

# Sensing a Live Audience

Chen Wang<sup>1</sup>, Erik Geelhoed<sup>2</sup>, Phil Stenton<sup>2</sup>, and Pablo Cesar<sup>1</sup>

<sup>1</sup>CWI: Centrum Wiskunde & Informatica  
Science Park 123, 1098 XG Amsterdam  
Netherlands

<sup>2</sup>Falmouth University  
25 Woodlane, Falmouth, Cornwall TR11 4RH  
United Kingdom

[cw@cwi.nl](mailto:cw@cwi.nl), [erik.geelhoed@falmouth.ac.uk](mailto:erik.geelhoed@falmouth.ac.uk), [phil.stenton@falmouth.ac.uk](mailto:phil.stenton@falmouth.ac.uk), [p.s.cesar@cwi.nl](mailto:p.s.cesar@cwi.nl)

## ABSTRACT

Psychophysiological measurement has the potential to play an important role in audience research. Currently, such research is still in its infancy and it usually involves collecting data in the *laboratory*, where during each experimental session *one* individual watches a video *recording* of a performance. We extend the experimental paradigm by *simultaneously* measuring Galvanic Skin Response (GSR) of a *group* of participants during a *live* performance. GSR data were synchronized with video footage of performers and audience. In conjunction with questionnaire data, this enabled us to identify a strongly correlated main group of participants, describe the nature of their theatre experience and map out a minute-by-minute unfolding of the performance in terms of psychophysiological engagement. The benefits of our approach are twofold. It provides a robust and accurate mechanism for assessing a performance. Moreover, our infrastructure can enable, in the future, real-time feedback from remote audiences for online performances.

## Author Keywords

Audience engagement; Galvanic Skin Response.

## ACM Classification Keywords

H.1. Human Factors, J.4. Psychology, J.5. Performance arts.

## INTRODUCTION

The primary motivation for the current study, and our research question, is to explore the viability of using Galvanic Skin Response (GSR) to monitor audience feedback during a live performance. We took GSR measurements of 15 people watching a live theatre performance simultaneously. The readings were synchronized with video recordings of the performance and the audience. The audience filled out questionnaires aimed to evaluate the emotions that the performance evoked. This resulted in a high volume of useful data of around 1680 data points for each participant.

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. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2014, April 26 – May 1, 2014, Toronto, Canada.

Copyright 2014 ACM XXX-X-XXXX-XXXX-X/XX/XX...\$5.00.

Results indicate that our approach – gathering GSR data during the play - is valid, as such data accurately reflects the engagement of the audience members. Moreover, it proves to be a useful tool for temporally unfolding the experience of the public, as the reactions of the public can be mapped to specific events during the play. In principle, we can conclude that our solution of using GSR data for monitoring audience feedback is novel and very valuable.

The paper is structured as follows. First, we highlight the novelty of our approach, when compared to prior works. Then, we describe the methodology we use in the study, followed by the results that support our hypothesis. Finally, the results are analyzed and discussed, focusing on the most important implications for next-generation video-mediated performances.

## RELATED WORK

Jennifer Radbourne details the importance of audience feedback through an extensive literature review and in-depth interviews [8]. The study justifies our hypothesis, as it shows that audiences are not primarily passive and that gauging the audience experience might provide an important measure of quality in the performing arts.

GSR measures excitation of the sympathetic nervous system and combined with other types of physiological and neurological, as well as self-report measures, have been applied in many areas of research, e.g. psychology, medical, gaming and education. Pejman et al. have explored how GSR data can be used for improving game design [6]; GSR sensors were also extensively used in research with hyperkinetic children [3]. For the purpose of this paper, we narrow the scope to audience feedback and how such measures open up new possibilities for interaction.

There are few studies using physiological measurements for learning about audience engagement per se (e.g. [1]). There seems to be more interest in *applying* sensor feedback *creatively*, e.g. to influence the outcome of a movie [2]. The most related work is the one from Celine Latulipe [5] who provides an extensive overview of research in this area. Her work draws on the empirical and theoretical work of Peter Lang [4], who describes a two dimensional space of different emotional states where one dimension runs from low to high (GSR) arousal and the other from low to high pleasure. Latulipe and colleagues explore how bio-feedback (in particular GSR) can be used to provide real-time visual

feedback to performers. In interviews with dance and theatre experts the notion of “valence” is introduced, i.e. how GSR arousal validates audience response.

Nevertheless, all these studies involve experimental sessions in which a single person watches a video recording of a performance. We instead took GSR measurement into the “field”, the natural habitat, and took collect the data of the audience simultaneously during the play. As such, we believe that our system supports ecological validity much better than the previous laboratory experiments.

## METHOD

Seven females (mean age 28.29) and eight males (mean age 23.13) formed the audience for a 28 minute theatre performance. Their GSR was measured every second throughout, resulting in 1680 data points for each participant. Actors devised and performed a comedy that was aimed at audience participation and produced occasional “shocks” (e.g. a popping balloon) to elicit the occurrence of GSR spikes during the performance (fig. 1).

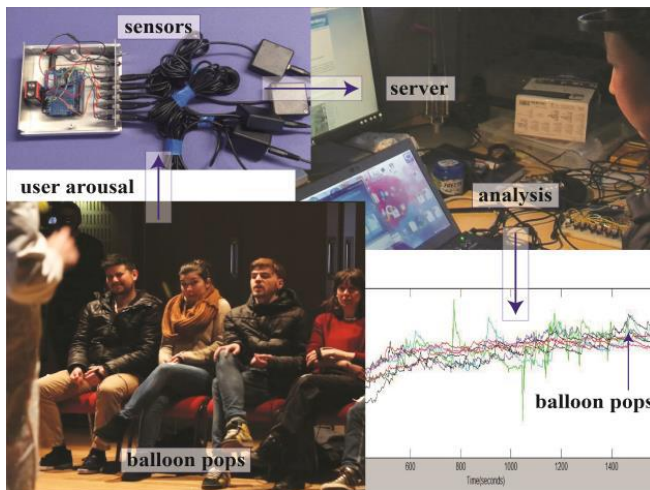


Figure 1: GSR system

For the results reported in this paper we used our own built sensors consisting of 1 Arduino UNO board and 1 Xbee wireless module (per five users), 1 47uF capacitor, 1 2Mohm resistor, and some aluminum foil. Two electrodes were placed in the users’ palm: in the case of the Arduino, one was connected to the 5V and the other to one of the analogues. There is a parallel circuit built with the resistor and the capacitor between the analogue and the GND. Arduino analogue readings range from 0 to 1023: when the two electrodes are open, the readings should be always 0; when the two electrodes are closed without the users connected, the readings should be always 1023. This mechanism allowed us to test the noise level (no noise after several hours of checking).

In addition to the noise proof testing, another validation method we used was to observe the distribution of the readings from the sensors. The sensor data distributions saw

typical linear patterns. Moreover, we applied the Filtering and Smoothing Matlab function in order to avoid the impact of hand movements during the experiment.

Groups of five sensors were each connected to one of three Arduino UNO boards (sample rate 1Hz). Xbee RF modules were used to create a wireless network such that the GSR data were sent directly to a laptop. This ensured the synchronization of all GSR readings. Cameras recorded the audience and the performance. Video streams were synchronized (post production) with GSR data.

Before the performance, participants filled out a short questionnaire asking about the type and intensity of the emotions they had experienced during the day. Afterwards participants filled out a similar questionnaire asking about emotions experienced during the play. The questionnaires were in the form of graphic rating scale [10] and measured 100mm. Participants were asked to make a mark between two extremes, i.e. between “not at all” and “very much”.

Participants were seated in one row with three sections of five seats each, arranged in a semi-circle around the stage. GSR modules were attached to the palm of the left hand. Before the performance started, participants took part in a meditation exercise to establish a baseline GSR level.

Questionnaires were analyzed using Analysis of Variance (ANOVA) and correlations. The synchronized GSR and video streams enabled us to relate events during the performance to corresponding GSR readings. GSR readings were analyzed using the Multidimensional Scaling (MDS) method [7]. Correlations and ANOVA had some limitations to do a complete interpretation of the readings. They are fairly suitable if the audience is being treated as a whole, but they cannot properly explain relationships – similarity and dissimilarity - between objects in a multi-dimensional space. In our case, we were interested in understanding the relationships between 15 objects (each audience member) GSR responses’. We calculated the dissimilarities between the objects using Pearson Correlation Coefficients, and two-dimensional scaling was chosen for scaling. After 30 iterations, the final configuration graphs were achieved and Kruskal’s stress reported in the results. For the subject to subject (Stress=.06; RSQ=.99) and minute to minute (Stress=.05; RSQ=.99) MDS solution we used the (SPSS) ALSCAL algorithm with good validation values.

## RESULTS

### Audience clustering

A MDS solution (fig. 2), based on correlations of GSR readings between audience members, shows how ten participants correlated closely (on average  $r = .86$ ), showing an initial rise in GSR followed by a flattening towards the end of the performance (inset in fig. 2). In this plot the Kruskal’s stress value is 0.06 (less than 0.15), indicating that the configuration of the 15 participants’ GSR readings can be considered as reliable. Questionnaire results and

brief interviews after the performance indicated that this group had been deeply engaged in the performance, reporting high levels of enjoyment.

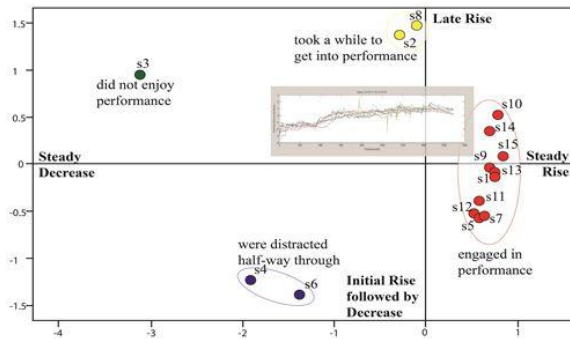


Figure 2: MDS Audience clustering based on GSR data

Five participants displayed different patterns. Two showed an initial rise in GSR followed by a decrease, i.e. after an initial engagement with the performance their attention waned; for one this was related to receiving sad news during the day. Two showed an initial lack of rise in GSR followed by an increase; they reported to be confused initially by the purpose of the play and as such it took them a while to get into the performance. One participant displayed a consistent drop in GSR and reported not liking the performance. These characteristics enabled us to label the extremes of the X and Y-axes.

### Unfolding of the performance

For each minute, the GSR readings were averaged for each participant. Here MDS (Kruskal’s stress: 0.05) yielded an almost chronological minute by minute unfolding of the play (anti-clockwise in fig. 3) up to minute 19. Using the video footage we were able to identify the clusters based on the content of the performance. Thus, initially the GSR readings are low (minute 1) followed by a steady rise (minute 2 – 19) after which the intensity of the GSR flattens (minute 20 – 28). The first part of the performance (minute 2 – 16, in red in fig. 3) built up to an active and physical participation during which the participants were asked to raise either their left or their right leg in response to (silly) questions by the actors. In minute 17 – 19 (in green) the results of a competition were revealed, where the relatively higher GSR readings might indicate anticipation. After that the audience was not required to interact as actively as they listened to a trumpet player (dark blue) and watched a juggling act (yellow). The Y-axis reflected levels of GSR intensity and the X-axis ran between low and high audience participation. Spikes were identified that corresponded to the intended “shocks”, e.g. balloon popping, the sudden sound of a (badly played) trumpet.

The minute average GSR readings during this comic play correlated positively with participants being (very) cheerful (on average  $r = .62$ ) and correlated negatively with participants being sad (on average  $r = -.60$ ) at different

stages of the performance, in particular from minute 16 onwards the average GSR readings showed strong correlations with audience’s “cheerful” ratings.

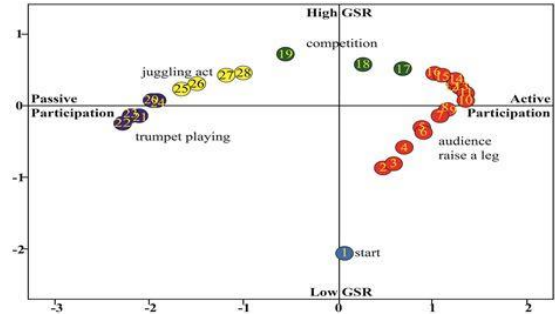


Figure 3: MDS minute by minute unfolding of the performance

Table 1 summarizes the significant differences between pre- (asking about their experiences during the day) and post-performance questionnaires. The ratings were given on a scale between “not at all” (= 0) and “very much” (=100). Thus participants rated that during the day, on occasion, they had a laugh with a mean intensity of 45 (Mean pre in Table 1) and they reported that the intensity of laughter generated by the performance was rated on average as 68.5 (Mean post), resulting in a significant difference,  $F(1,14) = 14.68, p = .002$ . Similarly, for cheerfulness, the difference between pre-and post- ratings was significant,  $F(1,14) = 7.12, p = .018$ . On average, participants had a reasonably cheerful day (Mean = 55) but these ratings increased to on average 74.5 after the performance. Lastly, although participants did not have a particularly sad day (with the exception of one participant) yielding a mean of 35, this was significantly reduced to a mean of 11.4 after the performance  $F(1,14) = 5.82, p = .03$ . There were also significant effects for gender and whether a participant knew another participant sitting in their row or not. However, due to the low numbers in each “cell”, we refrain from reporting these in this short paper.

### Pre- and Post- Performance Questionnaires

Item	p	Mean pre	Mean post
Laugh	.002	45	68.5
Cheerful	.018	55	74.5
Sad	.03	35	11.4

Table 1: Significant differences between pre-and post-questionnaires

### DISCUSSION

This paper describes an audience experience during a live theatre performance using a system to measure GSR of 15 people simultaneously. Psychophysiological measurement in audience research is usually carried out in a laboratory, where during an experimental session one individual watches a recording of a performance. There are advantages

to such studies as a range of physiological and neurological sensors can be used concurrently. However, being part of an audience is a group experience and it might not be straightforward to extrapolate from an individual's experience watching a recording to a larger audience watching a live performance. As such, we believe that our system supports ecological validity much better than laboratory experiments. It is not just an innovative contribution to audience research methods that makes this study of interest. We found that for most participants there was an unexpected and unusually high level of physiological closeness (GSR), which, to our knowledge, has not been reported before.

Analyzing the GSR data in conjunction with synchronized video recordings provided additional insights, e.g. we were able to link spikes in GSR to shock-effects during the performance. This validates the robustness and accuracy of our measurements. However, more interesting are the general (smoothed) shapes of audience engagement. We found that GSR readings of most of the audience followed a curve where in the initial stages readings were low and as the play progressed this increased steadily, reflecting an increase in engagement with the play across time. Returning to the notion of valence [5], low GSR does not necessarily imply a negative audience judgment, for peaks to happen troughs are essential, but it is informative to evaluate the overall shape of the response, as exemplified by the steady decrease in GSR of one participant who was not engaged by the play. In addition we could place the performance in a two dimensional space, not unlike Lang's [4] where one dimension runs from low to high arousal and the other from low to high audience participation. Embedding the audience experience in how their day had been also showed how a performance can lift an audience out of the ordinary. Cheerfulness and enjoyment were the main (more) emotional components linked to the GSR data.

### CONCLUSION AND FUTURE WORK

We are currently in the process of scaling up the system, taking advantage of small form-factor developments in wireless technology and GSR measurement. Heart rate and blood pressure might make useful additions.

There are some near future applications of a relatively low-cost system such as the one described here. In a "next-bench" type of fashion it can be used for further audience research. It makes it feasible for theatre companies to receive detailed (and time-stamped) early feedback during try-outs, to evaluate what works well and what does not or rather identify where audience engagement wanes.

In our work we are exploring how current video-mediated technologies can be used, and extended, for supporting novel interactive performances. In particular, our final goal is to provide the infrastructural components and support, so performing artists can reach a wider remote audience with their productions, but still maintaining the close relationship

between the actor and the audience. Live performances of big productions are already streamed to cinemas and to homes. However, these present limited opportunities for audience interaction. Wearable physiological sensors have the potential to open a whole array of creative solutions to suite this aim, e.g. the aggregate response of those who are deeply engaged in a performance can be used to provide visual, auditory or even haptic feedback [9] to performers.

### ACKNOWLEDGEMENTS

This work was supported in part by funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. ICT-2011-287760.

### REFERENCES

- [1] S. Bardzell, J. Bardzell, and T. Pace. 2009. Understanding affective interaction: Emotion, engagement, and internet videos *Proc. of IEEE International Conference on Affective Computing and Intelligent Interaction*, pp. 1-8.
- [2] C. Grant. 2013. Many Worlds: The movie that watches its audience. *BBC News*  
<http://www.bbc.co.uk/news/technology-21429437>
- [3] J.E. Hastings, R.A. Barkley. 1978. A Review of Psychophysiological Research with Hyperkinetic Children. *Journal of Abnormal Child Psychology*, 6(4): 413-447.
- [4] P. Lang. 1995. The emotion probe: Studies of motivation and attention. *American Psychologist*. 50(5): 372-385.
- [5] C. Latulipe, E.A. Carroll, and D. Lottridge. 2011. Love, hate, arousal and engagement: exploring audience responses to performing arts. *Proceedings of CHI*, pp. 1845-1854.
- [6] P. Mirza-Babaei, L. E. Nacke, J. Gregory, N. Collins, and G. Fitzpatrick. 2013. How does it play better? exploring user testing and biometric storyboards in games user research. *Proceedings of CHI*, 1499-1508.
- [7] G.P. Quinn, M.J. Keough. 2002. *Experimental Design and Data Analysis for Biologists*.
- [8] J. Radbourne, K. Johanson, H. Glow, and T. White. 2009. Audience experience: measuring the quality in the performing arts. *International Journal of Arts Management*, 11(3): 16-29.
- [9] S. Stenslie. 2013. Towards Telehaptic Performativity. *Remote Encounters: Connecting bodies, collapsing spaces and temporal ubiquity in networked performance*.
- [10] H. Stone, J. Sidel, S. Oliver, A. Woolsey, and R. C. Singleton. 1974. Sensory Evaluation by Quantitative Descriptive Analysis. *Food Technology*, 28. 24-34.