Linking Audience Physiology to Choreography

JIAWEN HAN, GEORGE CHERNYSHOV, MOE SUGAWA, DINGDING ZHENG, DANNY HYNDS, TAICHI FURUKAWA, and MARCELO PADOVANI MACIEIRA, Keio University
KAROLA MARKY, Leibniz University Hannover, University of Glasgow
KOUTA MINAMIZAWA, Keio University
JAMIE A. WARD, Goldsmiths, University of London
KAI KUNZE, Keio University

The use of wearable sensor technology opens up exciting avenues for both art and HCI research, providing new ways to explore the invisible link between audience and performer. To be effective, such work requires close collaboration between performers and researchers. In this article, we report on the co-design process and research insights from our work integrating physiological sensing and live performance. We explore the connection between the audience's physiological data and their experience during the performance, analyzing a multi-modal dataset collected from 98 audience members. We identify notable moments based on HRV and EDA, and show how the audience's physiological responses can be linked to the choreography. The longitudinal changes in HRV features suggest a strong connection to the choreographer's intended narrative arc, while EDA features appear to correspond with short-term audience responses to dramatic moments. We discuss the physiological phenomena and implications for designing feedback systems and interdisciplinary collaborations.

CCS Concepts: • **Applied computing → Performing arts;**

Additional Key Words and Phrases: Datasets, dance performance, electrodermal activity, heart activity

ACM Reference format:
Jiawen Han, George Chernyshov, Moe Sugawa, Dingding Zheng, Danny Hynds, Taichi Furukawa, Marcelo Padovani Macieira, Karola Marky, Kouta Minamizawa, Jamie A. Ward, and Kai Kunze. 2023. Linking Audience Physiology to Choreography. *ACM Trans. Comput.-Hum. Interact.* 30, 1, Article 9 (March 2023), 32 pages.
https://doi.org/10.1145/3557887

This work is conducted under the Cybernetic Being project supported by JST Moonshot R&D Program Grant Number JPMJMS2013.
JW is funded by a Leverhulme supported grant from The British Academy, Royal Academy of Engineering and Royal Society (APX101093), and a European Research Council (ERC) grant (Neurolive 864420).
Authors' addresses: J. Han, G. Chernyshov, M. Sugawa, D. Zheng, D. Hynds, T. Furukawa, M. P. Macieira, K. Minamizawa, and K. Kunze, Keio University, 4-1-1 Hiyoshi Kohokuku, 223-8526, Japan; emails: {hanjiawen, chernyshov}@kmd.keio.ac.jp, moe.sen.se321@gmail.com, {zheng208, dannyhynds, taichi, mapama, kouta, kai}@kmd.keio.ac.jp; K. Marky, Leibniz University Hannover, Hannover, 30060, Germany; email: marky@itsec.uni-hannover.de; J. A. Ward, Goldsmiths, University of London, New Cross, London, SE14 6NW, United Kingdom; email: jamie@jamieward.net.
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2023 Association for Computing Machinery.
1073-0516/2023/03-ART9 $15.00
https://doi.org/10.1145/3557887

1 INTRODUCTION

The most essential component in all performances, be it dance, music, or theater, is perhaps the live interplay between the performers and their audience. Even when following a script, or well-defined choreography, each rendition is different, partly due to audience reactions. Performers often describe it as an invisible link between them and the audience, involving the audience in co-creating art [12, 72].

Inspired by emerging concepts and technologies, performers have attempted to apply embodied interactions to enhance the notion of being connected. Within the HCI community, there is an established history of art-related research as well as a growing interest in collaborating with artists for interactive performances [8]. Many works investigate potential uses of sensing technologies to realize performers' creative visions, including tracking performer movements [46, 54], audience text messages [21], and performer and audience's physiological data [25, 38, 68, 92, 107]. Moreover, unobtrusive sensing technologies can help researchers and practitioners to understand the audience's experience in the wild.

A conventional way to investigate the subjective experience of a performance is measuring audience participation and engagement by using a combination of observations during the performance and the collection of audience responses afterwards [15]. This type of approaches is, however, are highly subjective and surveys collection can be unreliable depending on the audience's interest and difficult to deploy because some audience members might not want to answer and some may think they are marketing-related [54]. A complementary approach is to directly measure physiological responses of the audience using environmental or wearable sensors. The idea is to use unconsciously generated physiological data, such as changes in the **heart rate** (**HR**) or perspiration, as proxy measures for emotional state [55, 100, 107]. Yet, sensors might become a distraction from the performance and interpretation of the data is still open to discussion.

Our research reported in this article presents a path to explore live audience experience that grew out of a collaboration between researchers and performers. This article describes our process of sensing the audience's experience during a dance performance by capturing physiological data (**electrodermal activity** (**EDA**) and **heart rate variability** (**HRV**)) to enable the co-creation of the performance environment. A body of research have used these data and derived EDA and HRV features as measures of emotional affect across a variety of domains [3, 4, 14, 16, 22, 37, 96]. Most of the sensors we used for data collection are worn on the wrist, and were custom-built to reduce audience distraction.

While previous work have often explored audience experience by linking sensor data to concepts like engagement or attitude [55, 94, 101]. In contrast, our work explores the link between the audience's physiology and the choreography of the dance. Choreography encompasses compositional design and syntactical abstractions of movement to convey an underlying meaning or idea [2], and several works have used choreographic events to predict emotional arousal measured by continuous self-reporting [87]. By comparing physiological data to the choreographic structure, we investigate the complex interplay between planned live performance and audience experience. Additionally, we reflect on the overall interactive system design and how this might contribute toward embodied performance. We also summarize the lessons learned for both HCI researchers and performance artists.

Research Contribution: The main contributions of our work are the following:

(1) **A systematic overview of our iterative co-design process:** This article reports a three-year-long co-design process involving a modern dance ensemble, a choreographer and members of a contemporary theater that worked hand-in-hand with researchers to design a conceptual dance performance. Our process presents an interesting method to balance artistic

Fig. 1. A Scene of the dance performance, recording physiological signals from the audience.

and research interests during the collaboration. We report the process of integrating sensor feedback into the choreography, which contributed to creating both a novel artistic experience using audience physiological data to affect staging elements and an environmental method for researchers to analyze data collection in the wild, see Figure 1.

(2) **Linking data and art to explore insights into audience experience:** We explore links between audience physiology and choreography. We present our investigation into HR-derived features (specifically HRV features) and EDA features. We investigate the audience's experience from physiological, artistic, and subjective angles, linking HRV and EDA features during the performance. We report an exploratory interpretation of these observations and how they connect to choreography, audience subjective feedback, and interviews with the dancers. Our results reveal similar dynamic changes of physiology among different audience members, supporting the idea that choreographic design might be used to trigger expected physiological responses. Our findings as well as the general approach provide a promising template for future work on interactive performance.

(3) **Technical contribution for recording in the wild:** We describe a custom wearable sensor system that records motion, **blood volume pulse** (**BVP**), and EDA. The system can collect, store, and stream data from multiple participants simultaneously using wearable wristbands. This provides researchers a reproducible method to set up multi-modal physiology sensing systems in the wild.

(4) **An open source multi-modal dataset.** Finally, we contribute a freely-available, multi-person, multi-modal dataset (BVP, EDA, wrist acceleration and angular velocity) from 98 participants in total over three performances (duration about one hour per performance). The dataset (in total ca. 2.5 GB without video) together with the sample code for analysis is available for researchers under this link: http://bit.ly/audiencebiodata. This will give other researchers the chance to explore additional facets of the work beyond those covered in the current article.

The remainderof this article is structured as follows: Section 2 provides background information on how to utilize physiological sensing to explore live audiences and briefly introduces the sensing system that we used in our performances where audience physiological data were collected.

Section 3 describes the co-design process and iterative design considerations. Section 4 describes the prototype implementation and choreography for the final performance. Sections 5 reports the analysis and results of the audience's physiological data and their qualitative feedback. Sections 6 reports the methodology and results of analyzing the dance team's qualitative feedback. Sections 7 discusses the results and implications for interactive feedback design and co-design process. Sections 8 concludes the article.

2 RELATED WORK

Our research follows the broader research agenda of augmented humans, focusing on the overall experience of integrated feedback loops [23, 30, 36, 64, 105] rather than interactions between humans and computers. To set the scene for related work, we summarize research on feedback loops in *interactive performances*, *physiological sensing*, and *sensing live audiences*.

2.1 Interactive Performances

Early work in cybernetics inspired many art performances to incorporate physiological feedback in some way. Khut et al. identified the implications of physiological feedback for interactive art from the psychophysiology and presented a starting point for an exploration of participant-centred physiological feedback artworks [51]. Höök coined "Affective Loop Experiences" setting a research agenda for "bodily persuasion" [40]. She further introduced the so-called "Somaesthetic Appreciation" design concept, which means a correspondence by feedback and interactions that follow physiological rhythms. Our work takes these concepts as a basis, applying and extending them to a larger audience exploring bodily rhythms in the wild [41]. Collaboration between dancers and researchers is a common practice in interactive performances. During a co-creation process, researchers can benefit from artistic creativity to stretch the emerging technologies in unforeseen ways, a test-bed for embedding research content publicly [8]. The HCI community has investigated different approaches to designing for complex real-world experiences like those found in the performing arts [45, 73]. Dance in particular has been used to explore the use of technology to enhance aesthetic and affective expression. A common approach is to develop interactive elements that might be controlled or influenced by the performers [24, 31, 46, 54, 56], for example, incorporating changes in staging elements that are triggered by dancers' movements. An alternative approach is for interactions to be induced by the audience. For example, gauging audience reactions through text messages [21, 60], movements [63], or physiological data [38, 61, 68, 107].

Our study is based on an interactive dance performance where researchers and dancers collaborate to explore interactions through physiological sensing technology. This collaboration provides researchers and dancers an opportunity to accomplish novel dance performances and understand live audience experience in the wild. In the performances where we collected the dataset for this article, physiological feedback from the audience was used to trigger changes of stage elements. In a previous article, we described our design concept of streaming the physiological data (HR and EDA) in real time and integrating this data to staging elements, such as projected visualizations [88]. We did not investigate the collected data in our previous article because we considered the artistic aspects and concept design as the main contribution. In this article, we focus on the exploration of the dataset collected to get deeper insights into live audience responses in the wild and reflect on the long-term collaboration between HCI researchers and dancers.

2.2 Physiological Sensing

Related to the concepts of computational social science and social neuroscience, there is a lot of work expanding sensing beyond the individual and exploring the relationship between social experiences and physiological data [18, 57]. Some approaches used physical signals, such as location or movement data [29] and computer vision analysis, to detect facial landmarks, expressions, or postures [6]. The development of wearable sensing technology [34, 50] provided researchers possibilities to use internal signals associated with the involuntarily activated **autonomic nervous system** (**ANS**).

Emotions that humans experience while interacting with their environment are associated with varying degrees of physiological arousal where ANS plays a crucial role [5, 58]. ANS is mediated by two branches: (1) the **sympathetic nervous system** (**SNS**), and (2) the **parasympathetic nervous system** (**PSNS**) [59]. Parasympathetic nerves can exert their effect more rapidly (<1 s) compared to sympathetic nerves (<5 s) and mediate sudden large changes in HR [1, 27, 65]. Emotional states associated with ANS responses can be inferred using physiological data like **Electrocardiography** (**ECG**), EDA, and BVP [13, 78]. In our study, we used BVP and EDA, which are described below.

BVP is a pulse-based method of calculating the cardiac cycle from which the HR can be inferred [13]. HR and HRV are considered to result from ANS activities. The neurovisceral integration model describes HRV as the result of prefrontal cortex activities that affect modulation of the PSNS and SNS nervous systems' balance [93]. In other words, the neural circuitry that affects HRV goes very deep into our brain and reflects higher level cognitive processes and emotional states. Hence, HRV has been shown to be an indicator for reflecting emotions [16, 22].

EDA measures variations in skin conductance related to sweating and is a measure of the sympathetic nervous system. It has been used for over a century [47] and remains one of the most widespread tools for the measurement of autonomic nervous system responses in psychology and psychotherapy [69, 95]. EDA is often used to assess emotional arousal [3, 14, 86, 96]. A number of works explore the use of HR and EDA to gauge how groups of students respond to lectures and academic life [28, 33]. Sano et al. recorded mobile phone data, acceleration, and EDA (both wrist-based) of students during everyday university activities [77]. They were able to use this data to predict academic performance, sleep habits, and mental wellbeing. Recently, Gao et al. presented a viable system that detects students' in-class engagement using multidimensional readings of EDA, BVP, and other measurements like environmental sound [33]. Their choice of using both EDA and BVP supports our concept of using these two signals to track audience engagement. Our study differs from this prior work in that we do not seek to make explicit predictions, but rather present a process through which the physiology and behavior of a live audience might be interpreted and used.

2.3 Sensing Live Audience

Sensing live audiences has been explored using a variety of different sensor technologies including EDA [55, 84, 101], HR /HRV [7, 82, 98], **brain computer interfaces** (**BCI**) [38], and body movement [35, 90, 94] (see Table 1). Benford et al. [8] refer to the fusion of computing technology and public artistic projects (including performance) as "in the wild" research, in the sense of engaging "real" users with emerging technologies in real settings under conditions of actual use, as opposed to more constrained lab environments. We adopted this terminology and classify related work on audience sensing into either "lab" or "in the wild", with the latter referring to recordings during actual live theatrical performance. Table 1 provides on overview of this related work classification. Compared to physiological methods, physical signals, like body movements, facial expressions, and so on, are easier to record in the wild and thus feature prominently in the literature [35, 90, 94]. Theodorou et al. extracted face, hand, and body movement data collected from four contemporary dance performances together with two follow-up surveys on selected audience members for ranking performance and reporting engagement [94]. By comparing motion

Table 1. Recent Work About Sensing Live Audiences

Measurements	Scenario	Audience Size	Recording Duration	Sensing Technology	Method
GSR/EDA	Dance performance video [55]	49	11 minutes	Thought Technology GSR fingerwraps	Lab
	Films in theater/festival [84]	34	130 minutes	Affectiva Q Sensor	In the wild
	Live performance [101]	15	28 minutes	Customized sensor	In the wild
HR/HRV	Dance performance [7]	24	63 minutes	Bioharness 3 Sensor	Lab
	Piano performance (live/recorded) [82]	37	70/50 minutes	Win Human Recorder	Lab
	Dance performance [98]	101	35 minutes	Empatica E4	Lab
BCI	Live presentation [38]	11	35 minutes	Neurosky Mindwave	In the wild
Body Movement	Dance performance [94]	38	100 minutes	Night vision cameras	In the wild
	Music concert [90]	49	8 songs	Passive optical motion capture system	In the wild
	Dance/talks/music [35]	75	79/42/22 minutes	Customized neck-worn sensors	In the wild
EDA/HRV/ACC/ Angular Velocity	Our study	98	3 × 70 minutes	Customized wrist band	In the wild

We classified collection methods by referring to the “In the wild HCI research” definition by Benford et al. [8]. In our study, we collected multi-modal data via our customized wrist band in real dance performance settings, while recent works rarely sensed audience physiological data especially HR/HRV data in the wild.

data with survey results, they suggest that highest audience engagement corresponds to lowest overall movement. Yet they found no systematic relationship between audience movement and the dancers. Gedik et al. developed an approach to predict negative and positive experiences self-reported by the audience using accelerometer and proximity sensor data [35]. They linked body movements to memorable moments reported by audience members. In live music contexts, head movements were faster during live concerts than album-playback concerts. Swarbrick et al. explained this as higher engagement [90]. Clearly, different types of performance can elicit very different types of audience movement. In our work, we opted instead to use physiological instead of physical recordings where the response is arguably more straightforward to interpret.

Previous work using EDA to track audiences includes Silveira et al.'s exploration of using EDA to classify movie ratings [84]. For performances, Latulipe et al. used wearable EDA sensors to record 49 participants watching a video of a dance performance. Their results show strong correlations between the EDA and self-reported data, which supports the validation of temporal EDA data as reflection of audience engagement [55].

Similar to our work, Wang et al. recorded EDA from a live audience (15 participants for a 28-minute comedy performance) using wired electrodes on the palms [101]. From questionnaires' and EDA data, they clustered audience members and identified a main cluster of 10 audience members to represent the audience experience. They uncovered performance events (e.g., “balloon pops”) as changes in EDA and posited this as evidence of psycho-physiological engagement. Their study inspired us to find connections between audience physiology and the choreography of particularly noticeable scenes.

Audience HR/HRV has been mostly been studied in lab settings. Shoda et al. conducted a series of experiments to explore how HR and the spectral features of HRV differ between music that is live versus recorded, and fast tempo versus slow tempo [82]. They show that audiences tend to have higher HR and lower sympathovagal balance when listening to faster live performances. Interactions between pianists and the audience could reduce audience's physiological stress. In

Vicary et al.'s study, they tracked the dancers' acceleration as movement data and the audience' HR as affective feedback over five live performances [98]. Their results indicate a movement synchrony among performers that could predict audience aesthetic appreciation. Instead of looking into the synchrony among performers, Bachrach et al. used Myriam Gourfink's choreography to modulate respiratory rate and internal temporal clock and investigated the entrainment of audiences and dancers during dance performances [7]. They designed four experimental sessions from which they collected respiratory rate, and questionnaires related to subjective engagement, and time perception. Their work suggests that attention to breathing is closely related to entertainment. Their experimental design and analysis also inspired us to apply choreography as a method to trigger audience responses and interpret physiological reactions.

The work presented in this article applies multi-modal on-body sensing on audience members. This physiological data were utilized during the performance and afterwards, which contributes not only an interactive performance but also a valuable dataset for researchers to investigate audience reactions in the wild. Furthermore, our dataset was collected across three live performances with the same choreography. Although it is challenging to control all variables during a study in the wild, our interpretation of audience physiological data could be more robust and less subjective by comparing over three performances.

3 APPROACH AND DESIGN CONSIDERATIONS

In this section, we give an overview of the wider Boiling Mind project of which this work is part. We then detail the artistic and scientific *co-design process* involved in the work, before expanding on the *iterative design approach* that was used.

Boiling Mind is an embodied performance project combining modern dance practices, wearable sensing, and audio visual design. Both performers and researchers adopted physiological sensing as a way to explore the relationship between mind and body, invisible inner states and visible external cues [51]. As an initial use-case for the work, we created a trial 15-minute dance performance [32]. Building on this, we developed a full 70-minute dance performance that was performed three times [88].

We specifically investigated the link between performers and audience members. For this, we followed a methodology grounded in in-situ and in the wild studies [82, 90] to quantify and analyze live events. We focus on the audience in this work, introducing minimally intrusive sensing technology. We are particularly interested in the physiological changes of all audience members: are they entrained or following any rhythm? When does it happen?

3.1 Co-Design Process

To fulfill our goals, we held regular meetings, discussions, workshops, and test performances over the course of three years. The process involved 18 dancers, three stage directors, one choreographer, three stage designers, four visual designers, five engineers, seven researchers (from HCI, wearable computing, neuroscience, and performing arts), and one audio designer. In total, we conducted weekly meetings, 15 internal workshops, one 15-minute public trial performance, and the final series of 3 70-minute performances. Figure 2 depicts the overall iterative co-design process over the three years leading to the research insights presented in this article.

As an interdisciplinary project, there were several challenges for the research and dance teams. For the research team, instead of developing purely hypothesis-driven work, there was a need to pay equal attention to the needs of the performers and their vision. Giving control to performers might increase the risk of creating experiences for which the research questions cannot be formulated in advance [8]. For the performers, extra effort and time was required throughout the rehearsal process to understand and cooperate with the technology, as well as to design and adjust

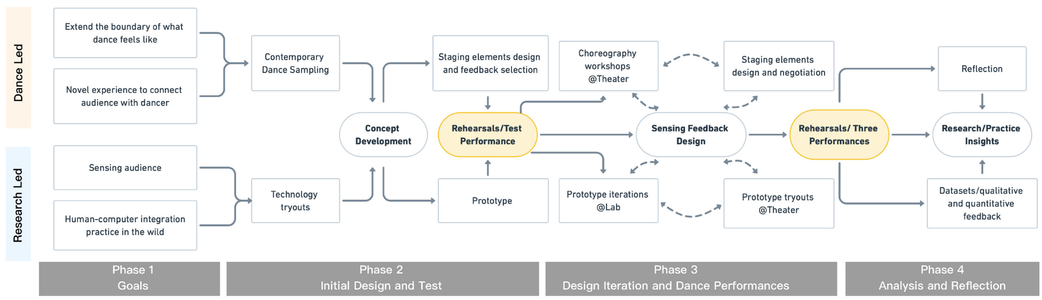

Fig. 2. This is a graphical overview of the ongoing iterative design process leading to the research insights described in this article. Five main components during the collaboration are presented in the middle and arrows show the dependencies and influences between the tasks.

their choreography accordingly. Meanwhile, the choreographer had to work hard to ensure that the research and technology components did not overpower the dance itself [54]. To offset these issues, we discussed the motivations of both the dancer and research team in advance. From this, we established a common goal: to probe the connection between the audience and dance during live performances. As this is fundamentally about performance, we also established that the design process and any final decisions about the production were made by the dance team.

The design process was centered around the five main components shown in Figure 2: (1) concept development, (2) initial test performance, (3) sensing feedback design, (4) the main performances, and (5) research insights. The *concept development* was to share domain knowledge and explore ideas about how to integrate sensing technologies with a dance performance. It started with practice reports from one of the co-authors who was also in the dance team, which included some sampling of contemporary dance performances to understand current practices in audience live responses. Technical members participated in this process, sharing the use of different technologies early-on to brainstorm about potential implementations. Workshops and rehearsals were organized during the prototype iterations for a suitable story arc to experience and record audience live physiological responses. Combined with the sensor tryouts and additional meetings related to the data visualizations, this led to the first prototypes and *test performance*. Insights from these activities led to the next stage of *sensing feedback design* in the form of an iterative cycle consisting of choreography workshops, staging elements design, prototype iterations, and prototype tryouts during rehearsals.

With the final design, we rehearsed and presented three "Boiling Mind Her Chair" main performances where audience physiological data were captured and used in real-time to provide visual and auditory feedback (described in Section 4). Combined with insights from the dance team, our key research insights are summarized in this article.

3.2 Iterative Design

In this section, we detail important design considerations that were made regarding choreography, sensing, feedback, audience involvement, and performance duration.

3.2.1 Choreography. The overall theme of the performance is to present women's struggle within Japanese society and to encourage others to find their own identities. The choreographer, an expert in dance performance design, helped craft both the interactive elements as well as physiological feedback effects of the work. Elements of the work were designed to break the "invisible wall" between dancers and the audience by including the audience's physiological reactions directly as a part of performance. Drawing on experience from our test performance [32], the

Fig. 3. Initial setup for 15-minute public test performance (as described in [32]). Audience physiological data were projected directly onto three screens on stage: HR, blink and pupil dilation (left), 12 audience members wore smart glasses (center), 4 members had their BVP readings measured (right). Additionally, 2 audience members wore pupil labs core eye trackers. The dancers wore accelerometer wrist bands.

choreographer planned specific scenes to trigger changes in audience physiological response as well as to highlight audience awareness of the effect their physiology can have on performance visualizations (see the final choreography in Section 4.1).

3.2.2 Sensing. We tried out different sensing technologies and conducted a test performance (see Figure 3) to evaluate the feasibility of large scale sensing in the theater. One technology we evaluated was eye-tracking using smart glass-based EEG [17, 43, 91]. We initially planned to have 60–100 participants wearing these; however, test participants reported feeling distracted by wearing something on their head and also did not like to look through the frame of the devices. Based on these reasons and the spread of COVID-19, we decided against a head-worn approach.

We also tried contact-free computer vision approaches to audience tracking. Kinect, LIDAR sensors, and OpenPose-based action recognition seemed promising [20, 48, 76, 108]. Unfortunately, lighting levels were too low to adequately capture the audience, and cost restrictions forbade the installation of additional (e.g., infrared) cameras. Even if this approach were feasible, some participants cited privacy concerns about the use of cameras. Instead, we turned to BVP and EDA sensing. Initially, we used commercial wrist-worn devices, like the E4 wrist bands from Empatica [34], but these showed poor data reliability and signal quality. This led us to develop our finger-based sensing hardware, described in Section 4.2). Audiences and test participants were more open to devices worn on their fingers and wrists than they were to head-worn approaches.

3.2.3 Feedback. One of the major challenges was implementing the feedback loop between dancers' movements and the audience physiological data in a noticeable but harmonious way. We selected the music and sound elements for our prime feedback loop in the main performances. The music for Boiling Mind consisted of our original compositions and existing pre-recorded tracks, including Maurice Ravel's Bolero [106]. The main feedback was implemented in our original composition sessions where the audience physiological data would affect three main aspects of the music: (1) rhythm, (2) timbre, and (3) texture [104]. Our initial design was to attach the audience's HR to the music's tempo. However, after we tested the feedback with physiological data streamed in real time, we found this direct link was not stable enough to dance with. Therefore, we finally decided to use the audience physiological data as a loose guide to trigger the changes of sound elements. Table 2 summarizes key feedback designs in the main performances and more detailed feedback design were included in our previous article [32].

3.2.4 Audience Involvement. A major difference between previous work and ours is the involvement of the entire audience.

Introducing novel interaction techniques to provide different means of live audience responses can have negative effects when they are not distributed to all or the majority of an audience. During a test performance, described in Fu et al. [32], we equipped between 10 and 20 participants from an audience with eye-tracking smart glasses, and used this data to incorporate some limited feedback on stage. Non-equipped audience members requested to be part of the performance and felt "left out." Although it is challenging to assess how strong this effect is in general, we want to highlight these issues as related work so far has not touched on the broad accessibility of novel interaction technologies in live performances.

These observations led us to equip all audience who wish to participate with sensors, and to favour the use of more unobtrusive, hand-worn sensors. As a result, 98 people (from a total audience count of 139) participated over the 3 performances.[1]

3.2.5 Performance Duration. There were several competing factors in setting the duration of the performances. Many individual performances by the dance groups we work with are around 10–20 min long. At first this seemed to be a suitable length, as it gives scope for creating an easily repeatable experimental setup. Yet, it was difficult to see longer term changes in the recordings of the physiological data. In particular, physiological changes in EDA and HR have been shown to undergo significant changes from 30–45 min [82, 84]. From a data analysis point of view, several hours of data would be ideal, but a contemporary dance performances lasting that long would be a strain for performers and the audience.

As a compromise, we decided on a duration of just over an hour (70 min). This provides sufficient data to observe interesting changes, while keeping in line with typical audience expectations (as surveyed during our experience with the Tokyo dance scene).

4 MAIN PERFORMANCES

Here, we describe the final setup and procedure used in the main performances. We first detail the choreography scene-by-scene. Next, we describe the design of our prototype implementation for a wearable sensing system, which records and analyzes the physiological data of the audience during the performance. We then outline the specific preparations made before each performance and report on the demographics of our participating audiences.

4.1 Choreography

Each performance involved seven female dancers and lasted for about one hour. For analysis, we divided the recordings into six sections, each containing one or more choreographic events. These sections are shown in Figure 4 and are described as follows:

Section 1: Suits. The performance starts with dancers playing the role of working women in suits and high heels trying to break out of societal pressures. At the end of this section, all dancers take off their suits and their heels. This intense movement was designed to raise the excitement level of the audience, mirroring the rhythmic, and dynamic crescendo of Ravel's Boléro.

Section 2: Cards. At approximately 11 min into the show, the dancers engage audience members in short conversations while handing out business cards. After the dancers return to the stage, they start hitting the floor in rhythm using their heels in hand. At 17 min, the dynamics and gestures of Boléro reach a final peak and one of the dancers rushes to the front of the stage to perform an aggressive solo (see Figure 4(2)).

Section 3: Puppet. At 18 minutes, the music turns to a more gentle and dark feel. At the same time, the previous solo dancer lays down in the center of the stage. One of the dancers brings a chair to the stage and the others gather around. All dancers start moving slowly and quietly

[1]The "Session House" in Tokyo and similar venues we are collaborating with can hold up to 100 visitors depending on stage design.

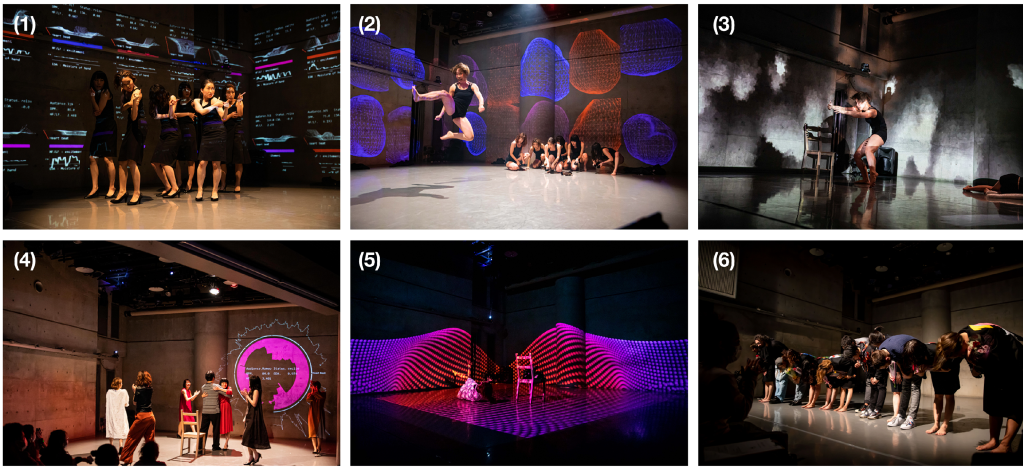

Fig. 4. Six sections: (1) Section 1: Suits, (2) Section 2: Cards (Solo), (3) Section 3: Puppet, (4) Section 4: Romeo, (5) Section 5: Growth, and (6) Section 6: Curtain.

along with the music. The chair represents everyone's position in the world. As the performance develops, each performer dances with a puppet that represents their alter ego. At the 24th minute, one dancer hands the puppet to an audience member. The choreography and music work together to create a mysterious and sombre tone.

Section 4: Romeo. At 37 minutes, one of the dancers invites a man from the audience to play the role of Romeo. He is led to the stage and sat on the chair, which was placed in the center of the stage. He is then asked to hug a dancer and dance together with the ensemble. The dancers start to improvise and reach out to Romeo to show they are happy to see him there. If he smiles, the dancers interact with him in an entertaining and playful way. If Romeo does not respond to the dancers accordingly, the dancers ask the rest of the audience to encourage him with applause. Some of the interactions between the Romeo and the dancers led to laughter among the audience.

Section 5: Growth. At 40 minutes, the second half of the performance develops into a deeper story. The dancers indicate the conflicting and complex feelings of instability, confusion, and joy that we all experience as we grow from childhood to adulthood. One of the dancers follows another one like a playful animal companion (e.g., a dog) willingly walking after its owner, clinging to her legs. This scene was designed to evoke the audience's sense of security and trust in being loved by others. The music for this part is quite sedated and relaxed, consisting of sparse synthesized textures and abstract rhythmic layers. As the final coda approaches, the dancers dress up as working women again but with different colorful designs embroidered into the back of their suits. This was intended to show a more positive meaning while referencing the beginning working women scene. The composed music reworked themes from Boléro into a more upbeat electronic treatment.

Section 6: Curtain. After 70 minutes, there is the curtain call, when all of the dancers and crew members line up in front of the stage and bow to the audience.

4.2 Prototype Implementation

We built wrist-worn devices measuring EDA from two electrodes on the fingers, and the HR using an optical BVP sensor placed on the fingertip (see Figure 5). The device uses an ESP32 module with Bluetooth and WiFi connectivity. It samples the BVP at 50 Hz and EDA at 4.545 Hz. The analog

Fig. 5. Left: Wristband with EDA, heart activity, acceleration, and gyro sensors. Right: Wristbands and consent from on the audience seats. The performance was held in an art studio with flexible uses—Session House. Before the performance, the staff from the Session House arranged seats for the audience. Cushion seats were used for the front rows to avoid obstructing the views of the audience who were sitting behind.

Table 2. Key Feedback Designs between the Audience Physiological Signal and the Staging Elements

Section	Visual Element	Sound Element	LF/HF ratio	EDA
Suits	one graph per audience member	–	value shows in the graph	value shows in the graph
Cards	one orb visual per audience member	–	control the orb's color	EDA difference controls speed of orb movement
Puppet	smoky fluid simulation	soundscape	average value controls the frequency and amount of smoke cloud's appearance	average value controls the amount of current smoke cloud and sub-frequencies into the soundscape
Romeo	one graph for Romeo	–	The value shows in the graph and controls graph's color	The value shows in the graph
Growth	one wave for all members	drum sounds	Average value controls the wave's color and dictates the pitch variance in the drum sounds	Average EDA difference controls the height of the wave and triggers the stuttering

More Details of the Implementation are Included in our Previous Article [88].

front-end of the EDA measuring circuitry consists of a Wheatstone measuring bridge connected to an AD8237 instrumentation amplifier feeding the data to the ESP32's internal 12-bit ADC. To save energy, each device buffers the data in 400 ms chunks and sends out the buffer 2.5 times per second. Both raw and filtered values were streamed and recorded through bluetooth. The analysis presented here uses only the raw values, while filtered values were used in the real-time visualisations. Transient response time of the digital filter output is guaranteed to be within 5% of the steady state in 10 seconds, which was necessary to provide smooth signal for the visuals generation without the noise from touching or adjusting the electrodes. More detailed technical implementation is included in our previous article [88].

In addition to the EDA and BVP, we recorded movement data using a 9-axis Bosch bmx160 absolute orientation sensor. The accelerometer and gyroscope were sampled at 50 Hz, magnetometer data were not recorded. For the feedback design, only EDA and BVP data were used as input to influence visual and sound elements on the stage (see Table 2).

4.3 Performance, Preparation, and Consent

Before each performance, all wristbands were disinfected with alcohol and placed on the seats together with performance flyers, consent forms and pens (see Figure 5).

Before the performance started, the host introduced the audience to the concept of our project and the design of the stage elements (see the stage visual design in Figure 4). The audience was briefed on the performance concept meaning the inclusion of physiological data into stage elements. All audience members were briefed about our experimental setup and were asked to read through the consent forms on their seats (see Figure 5). By signing the form, they agreed to participate and to record their physiological data. The experimental setup and data collection was conducted according to ethics rules and regulations of Keio University. We had 41–50 devices for each performance for data recording.

The connection of the physiological data from the wristbands and the visual projections, lights and music was demonstrated to the audience using a short section of pre-recorded data. The host explained how the physiological data could influence these changes, but we did not encourage them to consciously control it. Although there are some possible techniques to regulate heart rate or EDA, it is still challenging to control intentionally and such techniques or intentions might affect the audience's overall immersion and experience.

4.4 Audience Demographics

We recruited 98 participants (self-identified as male = 49; female = 49) from the audiences of three performances at the "Session House" during March 2020 (from total audience sizes of 57, 37, and 45). After removing incomplete or noisy data, we had physiological data from 80 participants (male = 38; female = 42). For the full details of these recordings, see Supplementary Material A.1.

Afterwards, the participants were asked to fill out an online questionnaire through a QR code on the performance flyers.

5 ANALYSIS OF AUDIENCE FEEDBACK

This section presents the pre-processing and results of analysing our audience feedback using both physiology (BVP and EDA) and post-show surveys.

5.1 Preprocessing of Physiological Data

5.1.1 Blood Volume Pulse (BVP). We extracted four commonly used HRV features from the recorded **Blood Volume Pulse** (**BVP**) data [81]. In a pre-processing step, we used acceleration data to help us identify and remove movement artifacts. To do this, we ran a peak detection algorithm on the euclidean norm of the accelerometer axes. If any peaks greater than 1.5 standard deviation were found, then we excluded the HRV data for 1s around each peak.

A 2nd order Butterworth low pass filter (from python package, *scipy.signal*) was then used to cut high frequency noise above 3.5 Hz [67, 99]. *Heartpy*, an python package for processing raw HR data, was used to get inter-beat (RR) intervals [97]. HRV features were calculated every 4 minutes with a 2-minute sliding window. These extracted features are:

— **LF/HF ratio:** the ratio of **low frequency** (**LF**) to **high frequency** (**HF**) power
— **SDNN:** the standard deviation of the inter beats intervals of normal sinus beats
— **RMSSD:** the root mean square of successive differences between normal heartbeats
— **PNN50:** the percentage of adjacent normal-to-normal intervals that differ from each other by more than 50 ms

Figure 6 shows the timeplots of each of these features' using the data from the third performance. The HRV features were divided by mean RR intervals of each participant for normalization to remove baseline differences between individuals [19, 74, 75]. For each minute, HRV features were averaged for each participant. The data were labelled in accordance with the six choreographed sections for analysis.

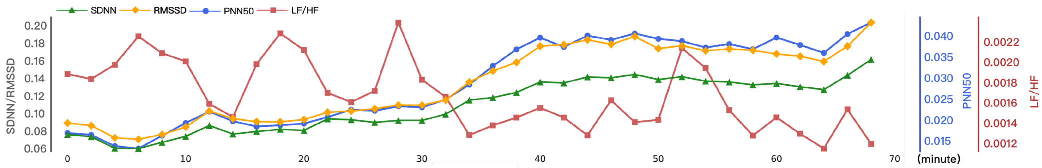

Fig. 6. Timecourse of the four HRV features from the third performance. Significant Pearson's correlations exist between PNN50 and SDNN ($r(35) = 0.99$, $p < .001$), and between PNN50 and RMSSD ($r(35) = 0.99$, $p < .001$) - with the later correlation also reported in other works [52, 80].

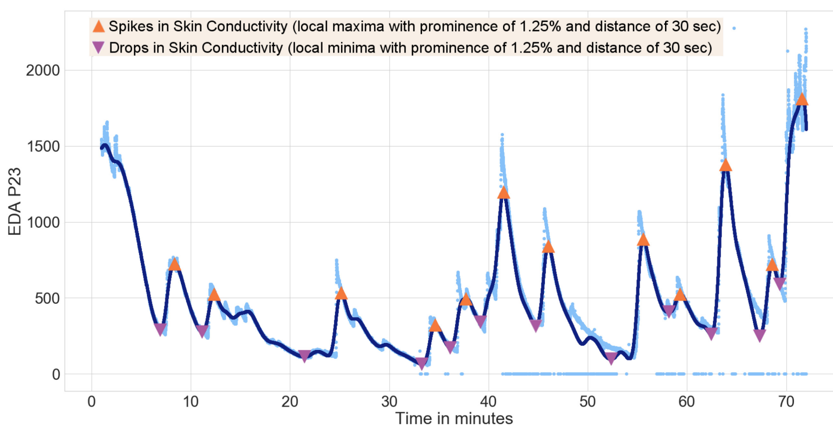

Fig. 7. EDA data of participant 23 (third performance). The raw EDA data from 12-bit ADC (0–4095 range) is depicted in light blue. The low-pass filtered data are depicted in blue; Orange markers depict recognized peaks in skin conductivity (local maxima); Purple markers depict recognized valleys (local minima).

5.1.2 Electrodermal Activity (EDA). Each participant's raw **Electrodermal Activity (EDA)** data were passed through a 2*nd* order Butterworth low-pass filter from the *scipy.signal* package (0.01 Hz) [99]. For EDA data analysis, we focused on the changes in EDA response, which is the first derivative of the EDA data. We refer to this as *EDA difference*. For each minute, the EDA differences were averaged for each participant. The data were labeled in accordance with six choreographed sections for analysis.

During the live performance, the EDA difference was used to affect the speed of change of the visualizations.

Because the onset of strong emotions is typically characterised by noticeably increased sweating on the skin, we looked specifically into the timings when skin conductance drastically increased. We refer to these points as *EDA extrema*. We detected peaks and valleys of skin conductance with prominence of 1.25% of the measurement range (0–4,095 due to 12-bit ADC) and inter-peak distance of at least 30 seconds. Since the performance venue is located underground with no cellphone coverage and during the performance, the only light was coming from the stage, we consider all spikes in skin conductivity to be very likely related to the subject's experience of the performance. Figure 7 shows an example of the process for locating EDA extrema. We counted the number of audience members who had experienced EDA extrema every minute to represent audience collective arousal feedback, which was addressed as EDA extrema count for describing our results and findings.

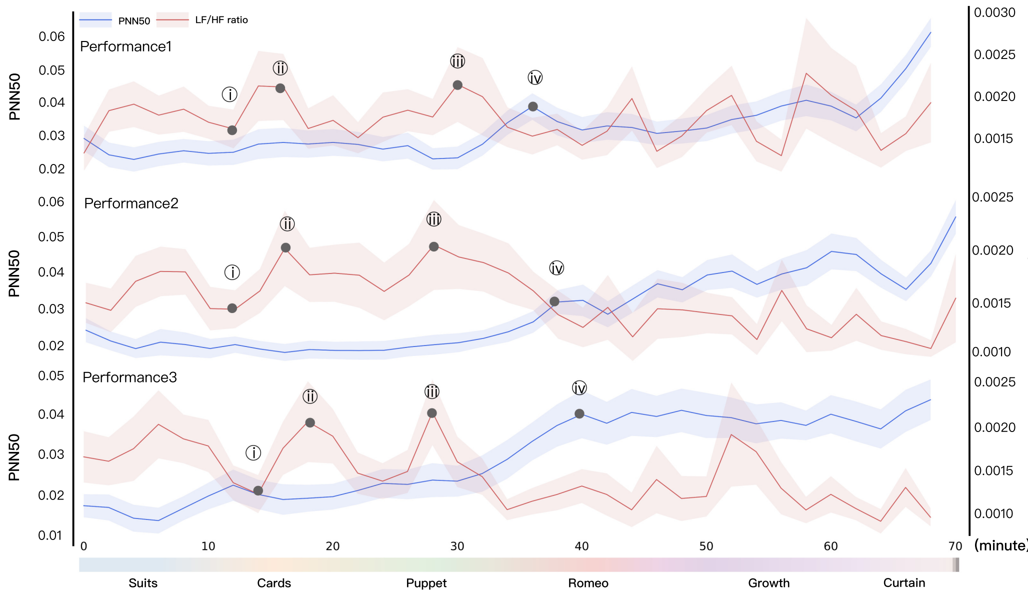

Fig. 8. Trends of PNN50 and LF/HF ratio with noticeable turning points. The timeseries shows noticeably similar patterns of HRV feature values across the three performances. (i) highlights the decline at the end of the section "Suits". (ii) marks peaks in the section "Cards". (iii) highlights the peak in the section "Puppet". Finally, (iv) marks the start of "Romeo".

5.2 Results: Physiological Data (HRV and EDA)

5.2.1 Heart Rate Variability (HRV). We initially extracted the four different HRV features as detailed above: LF/HF, PNN50, SDNN, and RMSSD. Whereas PNN50, SDNN, and RMSSD all reveal a similar timecourse, LF/HF ratio presents a very different pattern (see also Figure 6). Therefore, we choose only one of the three similar HRV features, PNN50, for further analysis for two reasons: Firstly PNN50 is easier to interpret because it represents **parasympathetic nervous system** (**PSNS**) only, which is associated with rest, and is consequently less influenced by **sympathetic nervous system** (**SNS**) (associated with excitement). In contrast, SDNN and RMSSD are driven by a mixture of PSNS and SNS, which can make interpretation more difficult [1, 27, 65]. Secondly, PNN50 is simple to calculate, which makes it useful as an indicator in designing a future real-time system.

We inspected the timeseries of the average LF/HF ratio and PNN50 over all audience members for each performance (see Figure 8). The timeseries shows noticeably similar patterns of HRV feature values across the three performances. For example, LF/HF ratios decline at the end of the section "Suits" and start rising at the start of the section "Cards". Then it first drops, and peaks again at around 30 minutes. The PNN50 is low at the start, but rises steadily throughout the performance. However, there is a sharper and drastic increase between the end of the section "Puppet" and the start of "Romeo".

5.2.2 HRV Scene Aggregate. As a further analysis, we aggregated the timeseries to produce statistics for each of the six main sections. A repeated measures ANOVA with a Greenhouse-Gessier correction was used to investigate the correlation and variance. For the post-hoc tests, we applied Bonferroni correction to prevent the inflation of type-I errors.

Fig. 9. EDA changes with EDA extrema counts (bar chart). EDA extrema peaks (highlighted in yellow) are selected as being over 1.5 standard deviation from the total EDA extrema counts, compared to EDA extrema within two or more minutes. The timeseries shows noticeably similar patterns of EDA feature values (EDA changes declined with outstanding EDA extrema peaks) at certain scenes across the three performances. The scenes are marked as ⓐ (the start of the section "Cards"), Around ⓑ (the end of "Cards"), ⓒ (the start of "Romeo"), and ⓓ (the end of "Romeo").

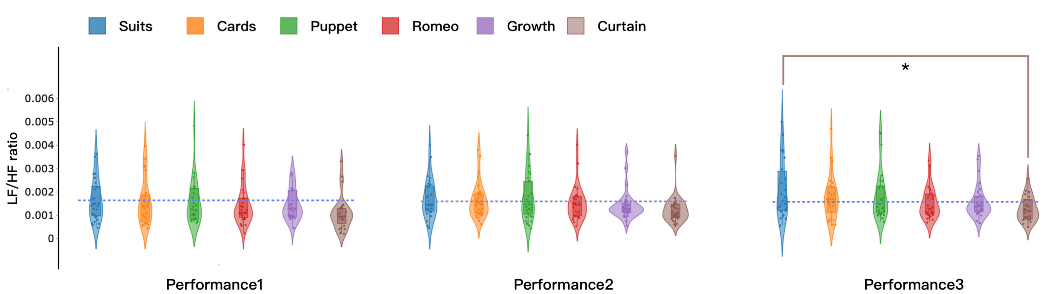

Fig. 10. Distribution of LF/HF ratio in six sections of three performances. The violin plots illustrate probability density, while individual observations are the dots within the violin graphs. The horizontal blue line represents the average LF/HF for each performance. The only significant pairwise difference is between Suits and Curtain in performance 3 ($^*p < .05$).

There were no statistically significant differences for LF/HF across the six sections over performance 1 (F(1.73, 44.93) = 0.351, p = .675) and performance 2 (F(3.15, 81.85) = 1.16, p = .33). Significant differences were present in performance 3 (F(3.37, 84.34) = 4.84, p = .003, η_p^2 = .162). In the post-hoc analysis we found that significant differences only existed between "Suits" (M = .0021, SD = .0012) and "Curtain" (M = .0012, SD = .0005) with p = .027 each. Since these two sections marked the beginning and end of only one performance, this effect is likely an anomaly. These results also depicted by Figure 10.

When analysing the mean PNN50 value, we found statistically significant differences between the six sections in performance 1 (F(2.81, 73.04) = 30.39, p < .001, η_p^2 = .539), performance

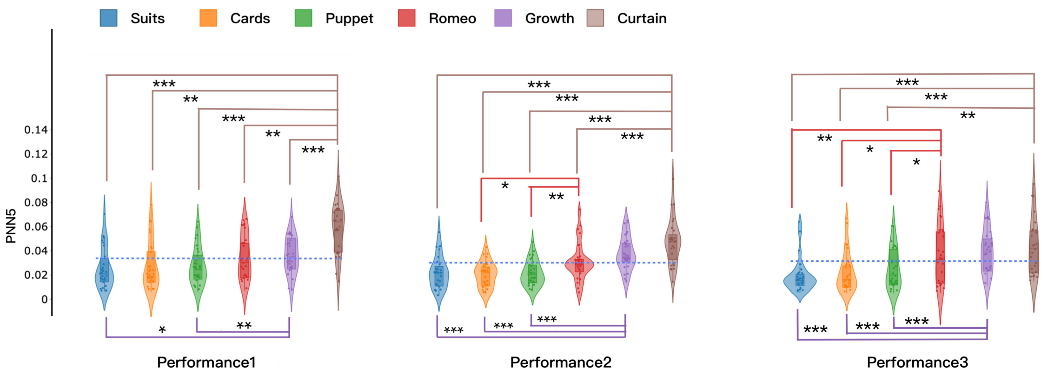

Fig. 11. Distribution of PNN50 in six sections of three performances. The violin plots illustrate probability density, while individual observations are the dots within the violin graphs. The horizontal blue line represents the average PNN50 for each performance. The vertical lines drawn between two graphs indicate the significant levels of pairwise comparison results ($^{*}p < .05$, $^{**}p < .005$, $^{***}p < .001$). The pairwise comparison shows distinguishable separation between the first half (before Romeo) and the second half (from the end of Romeo).

2 (F(3.10, 80.70) = 34.26, p < .001, η_p^2 = .569), and performance 3 (F(2.56, 64.08) = 18.40, p < .001, η_p^2 = .424). (Full descriptive statistics are provided in Supplementary Material A.3, Figure 11 depicts the distributions and pairwise comparisons.)

5.2.3 *Electrodermal Activity (EDA).* We inspect the EDA response using our two features: average EDA difference and EDA extrema counts (Figure 9). The timeseries reveals large changes at the beginning of each performance when the lights go off and the music starts, as well as at the end when the performers take a bow. Throughout the performances there are also common changes at around 13 minutes, marked in Figure 9 as a, 19 minutes b, 37 minutes c, and 41 minutes d. (Note that EDA extrema is shown in bar chart form to highlight that, unlike the other features, it represents a discrete count rather than an average.)

5.2.4 *EDA Scene Aggregate.* We calculated an aggregate pairwise comparison of EDA difference distributions between the 6 sections for each of the 3 performances (shown in Figure 12). After Bonferroni correction, we found that both Romeo and Curtain are statistically different to the other sections. According to a repeated measures of ANOVA with a Greenhouse-Gessier correction, mean EDA difference values differed in a statistically significant way between the sections in performance 1 (F(1.61, 41.93) = 13.41, p < .001, η_p^2 = .340), performance 2 (F(2.17, 49.97) = 8.15, p = .001, η_p^2 = .262), and performance 3 (F(2.16, 49.56) = 11.34, p < .001, η_p^2 = .330). (Full descriptive statistics are provided in Table 4.)

5.3 Audience Questionnaires

We used online questionnaires after each performance to gather audience feedback. These were accessible using a QR code on flyers handed out to each attendee. Responses were encouraged but not mandatory.

5.3.1 *Questionnaire Content.* The questionnaire assessed demographics, cultural background (how often do you visit theater/dance performances), and performance specifics (enjoyment of performance). Free-text answers were given to the specific question of "how much did you feel like participating in the performance", as well as general opinion on the piece. The full list of questions can be found in the Supplementary Material A.2.

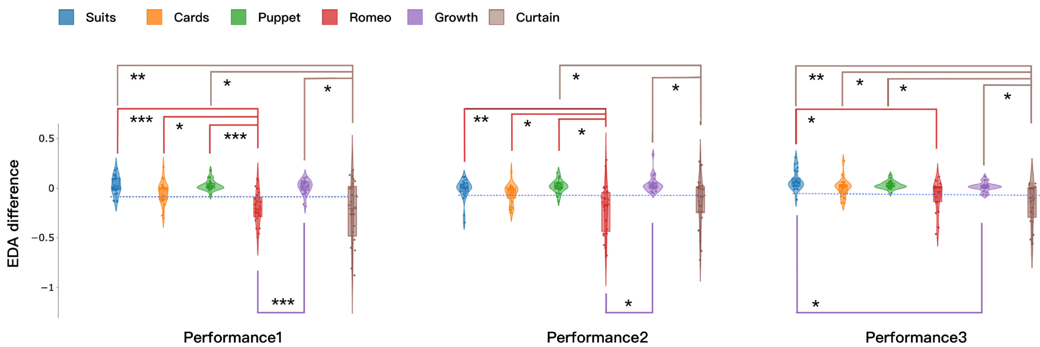

Fig. 12. Distribution of EDA difference in six sections of three performances. The violin plots illustrate probability density, while individual observations are the dots within the violin graphs. The horizontal blue line represents the average EDA difference for each performance. The vertical lines drawn between two graphs indicate the significant levels of pairwise comparison results ($^{*}p < .05$, $^{**}p < .005$, $^{***}p < .001$). Romeo and Curtain show significant differences to the other sections.

5.3.2 Questionnaire Participants. We received questionnaires from 35 participants in total (self-identified as male = 16; female = 18; prefer not to say = 1). Since the questionnaires were not completed by all participants, we consider these answers as supplementary information.

5.4 Results: Feedback from the Audience

Among the 35 respondents, 30 reported to experience of watching dance (every week: N = 3; every month: N = 8; every year: N = 19). For the open-ended questions, we categorized the feedback and reported as follows:

5.4.1 Participation in the Performance. A lot of feedback described strong feelings of participation compared to previous experiences. The simple knowledge that the audience was sensed might have played a part in this:

> *I was not sure if my heart rate really affected the visuals, but I was excited when thinking my heart rate was being measured. I felt like I was on stage at that time.*

> *The display of physiological data and the link between color and excitement were impressive and I felt I participated in it.*

Some participants reported a feeling of connection between audience and dancers:

> *Lighting and visuals changed in response to the audience's sensors, and I enjoyed the two-way interactions in this performance.*

However, others found the system confusing:

> *The music and visuals changed as our excitement changed. However, the lighting was a little difficult to notice.*

And some even felt a disconnect between how they felt and what they saw:

> *Sometimes the visuals from the sensor data matched my excitement while sometimes they did not match.*

5.4.2 Memorable Moments. Generally, participants considered the visualization, music, and dance as intriguing and meditative:

> *Whenever I watch their dance, something refreshing my memory happens. This time I had this feeling as well.*

Several audience members shared memorable moments:

> *I felt that music, sound, rhythm, and breaks tended to be the switches of excitement. During Bolero's gradation and explosion, rhythm and dance were connected closely.*

> *Dancers eye contacts when they using high heels to hit the floor were cool.*

> *I may be more excited in quiet and dark moments than when I'm feeling something intense will happen. I thought dancing the chair and the scene of the Japanese song were wonderful.*

5.4.3 Sense of Unity. When asked about their free opinions, most of participants mentioned they experienced a strong sense of unity between audience and dancer during the performance:

> *The abstract visual expression was very beautiful in connection with the dance. I felt that my senses were integrated with the dance through this indirect media.*

And among audience members:

> *I can feel not only my own sense of participation, but also other audience's reaction reflected. I was able to realize the sense of unity between the audience, which is usually hard to feel.*

However, one audience member doubted the need to enhance the sense of unity between dancers and audience suggesting that audience reactions may vary a lot due to different compositions of audience and this could make quality control harder:

> *In dance performances, "today's audience's feeling" and "sense of unity" seem to be less important to me. If the music and lighting change depending on the audience of the day, the impression of the work will change accordingly.*

6 DANCE TEAM FOCUS GROUP

To understand dancers' personal experience during the performance and attitudes toward the collaboration process, we conducted a focus group with five dancers which draws on the approach used by Huskey et al. [42].

6.1 Methodology

The semi-structured focus group was conducted via video conferencing three months after the performance when we had identified initial research insights. One of our co-authors, who is also in the dance team, was the facilitator. We encouraged the dancers to talk freely using the following pre-prepared questions as a guide:

(1) Share your experience and feeling when you saw the visualization, lighting, and the change of music triggered by the audience reactions. (Did you notice anything interesting, shocking, or disappointing?)
(2) Was this experience different from previous performances?
(3) What do you think about this collaboration? Share some experience of your memory about the collaboration.

All participants discussed in Japanese and videos were recorded for later transcription. We translated the transcripts and categorized the qualitative feedback.

6.2 Results: Dance Team Feedback

In general, dancers followed the choreography as they rehearsed without influence from the audience-derived visualizations. Two reasons were given for this. First, they had to focus on the performance itself with little time to care about changes in visualizations:

> *In the scene of Bolero, I need imagine the train's passing by during my dance and was not able to pay attention to visualizations until the scene changed,* (D1).

> *When I was waiting for my turn, audience heart rate displayed was very exciting. After I started dancing, though it was fun to see that, I could not spare my attention to the changes. I think we need more times to get used to it,* (D3).

The second is that they were sometimes confused by the meaning of the visualizations:

> *I felt it was tough to balance between something researchers want to show and something dancers want to show. It was difficult for me to fully understand the visuals' meanings and the formal performance day came,* (D1).
>
> *I did not understand the meaning when the lighting started to flicker. I did not see it as audience heart rates and it did not change that obviously,*. (The audience BPM data was mapped to the intensity of the lighting changes.) (D2)

Three dancers said that there were moments when they could sense audience reactions and one dancer were even influenced to adjust their movements:

> *I was aware of the visualizations when it came to the Romeo scene while the data used in generating the feedback loop was from the chosen audience,*. (D1)
>
> *It was very easy to see when audience felt more excited, but it was less noticeable for the calmer scene and I felt that audience's feelings did not change from the visualizations, (D2).*
>
> *When I danced with a puppet, I noticed the coloring of the visualizations were blue and I tried to dance intenser and faster, even hit the floor more painfully to get audience more sympathized and aroused,* (D4).

The dancers also thought there could be more space for improvisation where they could dance according to audience reactions, but it would be more difficult in terms of the choreography (D1, D2, and D4). To solve this, D2 mentioned they could "predetermine some triggers and reactions" accordingly during certain moments instead of improvising throughout the whole performance.

7 DISCUSSION AND IMPLICATIONS

In this section, we describe our main findings from our analysis of the physiological dataset and the qualitative feedback, followed by an interpretation of the physiological data against the artistic intent of the choreography.

7.1 Connecting Physiological Data to the Choreography

We first look at the overall timeline of the performance, highlighting notable moments from the choreography, and how these relate to changes in audience's physiological data. For ease of comparison, we summarise the four main physiological features—PNN50, LF/HF, EDA difference, and EDA extrema—from Figures 8 and 9 into one plot, Figure 13. We also highlight five noticeable moments in the data for further discussion. These moments are summarised in Table 3.

7.1.1 The Performance Timeline. According to the choreographer, the first half of the performance (through the section "Romeo") was designed to directly engage the audience and elicit

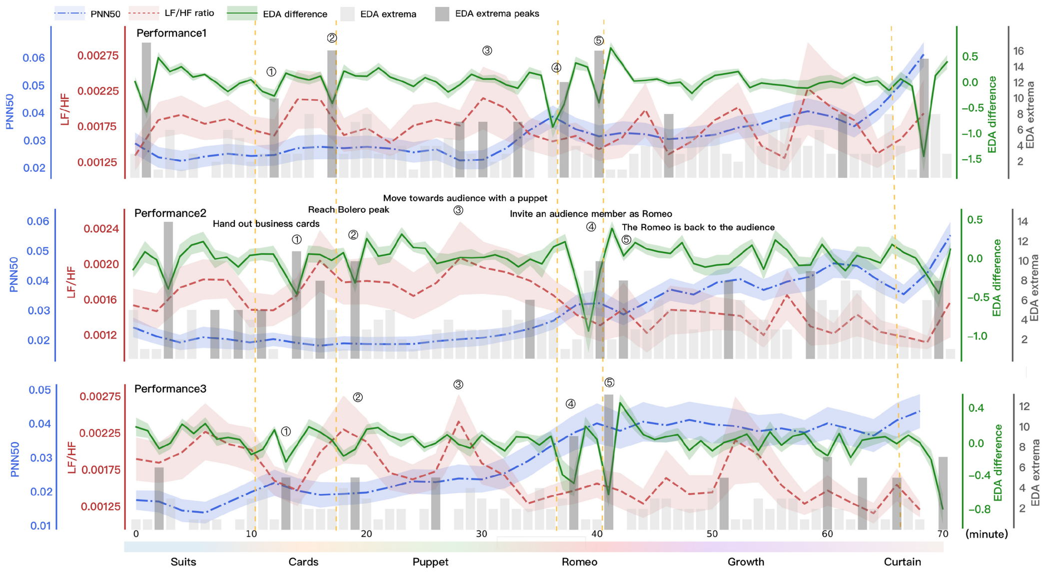

Fig. 13. The change of HRV features (Left Y scale) and EDA features (Right Y scale). EDA extrema peaks (highlighted in dark gray) are selected as being over 1.5 standard deviation from the total EDA extrema counts, compared to EDA extrema within two or more minutes. Five key moments are marked as ① (hand out business cards), ② (reach bolero peak), ③ (move toward audience with a puppet), ④ (invite audience member as Romeo), and ⑤ (the Romeo is back to the audience).

Table 3. Physiological Changes at Five Notable Moments Marked in the Figure 13

Moment	**Choreography elements**	**Changes in HRV**		**Changes in EDA**	
		LF/HF ratio	*PNN50*	*EDA Difference*	*EDA extrema*
①	Direct interaction, dancers to the audience, short time (5 s)	–	–	Abrupt drop	Outstanding peak
②	Music builds up, aggressive solo, strong rhythm by high heels	Noticeable spike	–	Abrupt drop	Outstanding peak
③	Dancers moving toward the audience, long time (20 s)	Noticeable spike	–	–	–
④	Direct interaction, one audience member to the stage	–	High Level	Abrupt drop	Outstanding peak
⑤	The audience member back	–	High Level	–	Outstanding peak

strong emotional reactions that might be more easily captured by the system. This includes direct interactive elements (e.g., “Cards”) as well as tense musical rhythms (e.g., “Puppet”). In the section “Romeo”, a direct and close interaction with one member of the audience was designed to elicit extreme involvement and empathy. The second half (from the start of the section “Growth”) was designed to be less interactive and more reflective. In contrast to the previous half, “Growth” did not involve any direct interaction between dancers and the audience in order to help the audience “digest” the piece and reflect on their own experiences. The choreographer’s intent here is to provide space to focus on the inherent message of the performance. In fact, several members of the audience reported that they experienced this as a process of “*meditation*”. Accordingly, the PNN50 shows a rising trend across all performances (see also Figure 13), with a significant increase between the second and first halves (see Figure 11). An increase in PNN50, being closely linked to

PSNS, is associated with relaxation [13, 70]. This trend over three performances could indicate the audience easing into the performance as it progressed.

7.1.2 Notable Moments from the Data. The changes of HRV and EDA features reflect the internal physiological reactions of the audience in response to certain choreographic elements. From Figure 13, it is hard to identify a clear long-duration trend of LF/HF over the three performances. The lack of any significant inter-scene difference in Figure 10 supports this. However, spikes of LF/HF ratios appear when the audience was subjected to relatively long scenes with an intensive crescendo (e.g., as the Bolero dance peaks ②, or when the Puppet is moved toward the audience ③). As an indicator of the balance between SNS and PSNS activity [81], LF/HF ratio could imply stress [78], anxiety [83], or excitement [9, 44, 66]. Even though the interpretation of the LF/HF ratio is controversial [11], it is still possible to explain the changes with careful consideration of the recording contexts [81]. Since the trend of LF/HF ratio does not synchronize with that of PNN50 (an established measurement of PSNS), we are inclined to believe that these changes of LF/HF ratio were mostly due to increased arousal under the influence of SNS activity. This is supported by previous work on the physiological responses to music, where significant LF/HF increases were observed during exciting, fast-tempo music [9, 66].

Across all performances, there is a steep increase in PNN50 starting from the end of Puppet and Romeo. This is also reflected, in part, in the pairwise scene aggregate results of Figure 11, where the change is significant for performances 2 and 3. Since increased PNN50 has been linked to engagement and sustained attention [39, 71], the change in PNN50 observed here seems to reflect the choreographic intention to create a sense of security and reflection during this section of the performance.

Direct and intense interactions between the audience and the dancers correspond to both abrupt drops in EDA difference and large peaks in EDA extrema counts (see Figure 13(①, ②, ④)). As an index of emotional activation, EDA can reflect arousal regardless of valence types [3, 14, 53, 79]. Previous audience studies have connected EDA to engagement [55] and shock-effects during the performance [101]. In our study, we interpret observed EDA changes as indications of audience's surprise ①, excitement ②, and tension ④ modulated by the choreography. EDA difference declines notably around the Romeo scene ④. Looking at the collective measurement for the audience as a whole, EDA extrema counts are also highest around the Romeo scene. This suggests that inviting someone on stage (e.g., to play Romeo) may trigger sudden tension among the remaining audience, whether in sympathy, or in anticipation that they might be invited next).

7.1.3 Mapping Audience Physiology and Choreography. Mapping different measures of audience physiology over the choreographic arc can help researchers and practitioners maintain a multi-modal, holistic view of live performance. Increased arousal related to SNS activation has been interpreted as high engagement or participation [55, 101]. However, increased arousal level should not be the only criteria for understanding the audience's reactions. Although the second half of our performance was expected to correspond to decreased arousal because of the meditative theme and atmosphere, the opposite was observed. We therefore propose considering both PSNS and SNS changes alongside choreographic structure (such as the overall story theme, interactive elements, and music crescendo) when investigating the physiological responses of the audience.

7.2 Implications for Physiological Feedback Design in Interactive Performances

Based on the above exploration of the dataset and qualitative feedback from the audience and dancers, we reflect on the physiological feedback design for interactive performances.

7.2.1 Choices of HRV and EDA Features. EDA difference is more sensitive to short, sudden elicitation compared to HRV-based features. Because it is directly linked to sympathetic activation and is generally quite discriminable [26], EDA difference is well-suited to gauging audience reactions in real time, particularly during short temporal moments like direct interactions or shock-effects [101]. EDA extrema, as an aggregated feature, might also be used with a threshold based on the majority of the audience's physiological reactions. However, the calculation relies on comparing to a global average and therefore a large chunk of data are required. This may make it unsuitable for real-time use when the data size is not adequate. However, as a post-recording analysis tool, it can help identify moments where the majority of the audience experience increased arousal. HRV features could be used to capture the audience's moment-by-moment experience as the theme of the piece develops. LF/HF ratio presents more frequent fluctuations compared to other HRV features. Although not as sensitive as EDA measurements, it can capture growing arousal during particularly emotional moments. PNN50, calculated as the difference between adjacent heart periods, is nominally independent of resting HR [10]. This makes PNN50 relatively representative of the audience's PSNS activity and associated reactions such as relaxation and sustained attention [13, 39, 70, 71]. We suggest using PNN50 as an indicator when there are obvious contextual affect changes, such as tension and relief, or conflict and reconciliation.

7.2.2 Interaction with the Audience's Physiological Feedback. The dancers reported focusing more on their movements than on the changing audience feedback. One reason could be the lack of the perceived agency within the interaction because the effects were triggered by the audience's physiology instead of the dancers' own responses as in previous works [31, 54, 56]. Another reason could be due to their lack of experience with the novel technology. Despite this, there were some scenes where the dancers responded strongly to the audience feedback, such as when they moved faster in order to elicit a change in the coloring. Drawing on this, it would be useful to explore a tighter integration between this technology and dance by incorporating the practice of audience feedback earlier in the rehearsal process. This would familiarise the dancers with the system and allow them to explore more nuanced and interesting responses to unexpected feedback.

7.3 Implications for Co-design Process

We summarize implications for interdisciplinary collaboration between HCI researchers and performers by reflecting on our three-year collaboration.

7.3.1 Balance around Goals. The goals of researchers and dancers can be very different even in the same project, so it is important for the project's success to uncover shared goals [54]. One shared goal in our case was to explore and enhance the invisible link between the dancers and the audience through performance. Although in this work we prioritized artistic values such as the consistency of the theme and the immersive experience of the audience, the choreographer worked closely with researchers to include performance sections that were explicitly designed to trigger clear emotional changes—changes that prior evidence suggested would trigger physiological responses.

7.3.2 Negotiate through Practices. Regular meet-ups at each stage of the co-design process are essential. The artistic director connected the dance team and research team and led the negotiations by conveying expected choreographic elements and showing sensor feedback samples. The two teams met regularly and organized workshops to make and revise design choices. A summary of the workshop and meetup schedule can be found in Supplementary Material A.4. Researchers attended major rehearsals, observed the stage conditions, and tested prototypes on the spot.

Considering the pivotal role of music in the work, the dancers were also given access to samples of audio feedback as it was developed.

7.3.3 Share Research Insights. It is important that any findings and insights uncovered by the research team are regularly shared with the performers. For example, following the performances, we shared a version of Figure 9 with the dance team. Revealing the mapping of physiology and choreography in this way helped provide a fertile ground for further discussion. During the discussion, dancers matched scenes to changes in the graphs and shared their feelings, experiences, and audience comments around those specific moments. Some of the dancers mentioned they had trouble understanding the data visualization during the performance. However, looking back on the data afterwards provided them more time and space to consider the effects. They even further reflected on how to improvise while referring to the feedback loop system and contributed valuable insights to the interpretation of physiological data. This process was crucial to our co-creation project and helped us plan the way for future collaboration.

7.4 Lessons Learned

The work presented here is primarily practice-led, where research methods, contexts, and outputs involve a significant focus on creative practice [8, 85, 89]. Based on our investigation of the audience's physiological data and the co-design process, we summarize the lessons learned for both HCI researchers and performance artists.

For HCI practitioners interested in performance and audience interaction, our approach explores an effective way to collect, analyze, and interpret audience experience during live performance. Live dance performance is a useful in-the-wild scenario to explore interaction paradigms that move away from the individual and toward interactions in larger-scale groups. Although academic research is usually conducted as goal-oriented, while artistic practice is more process-driven [85], both teams converged around the common goal to enhance the invisible link between dancers and audience through performance. This co-design process led to a series of novel performances and large-scale physiological data collections from the audience. Our exploration of the dataset reveals a link between the choreography and the audience's physiology. PNN50, being closely related to PSNS activation, shows a general rising trend and a significant increase from the second half sections. We found PNN50 could be a reliable and robust indicator of the audience's tension and relaxation during the performance. Moreover, LF/HF ratio, EDA difference, and EDA extrema could reflect the audience physiological reaction (e.g., excitement, surprise, and anxiety) elicited by choreographic elements such as strong rhythms and direct audience interaction. Our findings suggest the potential for a more holistic view on understanding and quantifying audience experience by cross-mapping choreography and physiology.

For performance artists, our research opens a viable method to incorporate audience physiological data within a live performance. Based on our post-analysis of the audience physiological data and the feedback from the dancers, we provide suggestions about choosing HRV and EDA features for live feedback. EDA difference is well-suited to gauge audience reactions in real time, since it is sensitive to short and sudden changes in arousal like shock effects. HRV features may be used to reflect the audience's moment-to-moment experience or long term growth of emotional arousal. PNN50 in particular is a robust measure for visualizing a sustained change in engagement, or a shift from tension to relaxation.

The choreographer and the artistic director suggested some focal points for future improvement. One focus is to investigate different forms of aesthetic interaction that might generate a clearer feedback loop between the audience's physiological reaction and the improvisations. Another focus is to improve the audiences' feeling of comfort during the performance—enhancing confidence and trust in the performance environment, both artistically and with the technology. Creating a suitable environment benefits the audience experience, their enjoyment of the performance, as well as enhances the potential to obtain better quality data for research.

7.5 Limitations and Future Work

First, our dataset was collected from a real performance where audience physiological data was used to trigger changes of staging elements. The existence of the feedback loop complicated the exploration and interpretation of the results. In the current stage, we could not find a feasible method to evaluate the effect of feedback system on audience physiological response. One possible method could be dividing the audience into control and experiment groups for baseline investigation. However, according to our previous feedback from the audience group, this can disappoint the audience and may not be acceptable in a real commercial performance. Another method could be organizing a separate lab study with the same or similar setup to quantify the effect of the feedback. Yet, in this case, the complexity of the audience response in the wild might be left out. We consider this is also a challenging but valuable topic for further investigations. Moreover, there were no negative questions in the questionnaire, which could result in incomplete feedback and we only analyzed open-ended questions this time. For our further studies, we will carefully design the questionnaire and also consider including an annotated post-viewing session to achieve a comprehensive understanding of memorable scenes quantitatively.

Additionally, we were not able to obtain valid data from people who were moving too much since both EDA and BVP are sensitive to movement artifacts. We used accelerometer data to help us inspect moving periods and remove noisy data. However, we still did not make full use of motion data collected from accelerometer and gyroscope sensor. In future work, we will look into audience dynamic synchrony from motion data (e.g., as in Ward et al. [102]). There are also alternative methods to analyze the dataset. Our initial analysis explored the data based on choreographic sections and elements. Further analysis could investigate the temporal component through autocorrelation or autoregressive models, quantify the group dynamics and synchronized behaviors [7, 62], and cluster the audience groups [101]. Future research could introduce alternative ways of obtaining subjective feedback from the audience to help interpret the physiological reactions such as questionnaires targeting memorable scenes and debriefing sessions on bodily experience [61, 72].

As a participatory performance, 41 people out of a total of 139 audience members did not sign the consent forms and so their data had to be destroyed. This would let us reflect on how to organize the participatory performance considering potential concerns about physiological sensing, such as privacy and safety of both the audience and dancers. We will further improve the guidelines for creating interactive performance referring to the instructions regarding improvisations [49] and participatory interactions [61].

In future, we will continue our work on investigating the feedback loop, translating remote audience responses, and exploring distributed liveness [103] to try to reconnect the essential invisible link between performers and their audience that has been lost in this era of Covid and online performance.

8 CONCLUSION

In this article, we report an in the wild study involving performers and public performances. This article describes our three-year co-design process of creating a novel art performance with physiological sensing technology as a trigger to affect staging elements, the study to understand performers and audience's experience in the wild, and corresponding reflections. For the practice, we contribute our iterative design considerations regarding choreography, sensing feedback considerations, and audience involvement for researchers' collaboration with performers to create novel artifacts and understand audience experience in the wild. For the study, we present a large-scale dataset to utilize and analyze physiological data from 98 audience members during the final performance. Compared to recent works on sensing live audiences from multiple sensing aspects (sensing modality, scenario, participants, duration, sensing technology, and collection methods), we argued the challenges and significance of physiological sensing in the wild. We further describe the implementation set up for data collection and the procedures for data analysis. The exploration of the dataset and collected qualitative feedback enabled us to discover the link between the audience physiology and the choreographic design. Through this reproducible approach, we are progressing toward understanding and enhancing the invisible connection between performers and audience members. We will continue to explore feasible methods and techniques to collaborate with performers and implement physiological sensing in the wild.

A SUPPLEMENTARY MATERIALS

A.1 Dataset Information

The dataset is consisted of audience multi-modal signals (EDA, BVP, wrist acceleration and angular velocity) over three performances. We have 98 recordings in total (male = 49; female = 49). In 1st performance, we have 34 recordings (male = 17; female = 17). In 2nd performance, we have 31 recordings (male = 13; female = 18). In 3rd performance, we have 33 recordings (male = 19; female = 14).

By ruling out incomplete or noisy data records, we had 80 (male = 38; female = 42) sets of data from the recruited participants for the HRV analysis of this project. The breakdown for each performance was: 1st, 27 (male = 12; female = 15), 2nd, 27 (male = 11; female = 16), and the 3rd, 26 (male = 15; female = 11).

Each data file contains:

- LocalTime: Local timestamp in milliseconds of the recording server at the time of data packet arrival to the server. Each packet of samples (approx 400 ms window) is labeled with the same local time.
- RemoteTime: Number of milliseconds passed since the recording device was turned on. Each sample is labeled with the exact time of measurement.
- Label: Labels for syncronization with the video recordings.
- Other columns are data fields with sensor readings

A.2 Audience Questionnaire

The full list of questions in the questionnaire delivered to the audience were as follows:

- How much did you enjoy this performance overall? (Likert scale: "1-not at all"–"9-very much")
- How much did you enjoy the visualization/music/lighting/dance? (Likert scale: "1-not at all" – "9-very much")
- Compared to other performances, how much did you feel participating in the performance? (Likert scale: "1-nothing" – "9-strongly") and Why did you have this feeling? (free-text answer)
- Which staging elements excited you most? Why did you have this feeling? (single choice: visualization, music, lighting, and dance)
- Please leave your opinions freely on this performance. (free-text answer)

A.3 Descriptive Statistics

Table 4 below presents mean and standard deviation values of features we used in the analysis.

Table 4. Descriptive Statistics of HRV and EDA Features Over Six Sections

		LF/HF	*PNN50*	*EDA difference*	*EDA extrema counts*
		Mean (SD)	*Mean (SD)*	*Mean (SD)*	*Mean (SD)*
Performance 1	Suits	.0017 (.0009)	.025 (.018)	.023 (.096)	4.18 (4.71)
	Cards	.0019 (.0014)	.026 (.021)	−.049 (.109)	5.75 (5.23)
	Puppet	.0018 (.0013)	.027 (.017)	.026 (.051)	4.29 (2.31)
	Romeo	.0015 (.0009)	.035 (.020)	−.231 (.203)	8.20 (5.59)
	Growth	.0017 (.0011)	.037 (.017)	.007 (.071)	4.54 (2.06)
	Curtain	.0019 (.0024)	.061 (.023)	−.344 (.462)	6.75 (6.24)
Performance 2	Suits	.0016 (.0007)	.022 (.014)	−.020 (.106)	4.00 (3.82)
	Cards	.0016 (.0008))	.020 (.011)	−.058 (.098)	4.13 (3.87)
	Puppet	.0018 (.0012)	.021 (.011)	.016 (.065)	3.82 (2.01)
	Romeo	.0015 (.0007)	.031 (.017)	−.200 (.234)	5.20 (4.09)
	Growth	.0014 (.0007)	.038 (.015)	.037 (.084)	4.85 (2.24)
	Curtain	.0014 (.0011)	.049 (.020)	−.193 (.300)	4.25 (3.30)
Performance 3	Suits	.0021 (.0012)	.015 (.013)	.060 (.100)	1.64 (1.75)
	Cards	.0019 (.0012)	.019 (.015)	−.012 (.090)	2.88 (1.36)
	Puppet	.00183 (.0011)	.022 (.017)	.016 (.039)	2.47 (1.46)
	Romeo	.0015 (.0007)	.035 (.026)	−.130 (.207)	3.20 (3.35)
	Growth	.0015 (.0007)	.038 (.019)	−.002 (.047)	2.96 (2.58)
	Curtain	.0012 (.0005)	.044 (.025)	−.246 (.306)	2.75 (3.10)

A.4 The Schedule of Main Workshops and Key Meetups

The following summarized the schedule of three workshops and key meetups when we were designing the feedback after the test performance:

— Meetup (2019. November): Shared feedback from the test performance and discussed the sensing feedback design schedule.
— Workshop (2019. December): The iteration started from a workshop where the choreography and the main piece—bolero were introduced to the researcher team. Meanwhile, the researcher team prepared the hardware tryouts to help the dancers generate intuitions about physiological sensing.
— Workshop (2021. January): Discussed the performance choreographic sections' plan.
— Meetup (2021. January): Recorded sound elements used in the feedback loop. Rehearsed and adjusted the composed music with sound feedback.
— Workshop (2021. January): Mixed the sound, music, visual, and choreography together.
— Meetup (2021. March): Showed full performance together with all technological set-ups to the Session house staff.

ACKNOWLEDGMENTS

We want to thank Renate Kunze for her loving support, encouragement, and kindness.

REFERENCES

[1] Solange Akselrod, David Gordon, Jeffrey B. Madwed, Nancy C. Snidman, Danied C. Shannon, and Richard J. Cohen. 1985. Hemodynamic regulation: Investigation by spectral analysis. *American Journal of Physiology-Heart and Circulatory Physiology* 249, 4 (1985), H867–H875.

[2] Sarah Fdili Alaoui, Kristin Carlson, and Thecla Schiphorst. 2014. Choreography as mediated through compositional tools for movement: Constructing a historical perspective. In *Proceedings of the 2014 International Workshop on Movement and Computing*. 1–6.

[3] John L. Andreassi. 2010. *Psychophysiology: Human Behavior and Physiological Response*. Psychology Press.

[4] Jenni Anttonen and Veikko Surakka. 2005. Emotions and heart rate while sitting on a chair. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 491–499.

[5] Bradley M. Appelhans and Linda J. Luecken. 2006. Heart rate variability as an index of regulated emotional responding. *Review of general psychology* 10, 3 (2006), 229–240.

[6] Sinem Aslan, Nese Alyuz, Cagri Tanriover, Sinem E. Mete, Eda Okur, Sidney K. D'Mello, and Asli Arslan Esme. 2019. Investigating the impact of a real-time, multimodal student engagement analytics technology in authentic classrooms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[7] Asaf Bachrach, Yann Fontbonne, Coline Joufflineau, and José Luis Ulloa. 2015. Audience entrainment during live contemporary dance performance: Physiological and cognitive measures. *Frontiers in Human Neuroscience* 9 (2015), 179. DOI: https://doi.org/10.3389/fnhum.2015.00179

[8] Steve Benford, Chris Greenhalgh, Andy Crabtree, Martin Flintham, Brendan Walker, Joe Marshall, Boriana Koleva, Stefan Rennick Egglestone, Gabriella Giannachi, Matt Adams, Nick Tandavanitj, and Ju Row Farr. 2013. Performance-led research in the wild. *ACM Transactions on Computer-Human Interaction (TOCHI)* 20, 3 (2013), 1–22.

[9] Luciano Bernardi, Cesare Porta, and Peter Sleight. 2006. Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: the importance of silence. *Heart* 92, 4 (2006), 445–452.

[10] Gary G. Berntson, Karen S. Quigley, Greg J. Norman, and David L. Lozano. 2017. Cardiovascular psychophysiology. *Handbook of Psychophysiology* (2017), 183–216.

[11] George E. Billman. 2013. The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. *Frontiers in Physiology* 4 (2013), 26.

[12] Miranda Boorsma. 2006. A strategic logic for arts marketing: Integrating customer value and artistic objectives. *International journal of cultural policy* 12, 1 (2006), 73–92.

[13] Patricia J. Bota, Chen Wang, Ana L. N. Fred, and Hugo Plácido Da Silva. 2019. A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals. *IEEE Access* 7 (2019), 140990–141020.

[14] Wolfram Boucsein. 2012. *Electrodermal Activity*. Springer Science & Business Media.

[15] Astrid Breel. 2015. Audience agency in participatory performance: A methodology for examining aesthetic experience. *Participations: Journal of Audience & Reception Studies* 12, 1 (2015), 368–387.

[16] Jos F. Brosschot, Eduard Van Dijk, and Julian F. Thayer. 2007. Daily worry is related to low heart rate variability during waking and the subsequent nocturnal sleep period. *International Journal of Psychophysiology* 63, 1 (2007), 39–47.

[17] Andreas Bulling, Jamie A. Ward, Hans Gellersen, and Gerhard Tröster. 2010. Eye movement analysis for activity recognition using electrooculography. *IEEE transactions on Pattern Analysis and Machine Intelligence* 33, 4 (2010), 741–753.

[18] John T. Cacioppo, Gary G. Berntson, and Jean Decety. 2010. Social neuroscience and its relationship to social psychology. *Social cognition* 28, 6 (2010), 675–685.

[19] A. John Camm, Marek Malik, J. Thomas Bigger, Günter Breithardt, Sergio Cerutti, Richard J. Cohen, Philippe Coumel, Ernest L. Fallen, Harold L. Kennedy, R. E. Kleiger, F. Lombardi, A. Malliani, A. J. Moss, J. N. Rottman, G. Schmidt, P. J. Schwartz, and D. Singer. 1996. Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Circulation* 93, 5 (1996) 1043–1065.

[20] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2019. OpenPose: Realtime multi-person 2D pose estimation using Part Affinity Fields. *IEEE transactions on pattern analysis and machine intelligence* 43, 1 (2019), 172–186.

[21] Teresa Cerratto-Pargman, Chiara Rossitto, and Louise Barkhuus. 2014. Understanding audience participation in an interactive theater performance. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. 608–617.

[22] Kwang-Ho Choi, Junbeom Kim, O. Sang Kwon, Min Ji Kim, Yeon Hee Ryu, and Ji-Eun Park. 2017. Is heart rate variability (HRV) an adequate tool for evaluating human emotions?–A focus on the use of the International Affective Picture System. *Psychiatry Research* 251 (2017), 192–196.

[23] Andy Clark. 2001. Natural-born cyborgs?. In *Proceedings of the International Conference on Cognitive Technology*. Springer, 17–24.

[24] Nuno N. Correia, Raul Masu, An Hoang Dieu Pham, and Jochen Feitsch. 2021. Connected layers: Evaluating visualizations of embodiment in contemporary dance performances. In *Proceedings of the 15th International Conference on Tangible, Embedded, and Embodied Interaction*. 1–12.

[25] Franco Curmi, Maria Angela Ferrario, Jen Southern, and Jon Whittle. 2013. HeartLink: Open broadcast of live biometric data to social networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1749–1758.

[26] Michael E. Dawson, Anne M. Schell, and Diane L. Filion. 2017. The electrodermal system. *Handbook of Psychophysiology* (2017), 217–243.

[27] J. P. A. Delaney and D. A. Brodie. 2000. Effects of short-term psychological stress on the time and frequency domains of heart-rate variability. *Perceptual and motor skills* 91, 2 (2000), 515–524.

[28] Elena Di Lascio, Shkurta Gashi, and Silvia Santini. 2018. Unobtrusive assessment of students' emotional engagement during lectures using electrodermal activity sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–21.

[29] Nathan Eagle and Alex Sandy Pentland. 2006. Reality mining: Sensing complex social systems. *Personal and ubiquitous computing* 10, 4 (2006), 255–268.

[30] Douglas C. Engelbart. 1962. Augmenting human intellect: A conceptual framework. *Summary Report, Stanford Research Institute*, Menlo Park, California (1962).

[31] Sarah Fdili Alaoui. 2019. Making an interactive dance piece: Tensions in integrating technology in art. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 1195–1208.

[32] Zhuoqi Fu, Jiawen Han, Dingding Zheng, Moe Sugawa, Taichi Furukawa, Chernyshov George, Hynds Danny, Tachi Furukawa, Marcelo Padovani, Kouta Minamizawa, Karola Marky, Jamie A. Ward, and Kai Kunze. 2021. Boiling mind-a dataset of physiological signals during an exploratory dance performance. In *Augmented Humans Conference* 2021 (2021), 301–303.

[33] Nan Gao, Wei Shao, Mohammad Saiedur Rahaman, and Flora D. Salim. 2020. n-Gage: Predicting in-class emotional, behavioural and cognitive engagement in the wild. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–26.

[34] Maurizio Garbarino, Matteo Lai, Dan Bender, Rosalind W. Picard, and Simone Tognetti. 2014. Empatica E3—A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. In *Proceedings of the 2014 4th International Conference on Wireless Mobile Communication and Healthcare-Transforming Healthcare Through Innovations in Mobile and Wireless Technologies*. IEEE, 39–42.

[35] E. Gedik, L. Cabrera-Quiros, C. Martella, G. Englebienne, and H. Hung. 2021. Towards analyzing and predicting the experience of live performances with wearable sensing. *IEEE Transactions on Affective Computing* 12, 1 (2021), 269–276. DOI: https://doi.org/10.1109/TAFFC.2018.2875987

[36] Donna Haraway. 2006. A cyborg manifesto: Science, technology, and socialist-feminism in the late 20th century. In *Proceedings of the International Handbook of Virtual Learning Environments*. Springer, 117–158.

[37] Mariam Hassib, Daniel Buschek, Paweł W. Wozniak, and Florian Alt. 2017. Heartchat: Heart rate augmented mobile chat to support empathy and awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2239–2251.

[38] Mariam Hassib, Stefan Schneegass, Philipp Eiglsperger, Niels Henze, Albrecht Schmidt, and Florian Alt. 2017. EngageMeter: A system for implicit audience engagement sensing using electroencephalography. In *Proceedings of the 2017 Chi Conference on Human Factors in Computing Systems*. 5114–5119.

[39] Jacob B. Holzman and David J. Bridgett. 2017. Heart rate variability indices as bio-markers of top-down self-regulatory mechanisms: A meta-analytic review. *Neuroscience & Biobehavioral Reviews* 74 (2017), 233–255.

[40] Kristina Höök. 2008. Affective loop experiences–what are they?. In *Proceedings of the International Conference on Persuasive Technology*. Springer, 1–12.

[41] Kristina Höök, Martin P. Jonsson, Anna Ståhl, and Johanna Mercurio. 2016. Somaesthetic appreciation design. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3131–3142.

[42] Sybil Huskey, Celine Latulipe, Melissa Word, and Danielle Lottridge. 2018. Postproduction focus groups in dance: A case study and protocol. *Journal of Dance Education* 18, 2 (2018), 47–54.

[43] Shoya Ishimaru, Kai Kunze, Yuji Uema, Koichi Kise, Masahiko Inami, and Katsuma Tanaka. 2014. Smarter eyewear: Using commercial EOG glasses for activity recognition. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. 239–242.

[44] Takuya Iwamoto and Soh Masuko. 2015. Lovable couch: Mitigating distrustful feelings for couples by visualizing excitation. In *Proceedings of the 6th Augmented Human International Conference*. 157–158.

[45] Rachel Jacobs, Steve Benford, and Ewa Luger. 2015. Behind the scenes at HCI's turn to the arts. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. 567–578.

[46] Jodi James, Todd Ingalls, Gang Qian, Loren Olsen, Daniel Whiteley, Siew Wong, and Thanassis Rikakis. 2006. Movement-based interactive dance performance. In *Proceedings of the 14th ACM International Conference on Multimedia*. 470–480.

[47] Carl Gustav Jung. 1918. *Studies in Word-Association*. Heinemann.

[48] Doris Jung, Marie Hermo Jensen, Simon Laing, and Jeremy Mayall. 2012. cyclic. an interactive performance combining dance, graphics, music and kinect-technology. In *Proceedings of the 13th International Conference of the NZ Chapter of the ACM's Special Interest Group on Human-Computer Interaction*. 36–43.

[49] Laewoo Kang, Steven J. Jackson, and Phoebe Sengers. 2018. Intermodulation: Improvisation and collaborative art practice for hci. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[50] Arvid Kappas, Dennis Küster, Christina Basedow, and Pasquale Dente. 2013. A validation study of the Affectiva Q-Sensor in different social laboratory situations. In *Proceedings of the 53rd Annual Meeting of the Society for Psychophysiological Research, Florence, Italy*, Vol. 39.

[51] George Khut. 2006. Development and evaluation of participant-centred biofeedback artworks. *Unpublished doctoral exegesis, University of Western Sydney* (2006).

[52] Tom Kuusela. 2013. Methodological aspects of heart rate variability analysis. *Heart Rate Variability (HRV) Signal Analysis: Clinical Applications* CRC Press, Boca Raton, FL, 10–42.

[53] Peter J. Lang, Mark K. Greenwald, Margaret M. Bradley, and Alfons O. Hamm. 1993. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30, 3 (1993), 261–273.

[54] Celine Latulipe, Erin A. Carroll, and Danielle Lottridge. 2011. Evaluating longitudinal projects combining technology with temporal arts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1835–1844.

[55] Celine Latulipe, Erin A. Carroll, and Danielle Lottridge. 2011. Love, hate, arousal,and engagement: Exploring audience responses to performing arts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1845–1854.

[56] Celine Latulipe, David Wilson, Berto Gonzalez, Adam Harris, Erin Carroll, Sybil Huskey, Melissa Word, Robert Beasley, and Nathan Nifong. 2011. SoundPainter. In *Proceedings of the 8th ACM Conference on Creativity and Cognition*. 439–440.

[57] David Lazer, Alex (Sandy) Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy, and Marshall Van Alstyne. 2009. Life in the network: The coming age of computational social science. *Science (New York, NY)* 323, 5915 (2009), 721.

[58] Robert W. Levenson. 2003. Blood, sweat, and fears: The autonomic architecture of emotion. *Annals of the New York Academy of Sciences* 1000, 1 (2003), 348–366.

[59] Robert W. Levenson. 2014. The autonomic nervous system and emotion. *Emotion Review* 6, 2 (2014), 100–112.

[60] Christopher Lindinger, Martina Mara, Klaus Obermaier, Roland Aigner, Roland Haring, and Veronika Pauser. 2013. The (St) age of participation: Audience involvement in interactive performances. *Digital Creativity* 24, 2 (2013), 119–129.

[61] Lian Loke and George Poonkhin Khut. 2014. Intimate aesthetics and facilitated interaction. In *Proceedings of the Interactive Experience in the Digital Age*. Springer, 91–108.

[62] Claudio Martella, Ekin Gedik, Laura Cabrera-Quiros, Gwenn Englebienne, and Hayley Hung. 2015. How was it? Exploiting smartphone sensing to measure implicit audience responses to live performances. In *Proceedings of the 23rd ACM International Conference on Multimedia*. 201–210.

[63] Helena M. Mentis and Carolina Johansson. 2013. Seeing movement qualities. In *Proceedings of the Sigchi Conference on Human Factors in Computing Systems*. 3375–3384.

[64] Florian Floyd Mueller, Pedro Lopes, Paul Strohmeier, Wendy Ju, Caitlyn Seim, Martin Weigel, Suranga Nanayakkara, Marianna Obrist, Zhuying Li, Joseph Delfa, Jun Nishida, Elizabeth M. Gerber, Dag Svanaes, Jonathan Grudin, Stefan Greuter, Harald Reiterer, Katrin Wolf, Jochen Meyer, Thecla Schiphorst, Dakuo Wang, and Pattie Maes. 2020. Next steps for human-computer integration. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.

[65] David Nunan, Gavin R. H. Sandercock, and David A. Brodie. 2010. A quantitative systematic review of normal values for short-term heart rate variability in healthy adults. *Pacing and Clinical Electrophysiology* 33, 11 (2010), 1407–1417.

[66] Yuuki Ooishi, Hideo Mukai, Ken Watanabe, Suguru Kawato, and Makio Kashino. 2017. Increase in salivary oxytocin and decrease in salivary cortisol after listening to relaxing slow-tempo and exciting fast-tempo music. *PloS One* 12, 12 (2017), e0189075.

[67] Mirja A. Peltola. 2012. Role of editing of RR intervals in the analysis of heart rate variability. *Frontiers in Physiology* 3 (2012), 148.
[68] Arttu Perttula, Pauliina Tuomi, Marko Suominen, Antti Koivisto, and Jari Multisilta. 2010. Users as sensors: Creating shared experiences in co-creational spaces by collective heart rate. In *Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments*. 41–48.
[69] TW Picton, I. Martin, and P. H. Venables. 1980. Techniques in psychophysiology. *John Wiley & Sons*.
[70] Stephen W. Porges. 1991. Vagal tone: An autonomic mediator of affect. *The Development of Emotion Regulation and Dysregulation* (1991), 111–128.
[71] Robert F. Potter and Paul Bolls. 2012. *Psychophysiological Measurement and Meaning: Cognitive and Emotional Processing of Media*. Routledge.
[72] Jennifer Radbourne, Katya Johanson, Hilary Glow, and Tabitha White. 2009. The audience experience: Measuring quality in the performing arts. *International Journal of Arts Management* 11, 3 (2009), 16–29 (14 pages).
[73] Asreen Rostami, Donald McMillan, Elena Márquez Segura, Chiara Rossito, and Louise Barkhuus. 2017. Bio-sensed and embodied participation in interactive performance. In *Proceedings of the 11th International Conference on Tangible, Embedded, and Embodied Interaction*. 197–208.
[74] Jerzy Sacha. 2013. Why should one normalize heart rate variability with respect to average heart rate. *Frontiers in Physiology* 4 (2013), 306.
[75] Jerzy Sacha and Władysław Pluta. 2008. Alterations of an average heart rate change heart rate variability due to mathematical reasons. *International Journal of Cardiology* 128, 3 (2008), 444–447.
[76] Makoto Sakamoto, Takahiro Shinoda, Takahiro Ishizu, Masahiro Hori, Hazuki Watanabe, Amane Takei, and Takao Ito. 2018. A proposal of interactive projection mapping using kinect. In *Proceedings of the 2018 International Conference on Information and Communication Technology Robotics*. IEEE, 1–4.
[77] Akane Sano, Andrew J. Phillips, Z. Yu Amy, Andrew W. McHill, Sara Taylor, Natasha Jaques, Charles A. Czeisler, Elizabeth B. Klerman, and Rosalind W. Picard. 2015. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. In *Proceedings of the 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks*. IEEE, 1–6.
[78] Philip Schmidt, Attila Reiss, Robert Duerichen, and Kristof Van Laerhoven. 2018. Wearable affect and stress recognition: A review. arXiv:1811.08854 (2018). Retrieved from https://1811.08854.
[79] Henrique Sequeira, Pascal Hot, Laetitia Silvert, and Sylvain Delplanque. 2009. Electrical autonomic correlates of emotion. *International Journal of Psychophysiology* 71, 1 (2009), 50–56.
[80] Fred Shaffer and J. P. Ginsberg. 2017. An overview of heart rate variability metrics and norms. *Frontiers in Public Health* 5 (2017), 258.
[81] Fred Shaffer, Rollin McCraty, and Christopher L. Zerr. 2014. A healthy heart is not a metronome: An integrative review of the heart's anatomy and heart rate variability. *Frontiers in Psychology* 5 (2014), 1040.
[82] Haruka Shoda, Mayumi Adachi, and Tomohiro Umeda. 2016. How live performance moves the human heart. *PloS One* 11, 4 (2016), e0154322.
[83] Lin Shu, Jinyan Xie, Mingyue Yang, Ziyi Li, Zhenqi Li, Dan Liao, Xiangmin Xu, and Xinyi Yang. 2018. A review of emotion recognition using physiological signals. *Sensors* 18, 7 (2018), 2074.
[84] Fernando Silveira, Brian Eriksson, Anmol Sheth, and Adam Sheppard. 2013. Predicting audience responses to movie content from electro-dermal activity signals. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 707–716.
[85] Hazel Smith. 2009. *Practice-Led Research, Research-Led Practice in the Creative Arts*. Edinburgh University Press.
[86] Erin T. Solovey, Marin Zec, Enrique Abdon Garcia Perez, Bryan Reimer, and Bruce Mehler. 2014. Classifying driver workload using physiological and driving performance data: Two field studies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 4057–4066.
[87] Catherine J. Stevens, Emery Schubert, Rua Haszard Morris, Matt Frear, Johnson Chen, Sue Healey, Colin Schoknecht, and Stephen Hansen. 2009. Cognition and the temporal arts: Investigating audience response to dance using PDAs that record continuous data during live performance. *International Journal of Human-Computer Studies* 67, 9 (2009), 800–813.
[88] Moe Sugawa, Taichi Furukawa, George Chernyshov, Danny Hynds, Jiawen Han, Marcelo Padovani, Dingding Zheng, Karola Marky, Kai Kunze, and Kouta Minamizawa. 2021. Boiling mind: Amplifying the audience-performer connection through sonification and visualization of heart and electrodermal activities. In *Proceedings of the 15th International Conference on Tangible, Embedded, and Embodied Interaction*. Association for Computing Machinery, New York, NY, Article 34, 10 pages. DOI: https://doi.org/10.1145/3430524.3440653
[89] Graeme Sullivan. 2010. *Art Practice as Research: Inquiry in Visual Arts*. Sage.
[90] Dana Swarbrick, Dan Bosnyak, Steven R. Livingstone, Jotthi Bansal, Susan Marsh-Rollo, Matthew H. Woolhouse, and Laurel J. Trainor. 2019. How live music moves us: Head movement differences in audiences to live versus recorded music. *Frontiers in Psychology* 9 (2019), 2682.

[91] Benjamin Tag, Andrew W. Vargo, Aman Gupta, George Chernyshov, Kai Kunze, and Tilman Dingler. 2019. Continuous alertness assessments: Using EOG glasses to unobtrusively monitor fatigue levels In-The-Wild. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.
[92] Paul Tennent, Stuart Reeves, Steve Benford, Brendan Walker, Joe Marshall, Patrick Brundell, Rupert Meese, and Paul Harter. 2012. The machine in the ghost: Augmenting broadcasting with biodata. In *Proceedings of the CHI'12 Extended Abstracts on Human Factors in Computing Systems*. 91–100.
[93] Julian F. Thayer and Esther Sternberg. 2006. Beyond heart rate variability: Vagal regulation of allostatic systems. *Annals of the New York Academy of Sciences* 1088, 1 (2006), 361–372.
[94] Lida Theodorou, Patrick G. T. Healey, and Fabrizio Smeraldi. 2019. Engaging with contemporary dance: What can body movements tell us about audience responses? *Frontiers in Psychology* 10 (2019), 71.
[95] Christian Tronstad, Gaute E. Gjein, Sverre Grimnes, Ørjan G. Martinsen, Anne-Lene Krogstad, and Erik Fosse. 2008. Electrical measurement of sweat activity. *Physiological Measurement* 29, 6 (2008), S407.
[96] Marieke van Dooren, J. J. G de Vries, and Joris H. Janssen. 2012. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology & Behavior* 106, 2 (2012), 298–304.
[97] Paul van Gent, Haneen Farah, Nicole van Nes, and Bart van Arem. 2019. Analysing noisy driver physiology real-time using off-the-shelf sensors: Heart rate analysis software from the taking the fast lane project. *Journal of Open Research Software* 7, 1 (2019), 32. DOI: http://doi.org/10.5334/jors.241
[98] Staci Vicary, Matthias Sperling, Jorina von Zimmermann, Daniel C. Richardson, and Guido Orgs. 2017. Joint action aesthetics. *PLOS ONE* 12, 7 (07 2017), 1–21. DOI: https://doi.org/10.1371/journal.pone.0180101
[99] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C. J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt & SciPy 1.0 Contributors. 2020. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods* 17, 3 (2020), 261–272.
[100] Chen Wang, Erik N. Geelhoed, Phil P. Stenton, and Pablo Cesar. 2014. Sensing a live audience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1909–1912.
[101] Chen Wang, Erik N. Geelhoed, Phil P. Stenton, and Pablo Cesar. 2014. Sensing a live audience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1909–1912.
[102] Jamie A. Ward, Daniel Richardson, Guido Orgs, Kelly Hunter, and Antonia Hamilton. 2018. Sensing interpersonal synchrony between actors and autistic children in theatre using wrist-worn accelerometers. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers* (Singapore, Singapore). Association for Computing Machinery, New York, NY, 148–155. DOI: https://doi.org/10.1145/3267242.3267263
[103] Andrew M. Webb, Chen Wang, Andruid Kerne, and Pablo Cesar. 2016. Distributed liveness: Understanding how new technologies transform performance experiences. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 432–437.
[104] Christopher Wm White, Jeffrey Fulmer, Brian Cordova, Alexandria Black, Chloe Danitz, William Evans, Aidan Fischer, Rashaad Greene, Jinhan He, Emily Kenyon, Joan Miller, Madeline Moylan, Abigail Ring, Emily Schwitzgebel, and Yatong Wang. 2021. A new corpus of texture, timbre, and change in 20th-century American popular music. (2021).
[105] Norbert Wiener. 2019. *Cybernetics or Control and Communication in the Animal and the Machine*. MIT press.
[106] Wikipedia contributors. 2021. Boléro. Retrieved September 2021 from https://en.wikipedia.org/wiki/Bolero.
[107] Shuo Yan, GangYi Ding, Hongsong Li, Ningxiao Sun, Yufeng Wu, Zheng Guan, Longfei Zhang, and Tianyu Huang. 2016. Enhancing audience engagement in performing arts through an adaptive virtual environment with a brain-computer interface. In *Proceedings of the 21st International Conference on Intelligent user Interfaces*. 306–316.
[108] Yingfang Zhang and Junren Zhou. 2020. Application of kinect-based multimedia interactive imaging in small theater. In *Proceedings of the ACM Turing Celebration Conference-China*. 243–246.

Received 19 April 2021; revised 28 April 2022; accepted 6 May 2022