantages are related to the simplicity and potential of using hand gestures to control sound parameters. The key elements include the device's low frame rate, limitations of the distance between the hand and controller, and occlusion issues between the user's hands and the device [41].

In addition, LM is used also for rehabilitation purposes as a facilitated system for managing musical parameters for people with disabilities [42].

These studies highlight the growing intersection of neuroscience, music, and technology, showcasing the innovative potential of BCIs in musical applications. They explore the potential of BCIs in enabling users to create or manipulate music using brainwaves, each employing unique approaches such as integrating audio, gesture, and visual stimuli, developing novel graphical user interfaces, and utilizing advanced signal processing techniques. However, none of the existing works, to our knowledge, incorporates ML techniques to develop systems that can compose music using BCIs, as proposed in our research.

## NeuralPMG Framework Architecture

In this section, we describe the NeuralPMG architecture.

According to Fig. 2, NeuralPMG consists of three macro-components. The *Data Acquisition* module is in charge of data acquisition from the two devices, i.e., the Leap Motion and the Emotiv headset. The *AI Engine* module receives the EEG power band values from the CortexV.2 REST API and elaborates them with ML techniques to first create a user's mental state training dataset and later classify the current mental state. The *Music Generation Engine* module is responsible for (i) creating the melodic pattern based on the data acquired by the Leap Motion; (ii) creating polyphony based on the melodic pattern and the mental stated prediction value provided by the AI Engine; and (iii) displaying a GUI with the widgets for user interaction and system output. The dashed lines between the architecture modules represent control signal.

The processes of the modules are controlled using a Python Flask library<sup>3</sup> which allows us, through the API-Rest technology, to manage a communication protocol based on TCP-IP capable of handling server-client calls with related data exchange on a specific IP address.

Two devices are used for acquiring user data:

- **Brain-Computer Interface.** The Emotive Insight device<sup>4</sup> is a five-electrode passive headset capable of detecting electrical voltages on the head surface (i.e., the EEG). The electrodes are Semi-Dry Polymers, as they have a conductive rubber coating. Generally, they can be used without gels or conductive solutions. The EEG signal quality is verified by the proprietary Emotive PRO<sup>5</sup> software interface, which provides a signal quality check by means of colored indicators, from "Green = excellent quality" to "Black = very poor quality." As shown in Fig. 2, the EEG headset interface communicates, with the Engine module by an API Gateway, using the native CortexV.2 REST API. Once the signal calibration phase has passed, data is received by the Cortex API SDK. Electrodes are placed according to the official 10-20 system in positions AF3, AF4, T7, T8, and Pz [43]. Figure 3 shows a schematic representation of the electrodes with their positioning on the user's head.

The Cortex V.2 API by Emotive enables interfacing the device with various development environments. For each electrode, the API provides a series of numerical values corresponding to the power spectrum in the different bands of interest: Delta (0–4 Hz), Theta (4–8 Hz), Alpha (8–12 Hz), Beta (12–35 Hz), and Gamma (35–43 Hz).

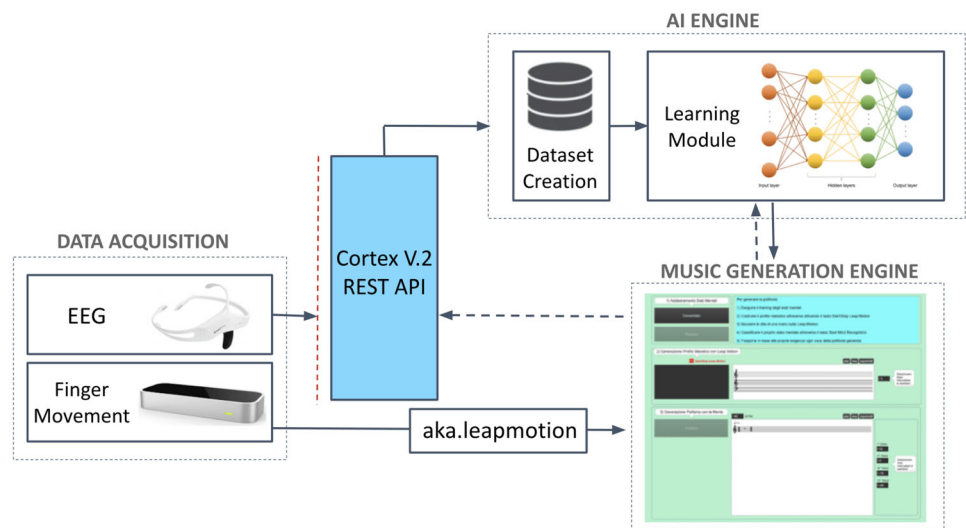
<sup>3</sup> Python Flask library: <https://flask.palletsprojects.com/en/2.0.x/>.

<sup>4</sup> <https://www.emotiv.com/insight/>.

<sup>5</sup> <https://www.emotiv.com/emotivpro/>.

<sup>2</sup> <https://gallery.leapmotion.com/geco-midi/>.

**Fig. 2** Overview of the NeuralPMG framework architecture main components



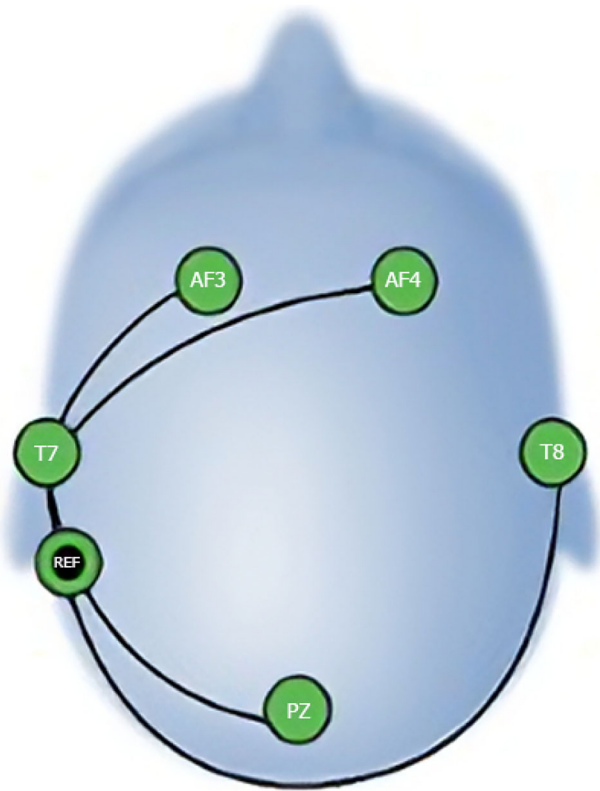
- **Leap Motion.** LM<sup>6</sup> is a powerful tool for recognizing the movement of different hand parts. Palm and fingers' data of one or both hands in the spatial coordinates on the  $x$ ,  $y$ , and  $z$  axes are detected by an infrared camera. The coordinates data streaming is received directly by `aka.leapmotion` object<sup>7</sup> in Max/Msp that visualizes them in a widget of the user interface.

sequences of rhythmic figures. Concerning dynamic progression in each polyphonic voice, amplitude variations corresponding to Alpha, Beta, Theta, Gamma, and Delta brainwaves are scaled within the 0–127 range, in accordance with MIDI protocol standards. The selection of rhythmic value sets and dynamic progression is subject to the user's discretion, controlled through mental states assumed during system usage.

## Polyphony Generation Process

The polyphony generation process is divided into two main parts, namely:

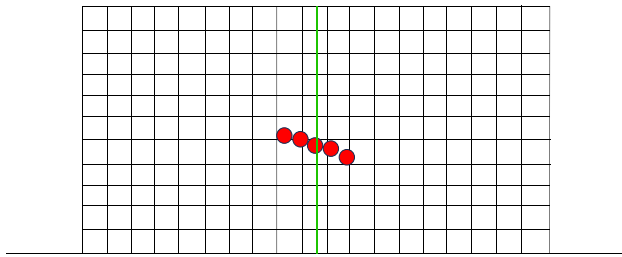
- **Generation of the basic melodic pattern:** It is generated by selecting a reference interval axis as proposed in Slonimsky's theory. The infra-inter-extrapolation process along with related permutations is determined by the movement of one hand's fingers (refer to "[Melodic Pattern Generation](#)" section for details).
- **Generation of polyphonic structure:** The system employs the previously generated melodic pattern to create a four-part polyphony. This involves overlaying four rhythmic profiles, derived from permutations of rhythmic figures found in two distinct reference sets, onto the four voices. The first set comprises rhythmic values of quarter notes, eighth notes, and sixteenth notes, along with their corresponding rest values. The second set includes rhythmic values equivalent to whole notes, half notes, and quarter notes, also accompanied by their respective rest values. These divided sets are then utilized to form



**Fig. 3** Electrode placement according to the 10-20 system

<sup>6</sup> <https://www.ultraleap.com/product/leap-motion-controller/>.

<sup>7</sup> <https://github.com/akamatsu/aka.leapmotion>.



**Fig. 4** Visualization of the user's left hand five fingers' position in the NeuralPMG GUI

## Mental State Generation

The system records power data from Alpha, Beta, Theta, Gamma, and Delta brainwave bands. This data is essential for defining mental states. As per literature [44], the *Focused* mental state correlates with increased power in Beta and Gamma bands, whereas the *Relaxed* state is associated with heightened Alpha band amplitudes. Initially, users must train the system to distinguish between these two mental states.

To elicit a Focused mental state, users engage in mental exercises that simulate an elevation in Beta and Gamma brainwave activities. The system sequentially displays numerical strings (e.g., [41-21-92-10-8-37-45-75-61-29-61-95-79-...]) for the user to observe and read. Conversely, to induce a Relaxed state, the system plays natural sounds such as flowing water and bird chirps. The data collected during the focused and relaxed mental state form a training dataset for various ML algorithms and simple feedforward neural networks with up to three hidden layers.

## Melodic Pattern Generation

Melodic pattern generation relies on the positioning of select fingers of one hand approximately 30 cm from a Leap Motion device. The Y-axis coordinates of the five fingers (left or right hand) are recognized. Figure 4 shows the position of the fingers at 45° angle relative to a horizontal reference plane. Three fingers are considered in building the melodic pattern: the first note is the reference tone; the second, third, and fourth notes are defined by the thumb, middle, and little fingers respectively. Figure 5 shows the same fingers translated onto the musical staff.

The melodic pattern generated through LM is scaled across two octaves using the `bach.mod`<sup>8</sup> library. This scaling process involves specific handling of input values relative to a defined threshold:

1. **Threshold processing:** The system employs a threshold of 2400 *cents*. For input values falling below this

threshold, the output value directly corresponds to the input value. Conversely, if the input value surpasses 2400 *cents*, the output is recalculated as *input value* - 2400 *cents*.

2. **Constant addition for note display:** To accurately represent the notes in both violin and bass clefs, a constant of +6000 *cents* is added to the output value. This adjustment ensures proper display and interpretation of the notes within the respective clef notations.
3. **Formation of the base melodic cell:** The base four-note melodic cell is derived initially. Subsequent notes are generated by adding constants that correspond to the selected interval axis. This process is pivotal in determining the melodic structure and is dependent on the axis choice:
  - **Unison interval axis:** If a unison interval axis is chosen, a value of 0 *cents* is added to the four base notes.
  - **Augmented fourth interval axis:** For an augmented fourth interval axis, each note of the base cell is incremented by +6000 *cents*.
4. **Display and notation:** The resulting pattern is then visualized in a software object that provides a transcription in mensural notation (as referenced in [45]). This display includes a consistent rhythmic profile, standardizing all notes to quarter notes.

Figure 6 illustrates the generated melodic pattern following these procedures.

## Polyphony Generation

The polyphony generation process starts with the analysis of the user's mental state, utilizing the Emotiv headset to capture EEG signals. These signals are transformed from the time domain to the frequency domain in the range 0.5–43 Hz. The system samples 50 time series at approximately 1-s intervals, each consisting of 25 features. This data, pertaining to Alpha, Beta, Theta, Gamma, and Delta bandwidth values, is then fed into a classification algorithm. The algorithm outputs a prediction value corresponding to a mental state class, which informs the rhythmic value selection in the Max/MSP software for generating the rhythmic profiles of the four voices. The rhythmic values are divided into the following two categories:

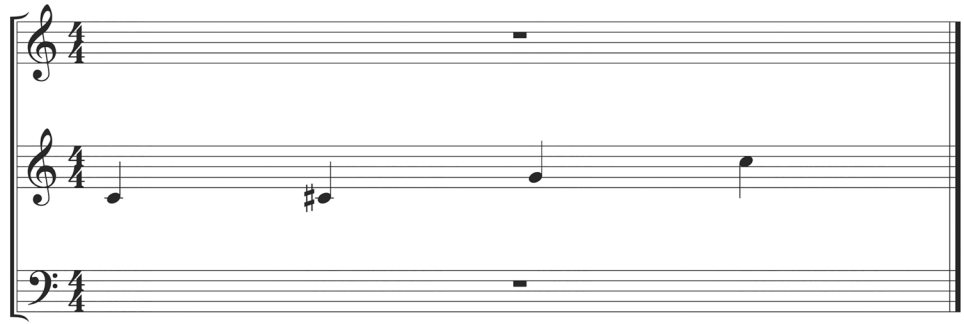
1. *Focused* state: consists of rhythmic figures 1/4, 1/8, 1/16 and their corresponding rest values;
2. *Relaxed* state: consists of rhythmic values of 4/4, 2/4, 1/4 and their corresponding rest values.

Random permutations of these rhythmic values generate strings of 49 rhythmic figures, assigning one rhythmic value

<sup>8</sup> <https://www.bachproject.net/>.



**Fig. 5** Finger transposition on the musical staff



**Fig. 6** Melodic pattern of infra-inter-ultrappolation on augmented fourth C-F# interval axis



to each note. These strings are then overlaid on the pitches of the basic patterns distributed across the four voices in the score, visualized in the `bach.score` object. This visualization facilitates observing the complete polyphonic score as shown in Fig. 6.

Subsequently, the score undergoes further refinement by transposing each voice along a selected interval axis, as per Slonimsky's theory. This involves adding a numerical constant to each note and voice, corresponding to the chosen interval axis. Figure 7 elucidates this transposition operation, showcasing the initial set of pitches in the top box and the altered set in the bottom box after adding a numerical constant of 300.

The rectangle labeled "I<sup>o</sup>voice transposition" in Fig. 1 performs the summation operation between the top height list plus the numeric constant. The input height list in the `bach.expr` method is represented by the variable  $x_1$ , while the constant is represented by the variable  $x_2$ . In the lower panel of Fig. 7, the output list on which the summing operation was performed can be observed.

Additionally, the amplitude values from the relevant EEG bands are utilized to set dynamic values for each note. These dynamics are translated into MIDI velocity values, ranging from 0 to 127. The Max/MSP's native scales object is calibrated to scale these values appropriately, with input ranges from 0 to 100 (reflecting the maximum bandwidths provided by the Emotiv Insight) and output ranges from 10 to 127. This ensures that notes are never assigned a velocity of 0, which would equate to silence.

Upon completion of polyphony generation, the piece can be audibly rendered. The tempo is adjustable via beats per minute (BPM), allowing for the input of the desired metronomic speed. Additionally, input boxes are provided for further modification of the score, enabling transposition of each polyphonic voice to preferred intervals. The finalized composition can be exported in MIDI or XML formats.

## NeuralPMG Engine

This section outlines the functioning of the NeuralPMG Engine. Two mental states are recorded (Focused/Relaxed) and a dataset is created. For each user, NeuralPMG identifies the best ML algorithm able to recognize the mental state. Using LM allows the user's finger movement to be recognized and generate the melodic pattern.

Different ML models are used and compared in the training phase. The best model is then used for predicting the user's mental state. The system then generates polyphony from the melodic pattern according to the user's detected mental state.

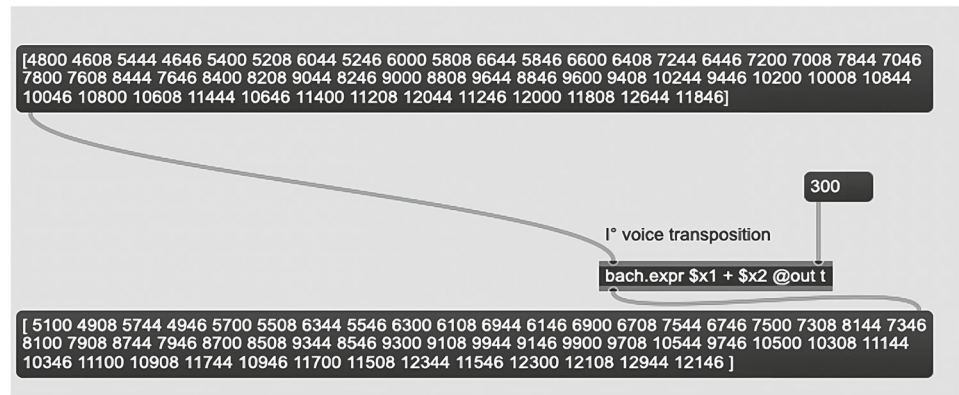
**Dataset Generation** The acquisition of the mental state is managed by specific routines. There is TCP-IP call to Flask<sup>9</sup> enabling communication between the BCI device and the Cortex API.<sup>10</sup> Once the connection has been established, the API starts streaming power band data for each band and electrode. They are absolute values whose unit is expressed in  $\mu V^2/Hz$ . Each user has a dataset consisting of 25 features based on 5 power bands for 5 channels, with 100 total acquisitions: the first 50 items collected while the user was stimulated to the Focused state by reading sequences of numbers, the next 50 while listening to natural environment sounds to induce the Relaxed state. As an illustrative example, the Supplementary Table 1 available in the Appendix to this article presents the mental state dataset acquired for a single user.

**Classification Model Evaluation** Determining the user's mental state is influenced by several factors, as it is well-known in the EEG analysis domain. First, values for delta,

<sup>9</sup> <https://flask.palletsprojects.com/en/2.2.x/>.

<sup>10</sup> <https://emotiv.gitbook.io/cortex-api/>.

**Fig. 7** Transposition operation of the pitch set



theta, alpha, beta, and gamma frequencies can vary among different users. Furthermore, users have different reactions on internal or external stimuli, thus having different abilities in relaxing or concentrating without being influenced by the surrounding environment. To mitigate these problems, we selected six ML models and tested their performances in classifying mental state by involving five users. The models are trained in parallel for each user; this approach is possible because the training dataset is not large and therefore it does require reasonable computational resources and time to execute. The following models were considered:

- Linear Discriminant [46];
- Decision Tree [47];
- Naive Bayesian [48];
- Support Vector Machine [49];
- K-nearest neighbors [50];
- Feedforward Neural Network [51] (1 hidden layer with 10-neuron);
- Feedforward Neural Network [51] (1 hidden layer with 25-neuron);
- Feedforward Neural Network [51] (2 hidden layers with respectively 10 neurons per layer);

To perform the training phases of the models, we used the Matlab toolbox - Classification Learner.<sup>11</sup> It performs the following steps: (i) Z-score standardization and (ii) training phase with k-fold = 5. The toolbox returns the model with the best accuracy among all models considering also the best performance of the model on the five folds obtained in the training phase. For each model, a random search is performed to select the best hyperparameters. The accuracy is calculated as in Eq. (1).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

<sup>11</sup> <https://it.mathworks.com/help/stats/classificationlearner-app.html>.

In Eq. (1), TP, TN, FP, and FN represent the number of true positive, true negative, false positive, and false negative predictions, respectively.

Table 1 shows test results. KNN showed extremely low accuracy values; therefore, it was not considered in the system implementation. The feedforward neural network 1 (FNN1), despite its weak performance, was implemented in the system.

## Evaluation Study

In this section, we describe the user study that aimed at evaluating the interaction with the NeuralPMG from various perspective.

## Participants and Design

The study on NeuralPMG specifically involved participants from the *Electronic Music Laboratory* of the *E.R. Duni Conservatory of Matera*. This targeted selection was based on two key prerequisites: expertise in algorithmic composition and proficiency in both algorithmic composition and sound manipulation software. This focus on electronic music students was intentional. Traditional music composers, typically from classical backgrounds, were not considered suitable for

**Table 1** Mental state classification performances of the considered models: LD, linear discriminant; DT, decision tree; NB, naive Bayesian; SVM, support vector machine; FNN, feedforward neural network

User	Accuracy performance							
	LD	DT	NB	SVM	KNN	FNN1	FNN2	FNN3
1	0.67	0.86	0.90	0.88	0.48	0.66	0.78	0.45
2	0.43	0.52	0.67	0.69	0.34	0.57	0.60	0.59
3	0.84	0.65	0.67	0.80	0.62	0.64	0.80	0.82
4	0.67	0.78	0.65	0.89	0.56	0.59	0.98	0.92
5	0.89	0.67	0.64	0.76	0.40	0.68	0.67	0.52



**Fig. 8** Emotive Insight headset positioning and calibration

this study. Their training often emphasizes manual composition methods, including hand-writing music, and frequently avoids the use of digital tools like music notation software. In contrast, our study required familiarity with computer systems integral to algorithmic composition.

All participants in the study possessed a moderate level of experience in algorithmic and computer-aided music composition. While some were professional instrumentalists, all participants were skilled in at least one musical instrument, aligning with the Conservatory's educational standards. Notably, none of the participants was previously acquainted with Slonimsky's method.

The participant group was diverse in terms of age and gender: age min = 19, age max = 61, age avg = 30, females = 10, males = 9, 1<sup>st</sup> academic degree level = 11 and, 2<sup>nd</sup> academic degree level = 8.

## Procedure

Three human-computer interaction (HCI) experts participated in the study: two as observers recording task execution times and problems, and one as a facilitator managing participant-system interactions. The experimental procedure had participants performing the study individually. The study was conducted in the "Rota" hall of the Conservatory under medium lighting conditions, optimal for task visibility without being invasive. The room climate was at a comfortable 21 °C, aided by its thick stone and tuff walls.

The study comprised six distinct stages:

1. **Initial Procedure:** The participant was first required to sign a consent form permitting photo documentation for research analysis and authorizing continuation of the study. Following consent, the participant was seated at the station where the study equipment was setup.

2. **Calibration Phase:** Participants donned the Emotive Insight headset, adjusted for electrode placement per the 10-20 standard system. The device was activated, and electrode impedance checked using EmotivePro software. Calibration duration varied due to individual hair types and skull shapes. EEG signal quality was ensured for subsequent tasks, maintaining device battery above 50% as recommended by the manufacturer.<sup>12</sup> Conductive gel was applied as necessary, and participants were advised against sudden head movements or expressive facial gestures. Figure 8 shows the facilitator committed to positioning the device on a participant's head and calibrating the device.

3. **Training Mental States:** The study utilized machine learning and neural network algorithms to differentiate between the participant's "Focus" and "Relaxed" mental states. For "Focus," participants read a sequence of randomly generated numbers on-screen. For "Relaxed," participants were asked to breathe diaphragmatically while listening to a nature soundscape. This phase's accuracy is critical for the success of subsequent stages, which rely heavily on the classifier predictions Fig. 9.

The training of mental states is a very important phase as it could significantly influence the subsequent phases of the study. The rhythmic generation of polyphony depends on the prediction of the classifiers, which must therefore be trained correctly with training data as consistent as possible with the user's mental states.

4. **Generation of the Melodic Pattern:** Post-calibration, participants were instructed on software commands to create a melodic pattern:

- Toggle the start/stop leap motion button.
- Move the hand and fingers above the leap motion sensor (at 10 cm) to compose a melodic pattern.
- Save the melodic pattern by holding a hand still over the sensor and pressing the start/stop button.
- Press "Play" to listen to the melody, modify the interval axis if necessary, and repeat until satisfied.

This phase averaged 10 min for each participant.

5. **Polyphony Generation:** Following Melodic pattern creation, participants generated polyphony through these steps:

- Press Dial to generate polyphony.
- Press Play to hear the polyphony adjust speed (BPM) or interval axes as needed, and repeat until satisfied.
- Save the composition by pressing the export midi button.

<sup>12</sup> <https://emotiv.gitbook.io/insight-manual/introduction/technical-specifications>.



**Fig. 9** Melodic pattern generation by moving the left (or right) hand over the Leap Motion device

The overall average completion time was approximately 20 min, including headset calibration.

## Data Collection

Both quantitative and qualitative data were collected through the answers to the questionnaires the participants filled in during the study and the notes taken by the observer on significant behaviors or externalized comments of the participants. All the interactions were audio-video recorded.

Initially, the participants filled in a questionnaire for collecting demographic data and their competences on IT, electronic music composition, and Slonimsky's method. Data about this have been already reported in "[Participants and Design](#)" section.

A second questionnaire consisted of 8 sections and aimed at evaluating the interaction with NeuralPMG from different perspectives. Each of the first five sections included a questionnaire proposed in the literature, as detailed in the following, while the questions in the remaining 3 sections were defined by the authors for investigating specific aspects:

1. **User eXperience** by *AttrakDiff* questionnaire [52]: 28 seven-step items whose poles are opposite adjectives (e.g., "confusing - clear," "unusual - ordinary," "good - bad"). It is based on a theoretical work model illustrating how the pragmatic and hedonic qualities influence the subjective perception of attractiveness giving rise to consequent behavior and emotions. In particular, the following system dimensions are evaluated: (i) *Pragmatic Quality (PQ)*: describes the usability of a system and indicates how successfully users are in achieving their goals using the system; (ii) *Hedonic Quality-Stimulation (HQ-S)*: indicates to what extent the system support those needs in terms of novel, interesting, and stimulating functions, contents, interaction, and presentation-styles; (iii) *Hedonic Quality - Identity (HQ-I)*: specifies to what extent the system allows user to identify with it; (iv) *Attractiveness (ATT)*: describes a global value of the system based on the quality perception.
2. **Support to creative design** by *Creativity Support Index* questionnaire [53]: 12-item psychometric survey to evaluate the ability of a tool in supporting users engaged in creative works and which aspects of creativity support may need attention. The CSI measures six dimensions of creativity support: *Exploration*, *Expressiveness*, *Immersion*, *Enjoyment*, *Effort*, and *Collaboration*.
3. **Workload** by *NASA-TLX* questionnaire [54]: Six-item survey that rates perceived workload in using a system through six subjective dimensions, i.e., *Mental Demand*, *Physical Demand*, *Temporal Demand*, *Performance*, *Effort* and *Frustration*, which are rated within a 100-points range with 5-point steps (lower is better). These ratings were combined to calculate the overall NASA-TLX workload index.
4. **User Engagement** by *User Engagement Scale (UES) short-form* questionnaire [55]: 12-item survey, derived from the UES long form, used to measure the user engagement, a quality characterized by the depth of a user's investment when interacting with a digital system, which typically results in positive outcomes [56]. This tool measures user engagement by summarizing an index that ranges from 0 to 5. It also provides detailed information about four dimensions of user engagement, i.e., *Focused Attention (FA)*, *Perceived Usability (PU)*, *Aesthetic Appeal (AE)*, and *Reward (RW)*.
5. **Emotional Response** by *Self-Assessment Manikin (SAM)* questionnaire [57]: Two picture-oriented questions to measure an emotional response in relation to *pleasure*, *arousal* and *dominance*. In particular, the two questions asked participants to indicate, in relation to the polyphony generated, respectively: (1) the positive/negative liking; (2) the excitement felt in relation to the emotion. The score could be expressed through a Likert scale with scores from 1 to 9 representing the degree of agreement ranging from "negative" to "positive" for the first question, and "low exciting" to "high exciting" for the second.
6. **Self-assessment of the Generated Polyphony**: Eight questions specifically designed by the authors of this article (scale 1–7), organized in three sections to investigate:
  - (a) Need for further modifications of the generated polyphony with respect to pitches and dynamics (two questions);
  - (b) Which compositional techniques are considered most suitable for further processing (three questions);
  - (c) Which application domains are considered most suitable given the harmonic/melodic context of the generated polyphony (three questions).



7. **NeuralPMG appreciation:** Two open questions asking the most and the least aspects that the participant appreciated.
8. **Comments and suggestions:** One open question allowing participants to provide comments and suggestions.

The average time taken by participants to complete the questionnaire was 12 min.

## Results

In the following subsections, data collected during the experimental study are analyzed and results are provided.

### Classification Performances of NeuralPMG Models

Table 2 shows the accuracy obtained during the experiment by each model for each of the 19 participants. It emerges that the different models for classifying the participants' mental state were selected homogeneously by the engine (LD, DT, FNN2 = 4; SVM = 5), except for NB which was used for only two participants. FNN1, which in the pilot test had weak performance, has never been the best classification model and was used for no participants.

**Table 2** Model accuracy for each of the 19 participants. *LD*, linear discriminant; *DT*, decision tree; *NB*, Naive Bayesian; *SVM*, support vector machine; *FNN*, feedforward neural network. The gray background indicates that model has been selected for that participant

Participant	Accuracy performance					
	LD	DT	NB	SVM	FNN1	FNN2
1	0.67	0.78	0.85	0.66	0.51	0.43
2	0.54	0.65	0.45	0.69	0.52	0.58
3	0.84	0.76	0.67	0.82	0.64	0.47
4	0.34	0.55	0.41	0.86	0.78	0.94
5	0.89	0.78	0.68	0.72	0.81	0.84
6	0.90	0.87	0.77	0.82	0.67	0.78
7	0.55	0.52	0.41	0.48	0.50	0.51
8	0.34	0.45	0.44	0.49	0.31	0.30
9	0.80	0.87	0.90	0.97	0.87	0.88
10	0.45	0.80	0.51	0.78	0.80	0.97
11	0.30	0.73	0.45	0.56	0.70	0.75
12	0.67	0.56	0.78	0.80	0.78	0.90
13	0.45	0.91	0.34	0.78	0.68	0.87
14	0.30	0.40	0.46	0.44	0.23	0
15	0.56	0.39	0.4	0.76	0.6	0.5
16	0.70	0.80	0.64	0.79	0.73	0.77
17	0.43	0.44	0.65	0.32	0.7	0.98
18	0.67	0.91	0.80	0.78	0.79	0.46
19	0.31	0.20	0.30	0.43	0.38	0.40

**Table 3** Mean (AVG score) and standard deviation (*SD* score) for each category of the AttrakDiff questionnaire

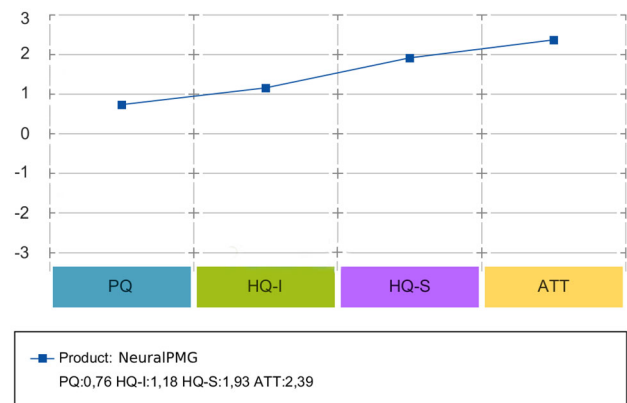
	AVG Score	SD score
PQ	5.83	0.81
HQ-I	6.48	0.60
HQ-S	7.25	0.73
ATT	7.50	0.63

### User Experience (UX)

An overview of the AttrakDiff results is presented in Table 3 and in the diagram shown in Fig. 10, which summarize the hedonic (HQ) and pragmatic (PQ) qualities of the system according to their respective confidence rectangles. In general, the larger the rectangle, the greater the uncertainty about the region to which the system belongs. NeuralPMG, therefore, has high HQ and PQ values; thus, it can be classified as a *desiderable* product with promising UX. Furthermore, the value of HQ (1.55, 0.29) is higher than that of PQ (0.76, 0.35), with partial values of attractiveness (ATT) (7.50, 0.63), hedonic quality-identity (HQ-I) (6.48, 0.60), hedonic quality-stimulation (HQ-S) (7.24, 0.73), and pragmatic quality (PQ) (5.82, 0.80) Fig. 11.

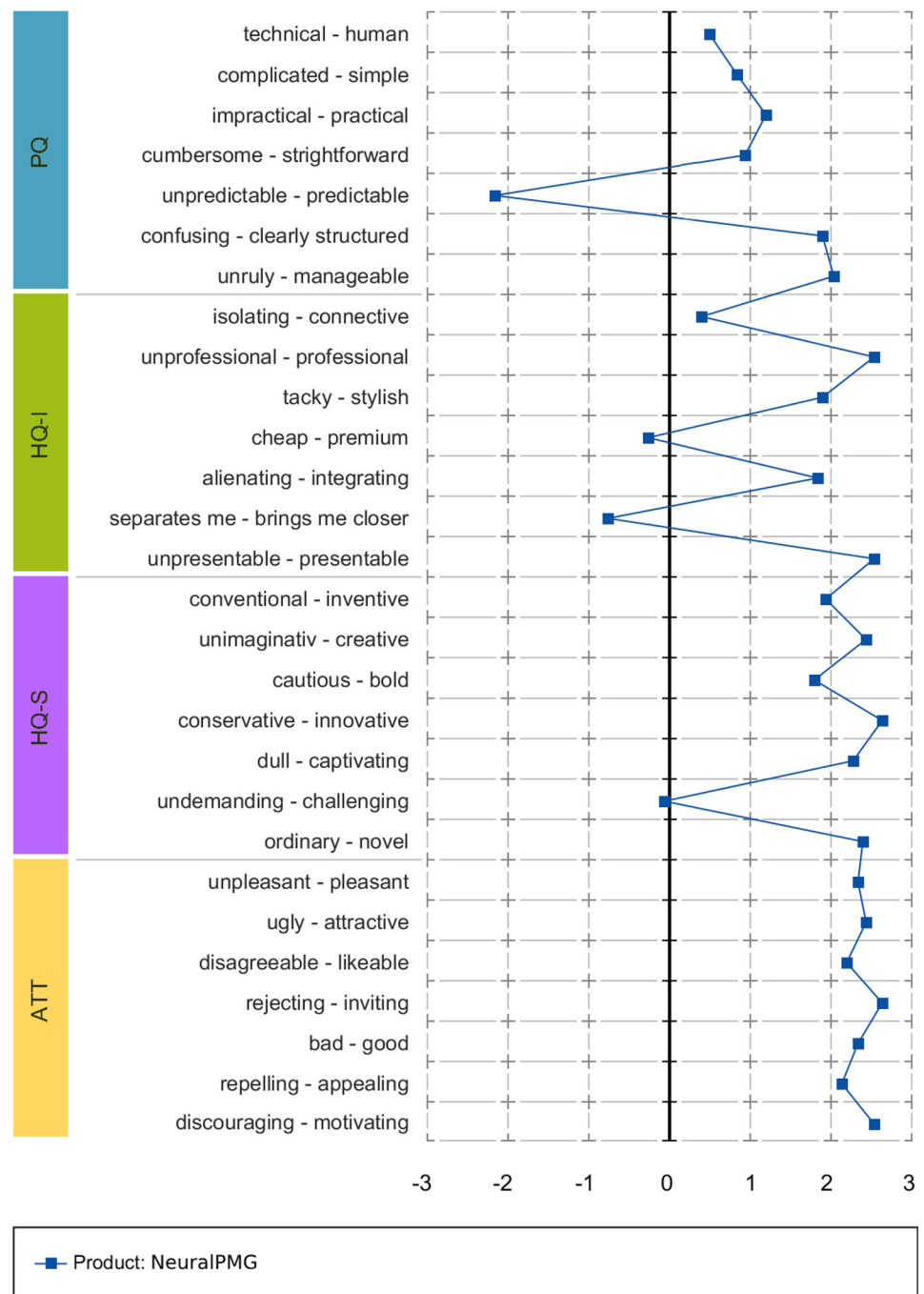
### Support to Creative Design

Using the CSI questionnaire, the participants' perceptions of creativity support were measured. From Fig. 12, it can be seen that the system achieved an average CSI score of 70/100, which means good support for creative design (CSI=70.00, STD=14.86). The mean and standard deviation of the system CSI dimensions were reported as 1.5.



**Fig. 10** Graphical representation of the average results of the AttrakDiff questionnaire

**Fig. 11** Graphical representation of AttrakDiff questionnaire word pairs



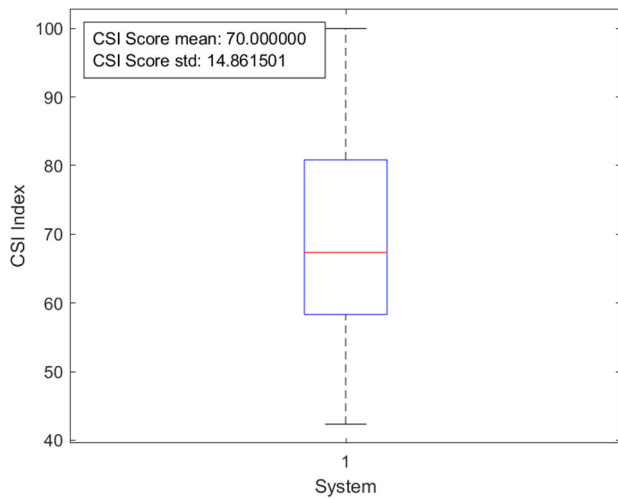
As can be seen from Table 4, the highest average result is immersion (8.18), i.e., the degree of immersiveness in the system, followed by expressiveness (73.2), exploration (76.3), and effort/reward trade-off (RWE 74.5). As expected, the lowest value was obtained for collaboration (50.3) since there is no user collaboration in the system.

In Fig. 13, we observe the proportion of relevance to the averages of the values for each field of the CSI. The fields of interest in exploration, enjoyment, and collaboration are evident.

## Workload

The workload data collected through the NASA-TLX are shown in Table 5, where the weighted average (AVG) and its standard deviation (SD) are indicated for each dimension. The Likert scale used in the questionnaire ranges (1–10).

The mental effort dimension presents a weighted mean value of 7.42 with an SD value of 7.68. These values denote a high mental workload. Physical effort was low with a weighted mean value of 0.63 and SD of 2.75. This result



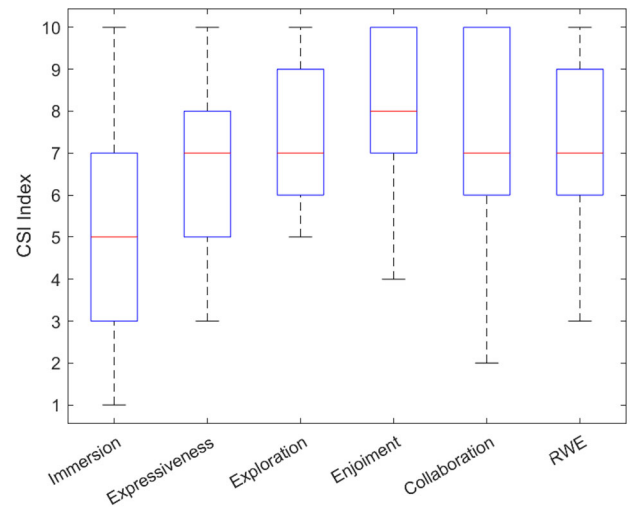
**Fig. 12** Box chart representation of the global support to creativity perceived by participants, based on the CSI questionnaire

is consistent with expectations since no physical effort was required to the users. The participants' perception of the time spent presents insignificant values as the mean value is 5.53 and the SD is 5.57. The best result is obtained with the Performance category, which presents average values of 8.11 and SD of 3.02, indicating positive feedback from users on the result. Like the perception of time, the Effort presents insignificant values (AVG 4.95, SD 6.84) which cannot be discussed as their SD is extremely high. Finally, excellent values can be observed with regard to "Frustration," which presents an average of 3.95 and an SD of 5.03. However, the SD of the latter factor is not negligible and could be due to some problems in the headset calibration and training phase.0

### User Engagement

The results of the UES short form provide a value between 1 ("I do not agree") and 5 ("I agree") and are based on participants' involvement data during the interaction with the system. Figure 14 shows that the overall user involvement has an average of 4.13 and SD of 0.41.

Furthermore, Table 6 shows averages and standard deviations for each dimension, i.e., focused attention (FA),



**Fig. 13** Box plot representation of overall user involvement according to the four dimensions of the CSI questionnaire

perceived usability (PU), aesthetic appeal (AE), and reward for interaction with the system (RW). Satisfactory results were obtained, as the average value of PU is 4.28 and the SD is 0.64.

The last two parameters, expressing respectively the aesthetic appeal and the reward for interaction with the system, turn out to be remarkably high. In addition to this table, it is also possible to observe in Fig. 15 the box plot of the UeS data is divided into categories.

In Table 6, we can see that the AE indicator presents a mean value of 4.28 (with an SD of 0.71), which denotes a high visual attractiveness on the part of students. Finally, the RW category presents the highest mean value of 4.46 (with an SD of 0.6). This value highlights how the participant, at the end of his tasks, felt satisfied with what he had produced.

### Emotional Response

Figure 16 shows the heatmap, average value, and standard deviation for participants' emotional response to the polyphony they generated.

Table 7 shows that two participants (nn. 1 and 13) stated that they are Sad, Depressed, and Bored. Twelve participants

**Table 4** Mean (AVG score) and standard deviation (SD score) for CSI questionnaire dimensions

	AVG count	SD count	AVG score	SD score	AVG weigh. score	SD weigh. score
Exploration	4	1.08	7.63	1.67	30.12	1.80
Collaboration	3	1.18	5.03	2.83	16.14	3.35
Immersion	2	1.60	8.18	1.72	18.52	2.74
Expressiveness	1.42	1.07	7.32	1.93	10.40	2.07
Enjoyment	3	1.33	6.74	2.01	20.21	2.68
Effort/reward tradeoff	1	1.61	7.45	2.34	8.62	3.77

**Table 5** Mean (AVG) and standard deviation (SD) for NASA-TLX questionnaire dimensions

	AVG count	SD count	AVG score	SD score	AVG weigh. score	SD weigh. score
Mental demand	1.21	0.92	4.63	2.83	7.42	7.68
Physical demand	0.11	0.46	2.68	1.73	0.63	2.75
Temporal demand	1.37	1.16	3	1.83	5.53	5.57
Performance	0.89	0.32	8.68	1.57	8.11	3.02
Effort	1	2.47	1	2.47	4.95	6.84
Frustration	0.74	2.35	0.74	2.35	3.95	5.03

indicated that they are Excited, Delighted, and Happy, while two felt Sleepy, Calm, and Content (nn. 16 and 19). Finally, three participants (nn. 3, 7, and 12) were neutral regarding the emotions they felt about the polyphony they created.

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (2)$$

Table 8 summarizes the results of the softmax activation function used to calculate percentages using the Eq. (2), where  $\vec{z}$  is the input vector to the softmax function made up of  $(z_0, \dots, z_K)$ ,  $z_i$  values are the elements of the input vector to the softmax function;  $e^{z_i}$  is the standard exponential function applied to each element of the input vector,

$$\sum_{j=1}^K e^{z_j} \quad (3)$$

is the term on the bottom of the formula is the normalization term. It ensures that all the output values of the function will sum to 1 and each be in the range (0, 1), thus constituting a valid probability distribution, and  $K$  is the number of

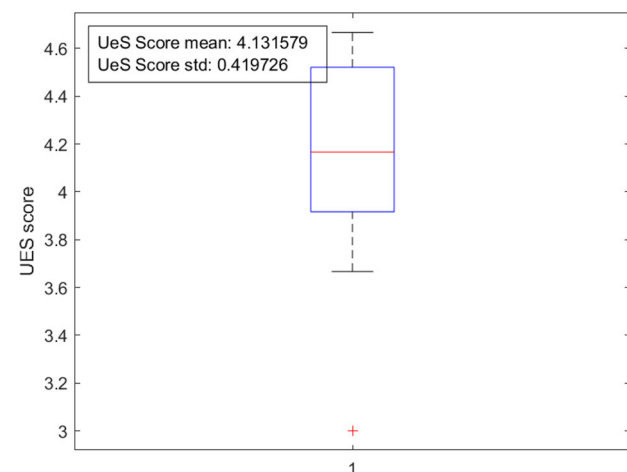
classes in the multi-class classifier. We can observe an high positive score.

### Self-assessment of the Generated Polyphony

Considering the answers to the two questions in the first section, i.e., need for further modifications of the generated polyphony with respect to pitches and dynamics that in Fig. 17 we have labelled as *Elaboration*, the average value of 3.42 out of 7 suggests that the generated polyphonies require substantial modifications when they are to be finalized for end use. Considering the answers to the three questions in the second section, i.e., whose compositional techniques are most suitable for the further elaboration of the polyphonies that we have labelled as *Compositional Techniques*, it appears that the most suitable techniques are aleatory (AVG = 5.42) and serial/atonal (AVG = 5.21), while aleatory techniques (AVG = 3.37) are not very suitable. Finally, the answers to the three questions in the third section, i.e., whose domains are considered most suitable for the application of polyphonies generated through the system that we have labelled as *Usage*, show the prevalence of the Multimedia (AVG = 5.84) and Academic (AVG = 5.16) domains, while commercial music for entertainment (AVG = 3.47) is less appropriate.

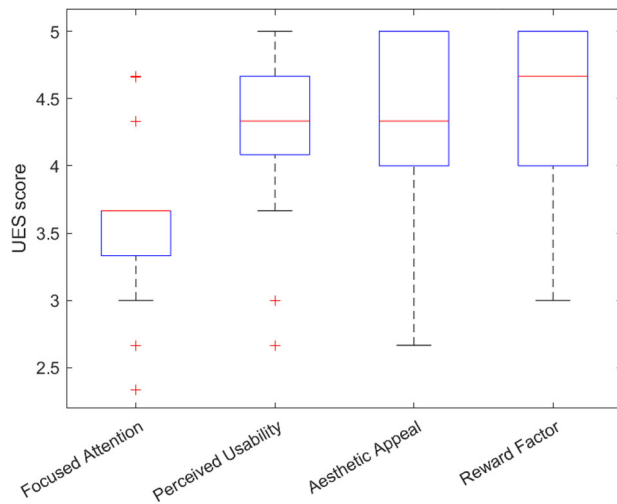
### NeuralPMG Appreciation

The two open answers provided by the participants were analyzed in a systematic qualitative interpretation using an inductive thematic analysis [58]. Two HCI researchers followed the six-step procedure proposed by Braun et al.

**Fig. 14** Box chart representation of the global participants' involvement score based on the UES questionnaire**Table 6** Mean (AVG) and standard deviation (SD) for NASA-TLX questionnaire dimensions for UES questionnaire dimensions

	AVG score	SD score
FA-S	3.51	0.56
PU-S	4.28	0.64
AE-S	4.28	0.71
RW-S	4.46	0.60





**Fig. 15** Box plot representation of overall participants' involvement score according to the four dimensions of the UES questionnaire

[58]: data familiarization, coding, themes generation, themes review, themes naming, and theme description. In the end, a set of themes representative of the most appreciated and least appreciated aspects of the system were identified. For the appreciated elements: *"Production of unconventional melodies," "Overall system novelty," "Power of generative grammar," "Production of unconventional polyphonies," "Use of BCI system,"* and *"Use of Leap Motion device"*; for

the critical elements: *"Use of BCI system," "Use of Leap Motion device," "System graphical user interface," "System appreciation,"* and *"System efficiency"*. The frequency of each theme in participants' responses was then calculated, as reported in Tables 9 and 10.

### Comments and Suggestions

Seven participants left an additional comment answering the last open-ended question of the questionnaire. All comments were positive. The respondents expressed appreciation towards the system (four comments), the wish to use the system in their work (two comments). One participant also suggested the possibility of using other biometrics, such as heartbeat.

### Polyphony Assessment by Music Experts

Two professors of the E.R. Duni Conservatory of Matera, both holding the title of Maestro, and with extensive experience in music composition were involved. They individually listened to the 19 polyphonies generated by the 19 participants in the previous study and, for each one, answered a questionnaire structured in six sections:

**Table 7** Participants' emotions about the generated polyphony

Participant	Emotion									
	Sad	Depressed	Bored	Excited	Delighted	Happy	Sleepy	Calm	Content	Neutral
1	✓	✓	✓	-	-	-	-	-	-	-
2	-	-	-	✓	✓	✓	-	-	-	-
3	-	-	-	-	-	-	-	-	-	✓
4	-	-	-	✓	✓	✓	-	-	-	-
5	-	-	-	✓	✓	✓	-	-	-	-
6	-	-	-	✓	✓	✓	-	-	-	-
7	-	-	-	-	-	-	-	-	-	✓
8	-	-	-	✓	✓	✓	-	-	-	-
9	-	-	-	✓	✓	✓	-	-	-	-
10	-	-	-	✓	✓	✓	-	-	-	-
11	-	-	-	✓	✓	✓	-	-	-	-
12	-	-	-	-	-	-	-	-	-	✓
13	✓	✓	✓	-	-	-	-	-	-	-
14	-	-	-	✓	✓	✓	-	-	-	-
15	-	-	-	✓	✓	✓	-	-	-	-
16	-	-	-	-	-	-	✓	✓	✓	-
17	-	-	-	✓	✓	✓	-	-	-	-
18	-	-	-	✓	✓	✓	-	-	-	-
19	-	-	-	-	-	-	✓	✓	✓	-

**Fig. 16** Participants' emotional response to the generate polyphony in terms of *Valence* and *Arousal*

Emotional aspect	Participant																			Descriptive statistics								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	AVG	SD							
Valence	4	8	5	7	9	7	7	6	9	7	9	7	2	7	9	6	7	7	6	6.79	1.78							
Arousal	3	8	2	7	9	6	5	6	9	6	6	5	3	6	7	4	7	7	4	5.79	1.96							
	Rating																			9	8	7	6	5	4	3	2	1

1. Three questions focusing on the evaluation of the aesthetic aspects of the polyphony;
2. Four questions addressing the technical evaluation of the harmonic aspects;
3. Three questions analysing the polyphony level of elaboration;
4. Three questions dealing with the evaluation of possible usage scenarios of the polyphony;
5. Nine questions assessing each polyphony with respect to its ability to arouse nine different emotions: happiness, tenderness, happiness, anger, sadness, fear, negativity, activity, positivity, tension.
6. One open question to provide further comments and opinions.

The first five sections were based on a Likert scale with values from 1 to 7, representing respectively the degree of agreement ranging from "not at all" to "very much."

Results of the music experts' assessment have been summarized by heatmaps, complemented with average values and standard deviation for every question and section. Figures 18 and 19 refer to the polyphony evaluation provided by Maestros 1 and 2, respectively.

## Discussion

In this section, we delve into the implications of the results from the NeuralPMGEvaluation. We explore various dimensions including *user experience (UX)*, *creativity*, *workload*, participants' *engagement*, participants' *self-assessment of the generated polyphony*, *application domains*, and participants' *feedback*. This discussion, which also takes into account the evaluation performed by the two Maestros, not only addresses the system strengths and areas for improvement but also reflects on its broader impact on musical composition and user interaction.

**Table 8** Emotional score softmax activation function

Excited, delighted, happy	0.99978
Sleepy, calm, content	0.00005
Neutral	0.00012
Sad, depressed, bored	0.00005

Regarding *UX*, the AttrakDiff questionnaire results indicate that participants generally provided positive feedback about NeuralPMG, highlighting its attractiveness. However, it is clear that there is a need for enhancements in its pragmatic quality.

In terms of *creativity*, as per the CSI results, the exploratory aspect scored the highest. This is in line with our expectations, given the participants' professional activity, applications of this type are perceived with particular interest. The coefficient of RWE (effort), despite achieving a fairly high score, played a marginal role for participants. Both *immersiveness* and *enjoyment* turn out to be particularly weighty components. However, it is worth noting that *expressiveness* did not emerge as a significant factor for them. This result is consistent with the sample of participants, since generally every musician or composer does not look for the expressive capacity in software or musical instruments but uses them to satisfy his own expressive need. We can therefore assume that the system stimulates and supports the creativity of the participants who, while using the system, remained focused on the final task, isolating themselves from their surroundings.

Regarding *workload*, *physical effort* was low, but some participant's *frustration* can come because of problems with BCI headset calibration.

Participant *engagement* while interacting with NeuralPMG was positively assessed by the UES questionnaire, as also confirmed by the measured *Emotional response*.

In the *self-assessment of the generated polyphony*, participants recognized system *effectiveness* in facilitating the complex process of managing and calculating note organizations within the tempered space. While the polyphonies based on Slonimsky's grammar require *additional refinement* in pitch and dynamics, they serve well the composer's initial creative needs, who wants to create rough polyphonies to be later refined according to a specific application. Also, the results align with the fundamental principles of serial and atonal *techniques*, which focus on exploring diverse organizational approaches in musical composition. The entire theoretical evolution of music focuses on one main question, namely that of providing the composer with the possibility of being able to explore the possible musical space both in the melodic/polyphonic direction and in the rhythmic and timbral direction. Finally, about the possible *application domains*, the participants found the polyphony generated more suitable for academic or multimedia application, rather

**Fig. 17** Participants' self assessment of the generated polyphony

Section	Specific Aspect	Participant																			Descriptive statistics				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	AVG	SD	AVG	SD	
Elaboration	No pitch editing required	4	6	2	4	4	5	3	6	4	7	6	6	4	4	7	3	2	7	3	4.58	1.64	4.58	1.62	
	No dynamic editing required	3	4	5	3	4	6	5	7	4	7	6	4	2	4	7	3	3	7	3	4.58	1.64			
Compositional Techniques	Tonal	4	1	2	3	7	5	1	4	4	7	4	2	2	3	7	2	3	1	2	3.37	1.98	5.21	1.36	
	Serial and/or Atonal	6	5	6	5	7	6	4	5	4	7	5	5	6	6	4	6	5	1	6	5.21	1.36		4.67	1.94
	Aleatory	5	5	6	6	7	6	5	7	1	7	6	4	7	5	7	6	7	1	5	5.42	1.80			
Usage	Academic Music domain	6	3	6	5	7	5	4	6	7	7	5	1	6	4	7	4	5	7	3	5.16	1.68	5.84	1.21	
	Multimedia domain	7	7	4	6	7	6	4	6	7	7	7	4	6	4	7	6	4	6	6	5.84	1.21		4.82	1.91
	Music Entertainment domain	4	1	3	3	7	3	4	2	1	7	4	2	2	3	7	1	3	3	6	3.47	1.98			
Rating		7	6	5	4	3	2	1																	

than commercial, reflecting a tradition in “cultured” music of rational experimentation and innovation. All composers belonging to the sphere of cultured music, throughout the history of music, have always experimented with new compositional solutions in order to search for useful compositional models to generate never-before-heard and innovative music. In multimedia contexts, particularly in movie soundtracks, these compositions effectively convey dynamic emotions and narratives, like portraying the passage of time or emotional undertones in a scene, through sound textures and harmonic layers.

Analysis of responses to open-ended answers of the questionnaire revealed insightful *participants' feedback* regarding the NeuralPMG system. Seven participants praised its effectiveness in decoding mental states for polyphony generation. However, concerns about the comfort of the BCI device were noted by five participants, underscoring the need for improved wearability. It suggests that while the introduction of BCIs in this domain is convenient and useful, the aspects concerning the wearability of the device itself need to be evaluated very carefully. In recent years, manufacturers of commercial BCI devices have been heavily investing in attractive and comfortable devices, making great strides in research.

Additionally, the system's innovativeness was recognized by seven participants. Feedback from two participants also pointed to the necessity of enhancing the *Graphical User Interface* in terms of colors and graphics in general. We welcome this observation, noting that in a future software upgrade, we plan to improve the interface, by implementing widgets specific to the music domain.

Regarding the assessment of *Slonimsky's grammar efficacy* in facilitating more accessible exploration of harmonic

and melodic domains, two participants observed that the system engendered unconventional melodies, while three noted its capacity to produce atypical polyphonies. These observations reinforce our hypothesis: the system acts as a catalyst in the discovery and analysis of novel compositional material, thereby aiding in the evolution of a more intricate and innovative compositional process.

The study reveals a noteworthy facet concerning the utilization of the *Leap Motion device*. Among the participants, four found it challenging to employ this device for crafting melodic patterns. Enhancements could be realized by moderating the tracking sensitivity to finger movements. Such adjustments would ensure a more stable generation of melodic patterns, effectively preventing minor finger movements from resulting in significant pitch alterations. This refinement would endow users with enhanced stability and control over their feedback.

The study elicited two critical responses concerning key aspects of the system. The first pertained to an inability for the user to “observe” the interplay between BCI technology and the generation of polyphony, specifically, understanding how various detected EEG signals are processed by AI algorithms to yield results. This challenge is acknowledged in the human-computer artificial intelligence (HCAI) field [59, 60], and underscores the broader issue of AI lack of explainability [61], which hinders user comprehension of AI-driven decisions. The second negative feedback concerned one participant's objection to the use of advanced computer systems in musical composition, perceived as overly alienating. This viewpoint highlights the intricate and sometimes contentious relationship between humans and technology, with some individuals viewing technology as a potential threat to

**Table 9** NeuralPMG aspects appreciated by participants

Theme	Frequency
Production of unconventional melodies	3
Overall system novelty	7
Power of generative grammar	2
Production of unconventional polyphonies	3
Use of BCI system	7
Use of Leap Motion device	3

**Table 10** NeuralPMG aspects criticized by participants

Theme	Frequency
Use of BCI system	5
Use of Leap Motion device	4
System graphical user interface	2
System appreciation	2
System efficiency	3

**Fig. 18** Maestro 1's evaluation of the polyphonies produced by the 19 participants

Section	Specific aspect	Participant																			Descriptive statistics			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	AVG	SD	AVG	SD
Aesthetic	Likeability	4	5	6	6	6	6	7	7	7	7	6	5	7	5	7	6	7	5	7	6.11	0.94		
	Elegance	6	6	6	6	6	6	7	7	7	7	7	5	7	7	6	7	5	7		6.42	0.69	6.32	0.78
	Consonance/Dissonance	5	6	6	6	6	6	7	7	7	7	5	7	7	7	6	7	6	7		6.42	0.69		
Innovativeness	Harmonic Interest	5	5	6	6	6	6	2	3	2	2	2	2	2	2	2	2	2	2		3.21	1.75		
	Harmonic Innovativeness	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2		2.11	0.46	2.37	1.03
	Innovativeness Vertical Intervallar Structures	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2		2.05	0.23		
	Innovativeness Horizontal Intervallar Structures	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2		2.11	0.32		
Elaboration	No Need to Adjust Parameters	2	5	6	5	5	5	6	6	6	6	6	6	6	6	7	6	7	6		5.74	1.10		
	Polyphony Thoroughness	2	5	5	5	5	5	5	4	5	4	4	5	5	5	4	6	4	5		4.63	0.83	4.18	1.84
	Polyphony Needs No Further Modification	6	3	4	3	3	2	1	2	2	1	2	2	1	2	1	2	1	2		2.16	1.26		
Usage	Music Improvisation Boosting	1	2	3	1	2	1	1	1	2	1	1	1	1	1	1	1	1	1		1.26	0.56		
	Music Composition Boosting	7	6	6	5	4	4	3	3	1	1	3	3	3	3	2	2	3	2		3.37	1.64	2.25	1.39
	Music Application Domain	1	1	1	3	2	1	3	2	2	2	2	1	3	2	3	2	3	3		2.11	0.81		
Emotion	Happiness	3	2	1	1	1	1	2	2	2	3	1	1	3	1	3	1	3	1		1.84	0.90		
	Sweetness	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1.00	0.00		
	Anger	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1.00	0.00		
	Sadness	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1.00	0.00		
	Fear	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1.00	0.00	1.61	0.96
	Negativity	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1.00	0.00		
	Arousal	5	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3		3.11	0.66		
	Positivity	3	2	1	1	1	1	1	1	1	2	1	1	2	1	3	1	1	1		1.42	0.69		
	Strength	4	2	3	3	4	3	3	3	3	3	3	3	3	3	3	3	3	3		3.11	0.46		
		Rating 7 6 5 4 3 2 1																						

human-centric activities. From our point of view, the problem can be related to the previous observation, i.e., the lack of awareness of computer processes, perceived as cryptic.

In response to these insights, we propose designing solutions that clarify the data flow management processes to the user, balancing transparency with the avoidance of information overload. Enhancements could include additional interfaces that explicitly outline the classification of mental states and the corresponding steps in generating rhythmic textures for each polyphonic voice. Such an approach aligns with HCAI principles advocating for AI systems to be reliable, safe, and trustworthy [62]. By making the back-end processes transparent, users gain both awareness and control over the system operations. Since this work is contextualized in the musical domain, it seems appropriate to place the sense of control alongside the Leitmotiv, a familiar tool in music and cinematography. The Leitmotiv, typically a melody representing a physical entity or character, serves as a perceptual cue, signalling a character's presence or actions, without the character being in the spectator's visual field. Leveraging the Leitmotiv concept, backend processes of the system could be

represented through distinct acoustic or visual signals, making the decision-making processes of the system perceptually evident to the user. This recognition would reduce uncertainty and stress associated with unseen backend activities.

The *evaluation by the two Maestros* provided valuable insights into the study. Regarding the *aesthetic quality* of the polyphonies, parameters like pleasantness and elegance are inherently personal assessments, rooted in each individual's interpretation of these terms. However, the Maestros concurred in their appreciation of the *aesthetic aspects* of the polyphonies created by the 19 participants, and they agreed on the consonance/dissonance of each piece, which is a more objective criterion.

In terms of *innovativeness*, however, the Maestros' opinions diverged, reflecting their distinct perspectives on what constitutes innovation and harmonic interest in polyphonies. Drawing from Vincent Persichetti's principles in "20th Century Harmony" treatise, the perception of musical intervals is influenced by their distribution within the composition. A dissonant interval can be defined as such if it is immersed in a context of consonant intervals. Conversely, if the entire

**Fig. 19** Maestro 2's evaluation of the polyphonies produced by the 19 participants

Section	Specific aspect	Participant																			Descriptive statistics			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	AVG	SD	AVG	SD
Aesthetic	Likeability	4	6	5	6	7	7	6	6	7	7	6	7	5	7	6	6	5	5	5	5.95	0.91		
	Elegance	4	7	5	6	7	6	5	6	7	6	5	7	5	6	6	6	4	5	5	5.68	0.95	5.95	0.91
	Consonance/Dissonance	5	7	6	7	7	6	6	7	7	7	6	7	6	7	6	6	4	6	5	6.21	0.85		
Innovativeness	Harmonic Interest	5	7	6	6	7	6	5	7	7	7	6	7	5	6	7	6	5	5	6	6.11	0.81		
	Harmonic Innovativeness	5	7	6	6	7	5	5	7	7	7	6	6	6	6	7	6	6	5	6	6.11	0.74	6.01	0.77
	Innovativeness Vertical Intervallar Structures	6	7	5	6	5	5	5	6	7	7	6	6	5	6	7	6	6	5	6	5.89	0.74		
	Innovativeness Horizontal Intervallar Structures	6	7	5	7	5	5	5	7	7	7	6	6	5	6	7	6	6	5	5	5.95	0.85		
Elaboration	No Need to Adjust Parameters	3	7	3	3	3	5	2	6	7	7	2	6	3	6	6	5	2	2	3	4.26	1.91		
	Polyphony Thoroughness	4	2	2	3	6	4	2	6	7	7	4	2	5	7	4	4	2	3	2	4.00	1.86	4.18	1.91
	Polyphony Needs No Further Modification	3	1	2	4	7	6	2	6	6	7	4	6	5	1	5	5	1	4	6	4.26	2.05		
Usage	Music Improvisation Boosting	5	7	6	6	7	5	2	4	7	7	5	6	5	7	6	5	2	4	5	5.32	1.53		
	Music Composition Boosting	6	7	7	6	7	5	3	7	7	7	5	7	6	7	7	6	3	5	6	6.00	1.29	5.88	1.36
	Music Application Domain	6	7	7	7	7	6	4	7	7	7	6	7	7	7	7	6	3	6	6	6.32	1.11		
Emotion	Happiness	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1.21	0.92		
	Sweetness	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1.05	0.23		
	Anger	1	1	1	1	1	1	1	2	5	4	2	2	1	4	7	1	2	1	3	2.16	1.71		
	Sadness	1	1	5	6	7	5	1	5	7	5	2	6	3	7	5	7	3	5	3	4.42	2.12		
	Fear	1	1	2	3	5	4	1	7	6	5	2	6	3	7	6	2	3	4	3	3.74	2.02	3.02	2.20
	Negativity	1	1	4	4	6	7	1	6	6	5	6	6	6	7	6	7	3	5	3	4.74	2.05		
	Arousal	5	5	2	2	1	1	6	4	6	7	5	5	3	5	7	1	5	5	7	4.32	2.06		
	Positivity	5	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1.26	0.93		
	Strength	6	7	5	2	3	3	2	2	6	7	2	5	3	6	7	4	3	5	4	4.32	1.83		
		Rating	7	6	5	4	3	2	1															



piece is based on dissonant intervals, consonant intervals are considered as dissonant. Given that both Maestros rated the polyphonies highly in terms of consonance/dissonance (averages of 6.42 and 6.21), this suggests a coherent harmonic structure. Maestro 1 viewed the polyphonies as not particularly innovative due to their harmonic alignment with traditional polyphonic styles. In other words, Maestro 1 did not consider disruptive the produced polyphonies. Conversely, Maestro 2 assessed innovativeness in light of Slonimsky's theory, focusing on the system ability to generate novel, interval, non-tonal structures. Their evaluations also differed in the *Elaboration* questionnaire section on the stimulation of composition and improvisation. Maestro 1's approach was more "vertical," considering the harmonic integration of melodic/polyphonic textures, while Maestro 2 adopted a more "contrapuntal" perspective, viewing each polyphony as an innovative overlap of independent melodies.

Regarding the need for further modification of the polyphonies, Maestro 1 felt minimal changes were necessary, consistent with his view of the coherence among the polyphonies. Maestro 2 believed many polyphonies required refinement, considering them as contrapuntal starting points for potential interval permutations aligned with Slonimsky's grammar.

The analysis of *Emotional perceptions* in the fourth section revealed diverse reactions, underscoring the subjective nature of musical interpretation. If one piece evokes tension for one musician and sadness for another, it will be utilized differently across various applications, such as in scoring a visual sequence or in pure composition. The differing emotional responses from the two domain experts affirm that the student-generated material indeed provokes a range of emotional reactions, highlighting its versatility and depth.

## Conclusions and Future Work

In this research work NeuralPMG, a framework to support the work of professional composers during the "compositional phase of music," has been presented. NeuralPMG is capable of calculating all possible interval permutations according to Slonimsky's grammar, as outlined in his treatise "Thesaurus of Scales and Melodic Patterns."

The evaluation study demonstrated that the framework can be an effective and useful tool for music composition. However, a limitation to the validity of the work is that the framework was only tested in a laboratory setting. Therefore, a longitudinal study with field experiments involving domain experts during the production phase of musical compositions in a real-world context would be desirable. This would also assess the timing of music production and the diversity of materials produced.

NeuralPMG shows potential for a numerous future works. From a hardware perspective, the type of BCI used is a major concern, as it is crucial in determining whether a musician would choose to purchase and use such a device. Currently, there are various off-the-shelf BCIs available, some of which are unobtrusive, comfortable, and still effective. In the study, a five-electrode device was used, as it was the one available in our laboratory. However, since the system only needs to classify between two states, namely Focused and Relaxed, a lightweight single-electrode headset, like the one offered by the Neurosky with a single frontal electrode, would suffice.

An interesting area for future investigation is expanding the current set of metrics. Interest could control an additional parameter of the music generative process, meanwhile monitoring user's *stress*. The detection and processing of *emotions* by ML algorithms can be considered in polyphony generation.

The overarching goal is to integrate NeuralPMG into an ecosystem of products useful to musicians/composers who utilize computer tools to enhance their work and connect with other art forms. In this direction, NeuralPMG could be integrated with other systems that consider mental states, interest, emotions, stress, and body movements detected by sensing systems like *OpenCV* or *Kinect*.

## Glossary

- **Augmented fourth interval:** Distance between two three-tone sounds.
- **Cents:** Cent is a logarithmic unit of measure used for musical intervals.
- **Chromatic variation:** Ability that a sequence of notes has to change its interval structure by altering the notes by a semitone or tone.
- **Heptatonic scales:** Scales having seven notes per interval of symmetrical repetition. In the case of scales made within the octave interval it is a scale for seven notes in an octave.
- **Infrapolation:** Term coined by Slonimsky to describe the insertion of a note below a reference interval progression.
- **Interpolation:** Term used by Slonimsky to describe the insertion of a note within a reference interval progression.
- **Interval axis:** Term that describes the imaginary line joining the notes of an interval progression theorized by Slonimsky.
- **Melody:** Succession of musical notes with complete meaning, with their own pitch and rhythm.
- **Octave:** Interval between two sounds, such that the higher (high) one has twice the frequency of the lower (low) one.

- **Pentatonic scales:** Scales having five notes per interval of symmetrical repetition. In the case of scales made within the octave interval, it is a scale for five notes in an octave.
- **Polymodality:** Term used in functional harmony to describe the superposition of multiple modes in different octaves in succession.
- **Polyphony:** In music, the union of several voices each performing its own melodic pattern.
- **Retrograde:** Procedure derived from contrapuntal technique to describe the reading of a series from the last note to the first.
- **Seriality:** Term coined by Schönberg to describe his compositional system based on the concept of seriality, that is, the arrangement of the chromatic scale in the different octaves and all the compositional rules about it.
- **Symmetrical repetition of interval scale structures:** This term refers to the point at which a scalar or interpolation pattern repeats symmetrically from that point-note to the next. In scales built on the “octave interval,” the octave interval represents the point of symmetrical repetition of the scale itself. In a scalar model based on an augmented fourth interval axis, the augmented fourth interval represents the point of symmetrical re-proposition of the scalar model.
- **Ultraposition:** Term coined by Slonimsky to describe the insertion of a note above a reference interval progression.
- **Unison:** Simultaneous sounds with equal pitch.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12559-024-10280-6>.

**Acknowledgements** This work was partially supported by the following projects: Italian P.O. Puglia FESR 2014 – 2020 (project code 6ESURE5) “SECURE SAFE APULIA,” Fincons CdP3, PASS-PARTOUT, Servizi Locali 2.0, ERP4.0. CTEMT - “Casa delle Tecnologie Emergenti di Matera.” This work has been carried out while Tommaso Colafiglio was enrolled in the Italian National Doctorate on Artificial Intelligence run by Sapienza University of Rome in collaboration with Politecnico di Bari.

**Author Contributions** Conceptualization, investigation, methodology, writing—original draft: Tommaso Colafiglio, Carmelo Ardito, Paolo Sorino and Domenico Lofù. Data curation, investigation, writing—review and editing: Tommaso Colafiglio, Carmelo Ardito, Paolo Sorino and Domenico Lofù. Supervision, writing—review and editing: Carmelo Ardito, Eugenio Di Sciascio, Tommaso Di Noia and Fabrizio Festa.

**Funding** Open access funding provided by Politecnico di Bari within the CRUI-CARE Agreement. Not applicable.

**Availability of Supporting Data** The experimental data and the simulation results that support the findings of this study are available on request.

## Declarations

**Ethics Approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

**Informed Consent** Informed consent was obtained from all individual participants included in the study.

**Competing Interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Mozarts. Mozarts musikalisches würfelspiel. 1787. Available at: <https://www.ensembleresonanz.com/task/mozarts-musikalisches-wurfelspiel> Last visit: January 2024.
2. Yearsley, D. Bach and the Meanings of Counterpoint, volume 10, Cambridge University Press. 2002.
3. Kerman J. The art of fugue: Bach fugues for keyboard, 1715-1750, University of California Press. 2005.
4. Colafiglio T. Dalle teorie compositive di Slonimsky ad una nuova impostazione dell’armonia, Euterpe Series, il Grillo Editore. 2011.
5. Giommoni M. Gli algoritmi della musica - Composizione e pensiero musicale nell’era informatica, Musica, Discipline dello spettacolo Series, CLEUP sc - Cooperativa Libreria Editrice Università di Padova, Italy. 2011.
6. Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. Clin Neurophysiol. 2002;113:767–91.
7. Ardito C, Bortone I, Colafiglio T, Di Noia T, Di Sciascio E, Lofù D, Narducci F, Sardone R, Sorino P. Brain computer interface: Deep learning approach to predict human emotion recognition. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE. 2022; pp. 2689–2694.
8. Houssein EH, Hammad A, Ali AA. Human emotion recognition from EEG-based brain-computer interface using machine learning: a comprehensive review. Neural Comput Appl. 2022;34:12527–57.
9. Bhiri NM, Ameer S, Alouani I, Mahjoub MA, Khalifa AB. Hand gesture recognition with focus on leap motion: An overview, real world challenges and future directions. Exp Syst Appl. 2023;120125.

10. Yi H. Efficient machine learning algorithm for electroencephalogram modeling in brain–computer interfaces, *Neural Comput Appl*. 2022; pp. 1–11.
11. Verdi L. *Organizzazione delle altezze nello spazio temperato*, vol. 4. Italy: Diastema Editrice; 1998.
12. Forte A. *The Structure of Atonal Music*, Yale University Press. 1977.
13. Schoenberg A. *Theory of harmony*, University of California Press. 1983.
14. Bell C. Algorithmic music composition using dynamic markov chains and genetic algorithms. *J Comput Sci Colleges*. 2011;27: 99–107.
15. Burraston D, Edmonds E. Cellular automata in generative electronic music and sonic art: a historical and technical review. *Digital Creativity*. 2005;16:165–85.
16. Santoboni R, Clementi M, Ticari AR. Virtual composer: a tool for musical composition. In: *Proceedings of the International Conference on Web Delivering of Music (WEDELMUSIC)*, IEEE. 2002; p. 228.
17. Kiloh LG, McComas AJ, Osselton JW. *Clinical electroencephalography*. Butterworth-Heinemann. 2013.
18. Preece J, Rogers Y, Sharp H, Benyon D, Holland S, Carey T. *Human-computer interaction*, Addison-Wesley Longman Ltd., 1994.
19. Hartson R, Pyla PS. *The UX book: Agile UX design for a quality user experience*, Morgan Kaufmann. 2018.
20. Collura TF. History and evolution of electroencephalographic instruments and techniques. *J Clin Neurophysiol*. 1993;10: 476–504.
21. Forcucci L. Music for brainwaves: Embodiment of sound, space and EEG data. *Body Space Technol*. 2018;17.
22. Blum T. Review of the Biofeedback and the Arts, Results of Early Experiments book. *Comput Music J*. 1989;13:86–8.
23. Cage J. *John Cage: an Anthology*, Da Capo Press. 1991.
24. Rosenboom D. *Biofeedback and the Arts*. Aesthetic Research Centre of Canada: Results of Early Experiments; 1976.
25. Gilden L. Instrumental control of EEG alpha activity with sensory feedback. In: Rosenboom D, editor. *Biofeedback and the Arts*. Results of Early Experiments: Aesthetic Research Center of Canada Publications; 1976.
26. Oliveros P. *Sonic meditations*, Smith Publications. 1974.
27. Sanyal S, Nag S, Banerjee A, Sengupta R, Ghosh D. Music of brain and music on brain: a novel EEG sonification approach. *Cogn Neurodyn*. 2019;13:13–31.
28. Teitelbaum R. In tune: Some early experiments in biofeedback music (1966–1974). In: Rosenboom D, editor. *Biofeedback and the Arts*. Results of Early Experiments: Aesthetic Research Center of Canada Publications; 1976.
29. Lucier A, Simon D. *Chambers: Scores by Alvin Lucier*, Wesleyan University Press. 2012.
30. Straebel V, Thoben W. Alvin lucier's music for solo performer: experimental music beyond sonification. *Organised Sound*. 2014;19:17–29.
31. Knapp RB, Lusted HS. A bioelectric controller for computer music applications. *Computer music journal*. 1990;14:42–7.
32. Brouse A. *The Interharmonium: an investigation into networked musical applications and brainwaves*. Master of Arts thesis, McGill University. 2001.
33. Miranda ER, Brouse A. Interfacing the brain directly with musical systems: on developing systems for making music with brain signals. *Leonardo*. 2005;38:331–6.
34. Hamadicharef B, Xu M, Aditya S. Brain-computer interface (BCI) based musical composition. In: *International Conference on Cyberworlds*, IEEE, 2010; pp. 282–286.
35. Wang M, Daly I, Allison BZ, Jin J, Zhang Y, Chen L, Wang X. A new hybrid BCI paradigm based on P300 and SSVEP. *J Neurosci Methods*. 2015;244:16–25.
36. Folgieri R, Zichella M. A BCI-based application in music: Conscious playing of single notes by brainwaves. *Comput Entertain*. 2012;10:1–10.
37. Pinegger A, Wriessnegger SC, Müller-Putz GR. Sheet music by mind: Towards a brain-computer interface for composing, in: *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE. 2015; pp. 1053–1056.
38. Deuel TA, Pampin J, Sundstrom J, Darvas F. The encephalophone: A novel musical biofeedback device using conscious control of electroencephalogram (EEG). *Front Human Neurosci*. 2017;11: 1–8.
39. Ehrlich SK, Agres KR, Guan C, Cheng G. A closed-loop, music-based brain-computer interface for emotion mediation. *PLoS one*. 2019;14: e0213516.
40. Croassacipto M, Ichwan M, Utami DB. Tone classification matches kodály handsign with the k-nearest neighbor method at leap motion controller. *Int J Inform Commun Technol*. 2019;5:40–5.
41. Polansky L. Interview with David Rosenboom. *Comput Music J*. 1983;7:40–4.
42. Baratè A, Ludovico LA, Oriolo E. Investigating embodied music expression through the leap motion: Experimentations in educational and clinical contexts. In: *Proceedings of the International Conference on Computer Supported Education (CSEDU)*, Springer. 2018; pp. 532–548.
43. Homan RW, Herman J, Purdy P. Cerebral location of international 10–20 system electrode placement. *Electroencephalogr Clin Neurophysiol*. 1987;66:376–82.
44. Bronzino JD, Peterson DR. *Principles of electroencephalography*, in: *Biomedical Engineering Fundamentals*, CRC press. 2006; pp. 445–456.
45. Lera L. Grammatica della notazione di notre-dame. *Acta Musicol*. 1989;61:150–74.
46. Gao J, Qi H, Xia X, Nie J-Y. Linear discriminant model for information retrieval. In: *Proceedings of the International Conference on Research and Development in Information Retrieval (ACM SIGIR)*. 2005; pp. 290–297.
47. Charbuty B, Abdulazeez A. Classification based on decision tree algorithm for machine learning. *J Appl Sci Technol Trends*. 2021;2:20–8.
48. Rish I, et al. An empirical study of the naive bayes classifier, in: *Workshop on Empirical Methods in Artificial Intelligence*. In *Conjunction with International Joint Conference on Artificial Intelligence*. 2001; pp. 41–46.
49. Yang Y, Li J, Yang Y. The research of the fast svm classifier method. In: *Proceedings of the International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, IEEE. 2015; pp. 121–124.
50. Sun S, Huang R. An adaptive k-nearest neighbor algorithm, in: *Proceedings of the International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, IEEE. 2010; pp. 91–94.
51. Svozil D, Kvasnicka V, Pospichal J. Introduction to multi-layer feed-forward neural networks. *Chemom Intell Lab Syst*. 1997;39:43–62.
52. Hassenzahl M, Burmester M, Koller F. Attrakdiff: Ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität, *Mensch & Computer*. *Interaktion in Bewegung*. 2003;2003:187–96.
53. Cherry E, Latulipe C. Quantifying the creativity support of digital tools through the creativity support index. *ACM Trans Comput Human Interact*. 2014;21:1–25.

54. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv Psychol*. 1988;52:139–83.
55. O'Brien H. Theoretical perspectives on user engagement. In: O'Brien H, Cairns P, editors. *Why engagement matters: Cross-disciplinary perspectives of user engagement in digital media*. Springer. 2016; pp. 1–26.
56. O'Brien HL, Cairns P, Hall M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *Int J Human Comput Stud*. 2018;112:28–39.
57. Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry*. 1994;25:49–59.
58. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3:77–101.
59. Schmidt A. Interactive human centered artificial intelligence: a definition and research challenges. In: *Proceedings of the International Conference on Advanced Visual Interfaces (AVI)*. 2020; pp. 1–4.
60. Shneiderman B. *Human-centered AI*, Oxford University Press. 2022.
61. Došilović FK, Brčić M, Hlupić N. Explainable artificial intelligence: A survey. In: *Proceedings International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, IEEE. 2018; pp. 210–215.
62. Shneiderman B. Human-centered artificial intelligence: Reliable, safe & trustworthy. *Int J Human Comput Interact*. 2020;36: 495–504.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.