UNIVERSITY *of York*

This is a repository copy of *On the use of AI for Generation of Functional Music to Improve Mental Health*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/170412/

Version: Published Version

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

On the use of AI for Generation of Functional Music to Improve Mental Health

Duncan Williams[1,2,3]*, Victoria J. Hodge[1,2] and Chia-Yu Wu[4]

[1]Digital Creativity Labs, University of York, York, United Kingdom, [2]Department of Computer Science, University of York, York, United Kingdom, [3]School of Science, Engineering and Environment, University of Salford, York, United Kingdom, [4]Department of Electronic Engineering, University of York, York, United Kingdom

Increasingly music has been shown to have both physical and mental health benefits including improvements in cardiovascular health, a link to reduction of cases of dementia in elderly populations, and improvements in markers of general mental well-being such as stress reduction. Here, we describe short case studies addressing general mental well-being (anxiety, stress-reduction) through AI-driven music generation. Engaging in active listening and music-making activities (especially for at risk age groups) can be particularly beneficial, and the practice of music therapy has been shown to be helpful in a range of use cases across a wide age range. However, access to music-making can be prohibitive in terms of access to expertize, materials, and cost. Furthermore the use of existing music for functional outcomes (such as targeted improvement in physical and mental health markers suggested above) can be hindered by issues of repetition and subsequent over-familiarity with existing material. In this paper, we describe machine learning approaches which create functional music informed by biophysiological measurement across two case studies, with target emotional states at opposing ends of a Cartesian affective space (a dimensional emotion space with points ranging from descriptors from relaxation, to fear). Galvanic skin response is used as a marker of psychological arousal and as an estimate of emotional state to be used as a control signal in the training of the machine learning algorithm. This algorithm creates a non-linear time series of musical features for sound synthesis "on-the-fly", using a perceptually informed musical feature similarity model. We find an interaction between familiarity and perceived emotional response. We also report on subsequent psychometric evaluation of the generated material, and consider how these - and similar techniques - might be useful for a range of functional music generation tasks, for example, in nonlinear sound-tracking such as that found in interactive media or video games.

Keywords: mental health, emotional states, feedback, biophysiological sensors, generative music, machine learning, artificial intelligence, algorithmic composition

INTRODUCTION

There is increasing evidence that mindfulness can form a positive contributor to mental health and general wellbeing (Baker and Bor, 2008; Economides et al., 2018). In this work we describe the design and evaluation of a system combining machine learning (ML) approaches with biophysiological metering and psychological evaluation of two descriptors which we consider to be at discrete ends of an affective space with positive mental health states at one side of the space (mindfulness, calmness, etc.), and negative mental states at the other side of the space (fear, anger, etc.) (Chambers et al., 2009).

The distinction between affective state, emotion, and mood, is complex, and is generally drawn between the duration of the response (Calvo et al., 2009). Various models of affective state exist, including models with dimensions for positivity and activation strength, such as the cirumplex model of affect (Russell, 1980). This model places valence (as a measure of positivity) and arousal (as a measure of activation strength) on the horizontal and vertical axes respectively. Emotional descriptors (e.g., happy, sad, angry) can be mapped on to this space. Other models exist, for example, multidimensional models which also include, for example, dimensions for dominance. This type of model might be useful when delineating between very intense and very negative emotional descriptors, such as the difference between anger and fear–both intense, and negative, but one being a more dominant response and the other more passive. Often, individual emotional descriptors can be plotted across these types of spaces (Williams et al., 2014). In the case of this work, we consider a general model with two specific descriptors as approximately at either end of a scale–mindful, and afraid. However, the descriptors themselves are open to debate and could certainly form the subject of further work. We intend to explore the use of AI to generate music intended to elicit differing emotional states in an abstract emotional space and to examine biophysiological markers in a synchronous manner.

Existing work has shown that there are responses to music in both the central and peripheral nervous system (in other words, both physiological, and neurological responses) (Aldridge, 2005; Calvo et al., 2009). When listening to enjoyable music, the listeners pupils may dilate, or they might experience a change in heart rate, blood pressure, and skin conductivity (Blood et al., 1999; Daly et al., 2015). Measurement of galvanic skin response (GSR) has been shown to be a robust metric for analysis of emotional responses to music (Shrift, 1954; Vanderark and Ely, 1993; Daly et al., 2015).

Thus, there is a potential crossover between mental state, physiological reaction, and musical stimulation. Chambers (Chambers et al., 2009) showed that states of mindfulness have correlations in GSR (otherwise known as electrodermal activity, or skin conductivity), heart rate variability, and the ratio of alpha and beta waves in electroencephalographic measurement. The electroencephalograph (EEG) is a technique for metering electrical activity from the scalp used to infer patterns of brain activity. Bondolfi (Bondolfi, 2013) and Economides et al. (Economides et al., 2018) proposed that proactive training and entrainment of mental states might
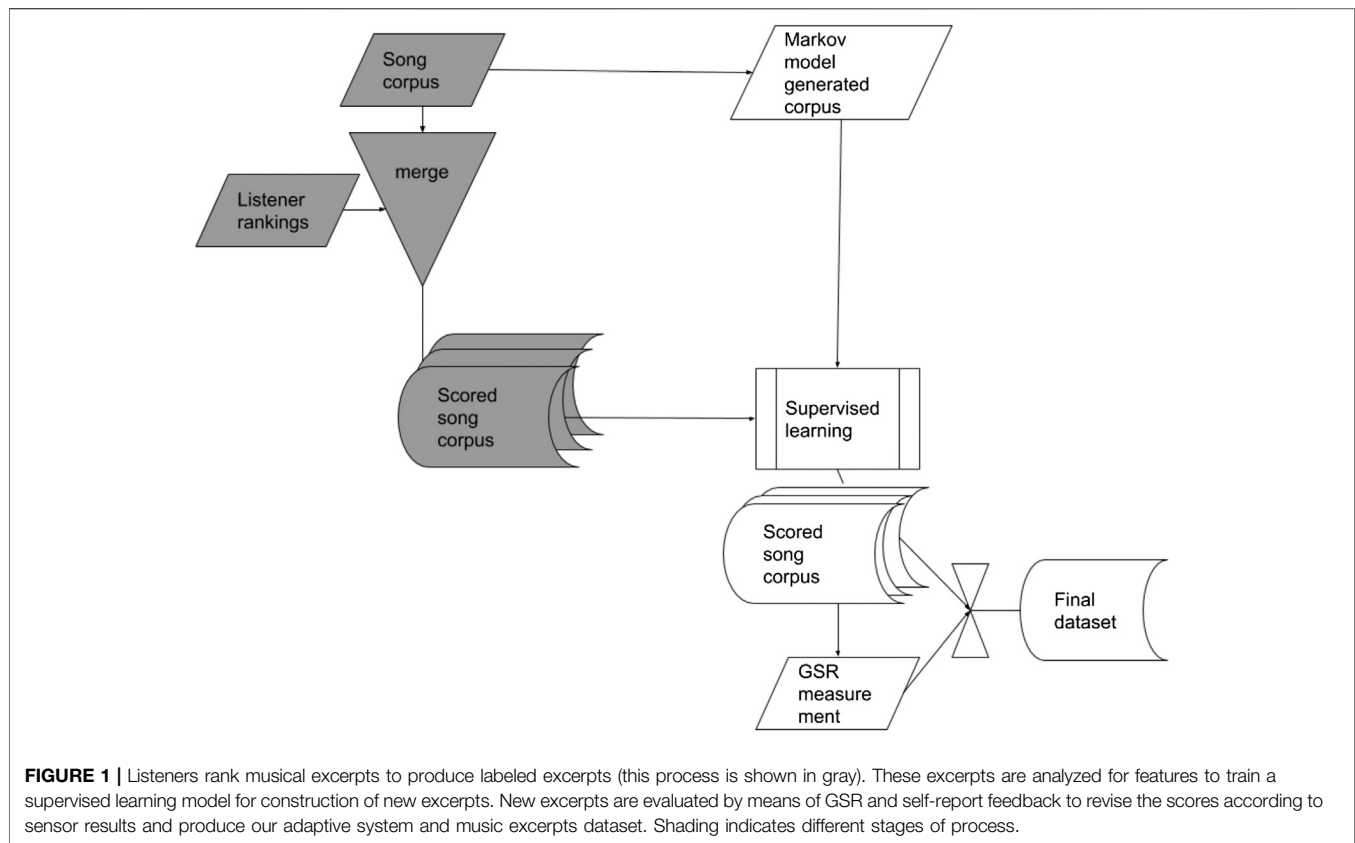
thus contribute to therapeutic treatment and physiological improvement.

We aim to harness these findings to create a machine learning based music-training system to encourage a change in affective state as measured through biophysiological correlations. For example, mood-based regulation (becoming less afraid or anxious) might be a useful mental health target for the user. Beyond mental health this type of system could have applications in the creative industries, for example, in film, television, or video games (Knox et al., 2008; Williams et al., 2015a), in which case, the viewer or player might be subjected to targeted mood disruption (i.e., rather than being calmed by the musical stimulus, the listener might preferably be deliberately excited, or even scared in the case of some types of gameplay).

In this paper we draw on previous experimental work documented in (Williams et al., 2019a; Williams et al., 2019b), and more widely, machine learning, a field of computer science covering systems that learn "*when they change their behavior in a way that makes them perform better in the future*" (Witten et al., 2016). These systems learn from data without being specifically programmed. Kim et al. (Kim et al., 2010) and Laurier and Herrera (Laurier and Herrera, 2012) give a literature overview of detecting emotion in music and focus on the music representations. Laurier and Herrera also provide an analysis of the machine learning algorithms used by systems in their survey. Classification algorithms used in the literature include C4.5, Gaussian mixture models, k-nearest neighbor, random forest, support vector machines, (Kim et al., 2010; Laurier and Herrera, 2012; Mostafavi et al., 2013). Regression techniques include Gaussian mixture model regression, multiple linear regression, partial least-squares regression and support vector regression.

ML has been used to retrieve music by mood and ML analyses found the personalized approach more consistent than a general approach (Mostafavi et al., 2013). An example is supervised learning. In supervised learning, the algorithm learns from a set of labeled inputs. It then generates a model associating the inputs with their respective labels or scores, and then classifies (or predicts) the likely label for previously unseen examples using the learned model. We use supervised learning to label newly generated material with potential emotional descriptors in the work documented in this paper.

Real-world testing of systems using bio-signal mappings in music generation contexts has become an emerging field of research, partly due to recent advances in portability, wearability, and affordability of biosensors. For example, Huang and Cai (Huang and Cai, 2017) generate simple music for specific emotional states using Markov chains. The Markov chains are used to generate music while the user wears a heart-rate sensor to monitor their bio-physiological response to the created music. The system was able to generate emotionally responsive music in a limited trial considering basic emotional descriptors. We have developed another such system, which assumes lower skin conduction variability as a correlate for positive affective state. It attempts to generate emotionally congruent music as a training tool to promote positive affective states in the context of mindfulness. In the future,

FIGURE 1 | Listeners rank musical excerpts to produce labeled excerpts (this process is shown in gray). These excerpts are analyzed for features to train a supervised learning model for construction of new excerpts. New excerpts are evaluated by means of GSR and self-report feedback to revise the scores according to sensor results and produce our adaptive system and music excerpts dataset. Shading indicates different stages of process.

this system could also work in reverse by using skin conductance variability as a control signal to inform musical feature mapping for non-linear music generation.

While the physical and mental health benefits of music use have increasingly been reported upon (including improvements in cardiovascular health (Szmedra and Bacharach, 1998)), reduction of dementia in elderly populations (Vink et al., 2003), stress reduction (Knight and Rickard, 2001) and so on), the use of existing music to target such outcomes can be problematic due to the influence of familiarity, or repetition of stimulus materials (Kim, 2011; Ladinig and Schellenberg, 2012). Thus a major focus of this work is to evaluate a system for the automatic generation of new musical materials with functional aims (improvement of listener affective state on a case-by-case basis as determined by self-report or biophysiological correlate). We therefore aim to:

(1) Measure musical features according to a similarity model from the human-labeled dataset and use these to inform Markov-model generation of new music to be evaluated by a supervised learning algorithm
(2) Evaluate the success of the supervised learning algorithm using self-report and biophysiological measurement of GSR

We hypothesize that music generated by the automated algorithm may be able to influence self-reported emotional state and GSR, when compared to music which listeners may already be familiar with from a corpus of popular film music.

MATERIALS AND METHODS

GSR is used as a marker of psychological arousal and as an estimate of emotional state to be used as a control signal in the training of the ML algorithm. This algorithm creates a non-linear time series of musical features for sound synthesis "on-the-fly," using a perceptually informed musical feature similarity model.

We use the system described in (Williams et al., 2017) to create functional music informed by biophysiological measurement across two case studies, with target emotional states at opposing ends of a Cartesian affective space (a dimensional emotion space with points ranging from descriptors from relaxation, to fear).

The system detects the user's current emotional level and the ML algorithm picks musical pieces to influence their future emotional level to achieve their desired mood. This whole process requires musical pieces that have an associated emotional label (score) to allow the selection of appropriate pieces. We use two tasks to achieve this, illustrated in **Figure 1**. Firstly we analyze a corpus for musical features using classification and regression. We use a multi-feature music representation combining analysis of symbolic musical feature data from a MIDI file, which represents the structure of

the melody, chords, rhythm and other musical properties concerning timing, dynamics, and pitch, with Mel-Frequency Cepstral Coefficients features (Logan, 2000) obtained from the entire piece to represent the timbral quality of the instrumentation. This dual representation is more flexible and richer than either MIDI or signal-based audio content features alone. We only use numerical data features to describe each piece and perform feature selection to identify the most significant set of features as described in (Hodge et al., 2016). Using this reduced set of significant features, the ML model can predict the likely emotional state score that a human listener would ascribe to newly input music pieces by determining the similarity between pieces using their respective sets of features.

We train the supervised learning algorithm to expand on a human-labeled corpus, which was labeled by means of a survey of 53 participants using a Qualtrics on-line survey (www.qulatrics. com). Each participant evaluated four musical excerpts, two calm or positive (N1 and N2) and two anxious or negative excerpts (S1 and S2), in a bipolar ranking across six pairs choosing the positive in each pair {N1vsS1, S2vsN2, S1vsN2, N1vsS2, S1vsS2 and N1vsN2}. The survey presented an initial question to allow the user to familiarize themselves with the format and then presented the six questions. The Qualtrics questionnaire allowed us to specify that each track played in full to each participant to ensure that the participant adapted fully to the track. We randomized the order of presentation of the questions (pairs of tracks) to each of the participants to reduce contextual effects. Participants were not required to answer every question in order to complete the evaluation. The algorithmic composition system uses hidden Markov models (HMM) to create new music to provide sufficient quantities of labeled pieces for the system to operate. We use a transformative algorithm based on a second order Markov-model with a musical feature matrix. New material is formed of permutations of the HMM with deliberate feature constraints following the procedure described in (Williams et al., 2015b; Williams et al., 2017). This allows discrete control over five parameters in a 2-dimensional model.

Human experiments are only feasible on a small set of music pieces as n pieces of music require $n!$ comparisons and enough human survey participants to provide enough responses for each of the $n!$ comparisons. Using human participants to generate a sufficiently large database of labeled pieces for this work would be very time consuming and complex. From our Qualtrics analyses we were able to train our ML model using supervised learning to map the musical sequences to scores where the sequences are represented by features as described previously. Our generative system can be used to create new musical sequences according to the likelihood of a particular affective state occurring after the current and preceding states measured in the listener and these can be scored by the trained ML model.

We then analyze the listener's GSR to select music which exercises the most influence of the listener's affective state. We incorporate a feedback loop to adapt the corpus scores according to the user's affective response, selecting musically consistent pieces and removing pieces that do not influence the user's emotional level (in essence a fitness function). We then compare the listener's GSR signal, the emotional tag they describe after listening and the calmness level of the piece the participant is listening to. To analyze GSR, we used the Shimmer3 wireless GSR + Unit[1] which has been validated for use in biomedical-oriented research applications. This device needs to be calibrated on each use to establish a baseline skin conductance signal, which varies due to many factors including skin dryness, nervousness (due to unfamiliarity with the experimental procedure) and ambient temperature. The captured reading for each user under analysis is their skin conductance response while undertaking the listening exercise, minus their individual skin conductance response baseline. After listening to each piece, the users completed a questionnaire describing the emotion they felt while listening which we compared to the GSR data.

RESULTS

Responses to the musical stimuli suggest that listeners found it relatively easy to discriminate the affective states between stimuli, which were rendered using different synthesized timbres. Generally speaking, shorter durations and larger pitch ranges were considered lower in positivity (for example, "more anxious") than longer durations with a more restricted pitch range, regardless of the musical timbres. 58.1% of participants thought S2 was more negative than S1 while 54.6% felt N1 was more negative than N2.

For S1, 94.5 and 93.2% of participants rated it more negatively than N1 and N2 respectively. For S2, 88.1 and 89.1% of participants rated it more negatively than N1 and N2 respectively. Yet, 58.1% of the participants rated S2 more negatively than S1 despite S2 having created more positive report than S1 when compared to the positive stimuli. Similarly, for N1, 4.6 and 10.9% rated it more negatively than S1 and S2, respectively, while for N2, 6.8 and 11.9% rated it more negatively than S1 and S2, respectively. This presents a similar contradiction as for the negative stimuli as N1 has lowest reported positivity yet was rated more negatively than N2 by 54.6% of participants.

From these comparisons, we were able to attain sufficient data that we can calculate a ranked order (score) for the pieces from these pairwise comparisons (Wauthier et al., 2013). In order for the biofeedback based evaluation to be feasible, we then use the supervised learning generated corpus to provide a large enough quantity of stimuli.

Our analyses revealed that there is a direct correlation between the reported negativity of a musical piece, the user's GSR readings and the emotions they describe feeling in a questionnaire survey conducted after listening. Users display elevated GSR for negative pieces which they also labeled congruently in the questionnaire and lower GSR and appropriate labels for calmer pieces. We also find an interaction between familiarity of existing material in the corpus, and the perceived emotional response.

---

[1]http://www.shimmersensing.com/products/shimmer3-wireless-GSRsensor

DISCUSSION

Our experiments highlighted that familiarity influences individual affective responses both in self-report and in GSR. For this reason, we have attempted to focus on generating novel music to create functional music which responds to a listener's biophysiological state rather than invoked or evoking memories (and removing some need to consider the influence of familiarity might affect listener responses).

Overall, we saw an increased GSR in each music excerpt, regardless of whether the excerpts were generated by the HMM model or not. We conclude that GSR is a suitable detection tool to evaluate emotional responses. DES-based self-report was used to allow listeners to report on different categories of emotions (Lane et al., 1990). However, the two measurements do not have consistent results when considered in response to music that listeners described themselves as being familiar with (samples of famous film music). Nevertheless, the emotional responses to generated music excerpt *g1* showed consistent results with both self-reporting and GSR. Thus we consider there may be an interaction between music and familiarity (perceived emotions). In self-reports, familiarity has insignificant effects. Conversely, in GSR data, there are differences in the simple effect of music between unfamiliar and familiar tracks. Familiar movie soundtracks also have higher GSR amplitude than unfamiliar ones but lower negative self-reports.

Hence, to induce calm states of mind reliably we believe further work should focus on unfamiliar music composed using artificial intelligence based approaches. Our main aim for this work is to develop a music generator for music therapy use that produces music which induces specific emotions in the listener but the approach described here might also be suitable in the design of a more generic music generator capable of inducing specific emotions in the audience, specifically when functional music with non-linear duration would be useful (e.g., videogame sound-tracking and the creative industries more broadly construed).

CONCLUSIONS AND FURTHER WORK

This work suggests that generative music technology has the potential to produce infinite soundtracks in sympathy with a listener's bio-signals, and in a biofeedback loop. There are promising applications for linking music with emotions, especially in the creative industries art and therapy, and particularly for relaxation. Enhancement of well-being using music and the emotions music induces is becoming an emerging topic for further work. We have applied a system for musical feature analysis from MIDI features and Mel-Frequency Cepstral Coefficients features (Logan, 2000) to train a supervised learning algorithm with listener responses to a corpus of training material. We use this algorithm to influence the generation of a larger corpus by means of a Hidden Markov Model algorithmic composition engine, and then analyzed the complete corpus by testing listener GSR and self-report in a DES evaluation. GSR is used as a marker of psychological arousal and as an estimate of emotional state to be used as a control signal in the training of the ML algorithm. This algorithm creates a non-linear time series of musical features for sound synthesis "on-the-fly", using a perceptually informed musical feature similarity model. These small case studies, with target emotional states at opposing ends of a Cartesian affective space (a dimensional emotion space with points ranging from descriptors from positive descriptor states such as calmness, to negative descriptors such as fear), show us an interaction between familiarity and perceived emotional response. We believe further work involves three major challenges.

(1) The extraction of meaningful control information from signals emanating from the body.
(2) Design of generative and performative music technology in order to respond to such information.
(3) Consideration of the ways in which such technology can be best deployed depending on the intended end-use; for example, in therapeutic contexts.

There is a tendency in human-computer interaction work for music generation to prioritize the technical implementation by focusing on increased speed or accuracy of a system, rather than the specific needs of the application. In a music therapy context, for example, one advantage of a functional music system is that it might be used by a patient with no musical ability and thereby potentially increases their own ability to express emotional states and have access to the pleasure of performing music with other people. Thus the use of biophysiological sensors is critical in the development of suitable systems for audio generation in the context of mindfulness or relaxation where improved affective state as part of mental health is an intended outcome. These are not trivial considerations in terms of application design, and subsequent evaluation. Methodologies for evaluating the success or failure of such systems remain a significant challenge for further work.

DATA AVAILABILITY STATEMENT

The datasets analyzed in this article are not publicly available. Requests to access the datasets should be directed to d.a.h.williams@salford.ac.uk.

ETHICS STATEMENT

The experiment was conducted with ethical approval from the University of York, Dept of Electronic Engineering review board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

DW and VH contributed conception and design of the study; DW composed the musical sequence database; CW led the human

analyses under guidance from DW and VH, VH developed the Qualtrics survey, CW and VH performed the statistical analyses; DW wrote the first draft of the article; VH and CW wrote sections of the article. All authors contributed to article revision, read and approved the submitted version.

REFERENCES

Aldridge, D. (2005). *Music therapy and neurological rehabilitation: performing health*. London UK: Jessica Kingsley Publishers, 272.

Baker, F., and Bor, W. (2008). Can music preference indicate mental health status in young people? *Australas. Psychiatr.* 16 (4), 284–288. doi:10.1080/10398560701879589.

Blood, A. J., Zatorre, R. J., Bermudez, P., and Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci.* 2 (4), 382–387. doi:10.1038/7299.

Bondolfi, G. (2013). [Depression: the mindfulness method, a new approach to relapse]. *Rev. Med. Suisse.* 9 (369), 32–39. doi:10.1176/appi.focus.20170039.

Calvo, R. A., Brown, I., and Steve, S. (2009). "Effect of experimental factors on the recognition of affective mental states through physiological measures," in Australasian joint Conference on artificial intelligence (Berlin, Heidelberg: Springer), 62–70.

Chambers, R., Gullone, E., and Allen, N. B. (2009). Mindful emotion regulation: an integrative review. *Clin. Psychol. Rev.* 29 (6), 560–572. doi:10.1016/j.cpr.2009.06.005.

Daly, I., Malik, A., Weaver, J., Hwang, F., Nasuto, S. J., Williams, D., et al. (2015). "Toward human-computer music interaction: evaluation of an affectively-driven music generator via galvanic skin response measures," in 7th computer science and electronic engineering conference, September 24–25, 2015. Colchester, UK: IEEE, 87–92. doi:10.1109/CEEC.2015.7332705.

Economides, M., Martman, J., Bell, M. J., and Sanderson, B. (2018). Improvements in stress, affect, and irritability following brief use of a mindfulness-based smartphone app: a randomized controlled trial. *Mindfulness* 9 (5), 1584–1593. doi:10.1007/s12671-018-0905-4.

Hodge, V. J., O'Keefe, S., and Austin, J. (2016). Hadoop neural network for parallel and distributed feature selection. *Neural Network.* 78, 24–35. doi:10.1016/j.neunet.2015.08.011.

Huang, C.-F., and Cai, Y. (2017). "Automated music composition using heart rate emotion data," in International conference on intelligent information hiding and multimedia signal processing (Cham, Switzerland: Springer), 115–120.

Kim, J. (2011). Affective states, familiarity and music selection: power of familiarity. *Int. J. Art Tech.* 4 (1), 74–89. doi:10.1504/ijart.2011.037771.

Kim, Y. E., Schmidt, E. M., Migneco, R., Morton, B. G., Richardson, P., et al. (2010). Music emotion recognition: a state of the art review. *Int. J. Arts Technol.* 4, 74–89.

Knight, W. E. J., and Rickard, N. S. (2001). Relaxing music prevents stress-induced increases in subjective anxiety, systolic blood pressure, and heart rate in healthy males and females. *J. Music Ther.* 38 (4), 254–272. doi:10.1093/jmt/38.4.254.

Knox, D., Cassidy, G., Scott, B., and Macdonald, R. A. R. (2008). "Music emotion classification by audio signal analysis: analysis of self-selected music during game play," in Proceedings of the 10th international conference on music perception and cognition, Sapporo, Japan, August 25–29, 2008, 581–587.

Ladinig, O., and Schellenberg, E. G. (2012). Liking unfamiliar music: effects of felt emotion and individual differences. *Psychol. Aesthet. Creat. Arts.* 6 (2), 146–154. doi:10.1037/a0024671.

Lane, R. D., Quinlan, D. M., Schwartz, G. E., Walker, P. A., and Zeitlin, S. B. (1990). The levels of emotional awareness scale: a cognitive-developmental measure of emotion. *J. Pers. Assess.* 55 (1–2), 124–134. doi:10.1207/s15327752jpa5501&2_12.

Laurier, C., and Herrera, P. (2012). "Automatic detection of emotion in music: interaction with emotionally sensitive machines." in *Machine learning: concepts, methdologies, tools and applications* Jyväskylä, Finland, August 12-16, 2009 (IGI Global), 1330–1354.

Logan, B. (2000). Mel frequency cepstral coefficients for music modeling. *ISMIR,* 270:1–11.

ACKNOWLEDGMENTS

Mostafavi, A. C., Ras, Z. W., and Wieczorkowska, A. (2013). "Developing personalized classifiers for retrieving music by mood," in Proc. Int. workshop on new frontiers in mining complex patterns, Prague, Czech Republic, September 27, 2013 (Cham, Switzerland: Springer International Publishing).

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39 (6), 1161–1178. doi:10.1037/h0077714.

Shrift, D. C. (1954). *The galvanic skin response to two contrasting types of music*. Lawrence, KS: University of Kansas,

Szmedra, L., and Bacharach, D. (1998). Effect of music on perceived exertion, plasma lactate, norepinephrine and cardiovascular hemodynamics during treadmill running. *Int. J. Sports Med.* 19 (1), 32–37. doi:10.1055/s-2007-971876.

Vanderark, S. D., and Ely, D. (1993). Cortisol, biochemical, and galvanic skin responses to music stimuli of different preference values by college students in biology and music. *Percept. Mot. Skills.* 77 (1), 227–234. doi:10.2466/pms.1993.77.1.227.

Vink, A. C., Bruinsma, M. S., and Scholten, R. J. P. M. (2003). Music therapy for people with dementia. *Cochrane Database Syst. Rev.* 4, doi:10.1002/14651858.CD003477

Wauthier, F., Jordan, M., and Jojic, N. (2013). "Efficient ranking from pairwise comparisons," in International conference on machine learning (Thousand Oaks, CA: SAGE), 109–117.

Williams, D., Kirke, A., Miranda, E. R., Roesch, E., Daly, I., and Nasuto, S. (2014). Investigating affect in algorithmic composition systems. *Psychol. Music.* 43 (6), 831–854. doi:10.1177/0305735614543282.

Williams, D., Kirke, A., Eaton, J., Miranda, E., Daly, I., Hallowell, J., et al. (2015a). "Dynamic game soundtrack generation in response to a continuously varying emotional trajectory." in Audio engineering society conference: 56th international conference: Audio for games, February 11–15, 2015. New York, NY: Audio Engineering Society.

Williams, D., Kirke, A., Miranda, E., Daly, I., Hallowell, J., Weaver, J., et al. (2015b). Investigating perceived emotional correlates of rhythmic density in algorithmic music composition. *Trans. Appl. Percept.* 12 (3), 8. doi:10.1145/2749466.

Williams, D., Kirke, A., Miranda, E., Daly, I., Hwang, F., Weaver, J., et al. (2017). Affective calibration of musical feature sets in an emotionally intelligent music composition system. *Trans. Appl. Percept.* 14 (3), 1–13. doi:10.1145/3059005.

Williams, D., Hodge, V., Gega, L., Murphy, D., Cowling, P., and Anders, D. (2019a). "AI and automatic music generation for mindfulness." in Audio engineering society conference: 2019 AES international conference on immersive and interactive audio, York, UK, March 27-29, 2019 (New York, NY: AES. Available at: http://www.aes.org/e-lib/browse.cfm?elib=20439.

Williams, D., Wu, C-Y., Hodge, V., Murphy, D., and Cowling, P. (2019b). A psychometric evaluation of emotional responses to horror music. in Audio engineering society convention 146, Dublin, Ireland, March 21-23, 2019 (New York, NY: AES). Available at: http://www.aes.org/e-lib/browse.cfm?elib=20270.

Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. (2016). *Data mining: practical machine learning tools and techniques*. Burlington, MA: Morgan Kaufmann.