The Perception of Live-sequenced Electronic Music via Hearing and Sight

Mehmet Aydın Baytaş
Koç University – Arçelik Research Center for Creative Industries
Koç University
Istanbul, Turkey
mbaytas@ku.edu.tr

Tilbe Göksun
Department of Psychology
Koç University
Istanbul, Turkey
tgoksun@ku.edu.tr

Oğuzhan Özcan
Koç University – Arçelik Research Center for Creative Industries
Koç University
Istanbul, Turkey
oozcan@ku.edu.tr

ABSTRACT

In this paper, we investigate how watching a live-sequenced electronic music performance, compared to merely hearing the music, contributes to spectators' experiences of tension. We also explore the role of the performers' effective and ancillary gestures in conveying tension, when they can be seen. To this end, we conducted an experiment where 30 participants heard, saw, or both heard and saw a live-sequenced techno music performance recording while they produced continuous judgments on their experience of tension. Eye tracking data was also recorded from participants who saw the visuals, to reveal aspects of the performance that influenced their tension judgments. We analysed the data to explore how auditory and visual components and the performer's movements contribute to spectators' experience of tension. Our results show that their perception of emotional intensity is consistent across hearing and sight, suggesting that gestures in "non-instrumental" live-sequencing can be a medium for expressive performance.

Author Keywords

Live-sequencing, spectator experience, tension, experiment, functional data analysis, eye tracking

ACM Classification

H.5.5 [Information Interfaces and Presentation] Sound and Music Computing

1. INTRODUCTION

Live-sequencing is a term that describes the currently prevalent paradigm for performing electronic music among mainstream professional artists. This way of performing music is characterized by a single performer using (hardware or software) synthesizers, samplers, and drum machines to generate or play back sound loops, while manipulating timbres and the musical form using audio mixers, effects, MIDI controllers, and recording equipment. This allows performers to assume a role that borrows practices from conductors, mixing engineers, DJs, and instrumentalists.

The *live-sequencing* paradigm has been explored by composers and performers since the 1940s [2], and currently it constitutes the norm among popular professional live electronic music performers [5]. Some authors consider interactions between live-sequencing performers and their musical devices "non-instrumental" and distinct from traditional instrument-playing, to the point that the use of the term "gesture" can be disputed in this context [12]. We would argue that gestures in live sequencing, even with user interfaces where interactions are limited to twisting knobs and pressing buttons, can convey expressive intent and warrant investigation. One particular aspect of this paradigm that has not yet been explored is the *spectator experience*. Previous research has looked into the experience of spectators who *watch* performances in the context of the "instrument-playing" paradigm or using qualitative methods [1, 3, 4, 6, 9, 10, 11, 12, 13]. However, in the context of live-sequenced music, a quantitative exploration of how watching a performer's interactions with their equipment can contribute to the audience experience remains an open question.

In this paper, we investigate the *spectator experience* of a *live-sequenced electronic music performance*. Specifically, we investigate how *watching* a performance, compared to merely hearing the music, contributes to the spectator's experience of *tension*. We also explore in how a performer's *effective gestures* and *ancillary gestures*—i.e. movements that must be performed to produce and manipulate sound, and movements that do not influence sound [19]—contribute to the experience of tension when they are seen.

To this end, we adapted an experiment design from previous work in cognitive psychology [18]. We randomly assigned 30 participants to one of three conditions—audio only (AO), visual only (VO), and both (AV). Each participant was presented the recording of a live-sequenced techno music performance, while they delivered real-time, continuous judgments on their experience of *tension*—a one dimensional proxy for the many factors that contribute to emotional response. Eye tracking data was also recorded during the task for the AV and VO groups. Our analyses reveal how watching a performance contributes to the audience's experience of emotion in music, and demonstrate a quantitative approach to evaluating the spectator experience. We used the data we collected to investigate the following research questions:

- *RQ1:* How do the experiences of tension (i.e. the magnitudes and contours of the tension curves) conveyed via sight, sound, and both compare?
- *RQ2:* How do performers' effective and ancillary gestures influence the experience of tension conveyed via sight?

NIME'16, July 11-15, 2016, Griffith University, Brisbane, Australia.
.

To most experiment particioants, the nature of the mappings in the musical devices in the video would be opaque [8]. This initially suggests that the resulting tension judgments would be dissimilar across presentation conditions. Any existing similarities could be expected to involve the performer's ancillary gestures, where features like rhythm and amplitude can be similar across modalities [18].

2. THE EXPERIMENT

Our experiment design is based on a 2006 paper by Vines et al. on the psychology of music perception [18], which investigates how the auditory and visual modalities and their interactions influence the perception of emotion in a clarinet performance. We adapted the experiment to investigate the same phenomenon in the context of live-sequenced electronic music. In our experiment, 30 participants heard, saw, or both heard and saw a live-sequenced music performance video. As a one-dimensional representation of the emotional intensity in music, we collected judgments on the *tension* experienced during the performance. Judgments were expressed and recorded real-time, during stimulus presentation, using a linear potentiometer. Augmenting the 2006 design to gain insights into which aspects of the performance captured participants' attention, we also collected eye tracking data during stimulus presentation.

2.1 Participants

30 participants (20 male, 10 female; ages 18 to 33, mean age 24 years) were recruited for the experiment. Per the 2006 paper, we ensured that all participants had at least 2 years of experience playing a musical instrument, in addition to experiences in DJing, music production, studying music theory, and teaching music. This criterion ensured that participants could readily feel comfortable with the musical term "tension." The literature [18] reports, and our own pilot testing confirms, that lay and musically experienced participants perform similarly on the tasks in this experiment, and the results generalize.

Participants were randomly assigned to one of three equally-sized treatment groups. The *Auditory Only (AO)* group experienced only the audio from the performance excerpt; they saw a blank (black) screen as they listened. The *Visual Only (VO)* group saw the performance with the audio removed. The *Auditory & Visual (AV)* group experienced the performance excerpt with both audio and visuals intact, as one would normally experience a performance.

2.2 Stimulus

Our stimulus comprised the final 4 minutes and 30 seconds of a video published on YouTube by the musician KiNK[1]. The video[2], titled *KiNK BeatStep Pro + Marsh + Microzwerg + Kraftzwerg*, was uploaded to the artist's own channel on July 1, 2015. We used an excerpt, not the whole video, to avoid exhausting the participants. (The entire video is over 13 minutes long.)

In the video, the artist live-sequences improvised techno music using three analog synthesizers (one of them a drum synthesizer), a delay effect pedal, an analog mixer, and an *Arturia BeatStep Pro* multi-purpose MIDI controller (more details on the setup can be found in the video's description). This video was selected since it is shot from the side, so that the performer's face and upper body, along with the controls on the musical devices he uses, are visible. Moreover, throughout the video, the camera does not move, and there are no other people visible; eliminating two factors that would confound the measurement of tension experiences.

[1]pbpm.net/artists/kink
[2]youtu.be/XtT00rxFsmc

2.3 Task

Participants performed the same task during stimulus presentation, twice: once for practice, and once for data collection. This ensures that participants familiarized themselves with the input device, and conforms to the procedure described in Vines et al.'s 2006 paper [18].

Participants were instruced to move a linear potentiometer between its minimum and maximum values to express their experience of tension in the performance. We showed the participants the same on-screen instruction used in the 2006 experiment, which reads as follows: *"Use the full range of the slider to express the TENSION you experience in the performance. Move the slider upward as the tension increases and downward as the tension decreases. Begin with the slider all the way down."* [18]

2.4 Apparatus

Tension judgments were collected using the one of the sliding potentiometers on an *Akai APC20* USB MIDI controller. Participants in the AV and AO conditions were given a pair of *AKG K271 Studio* headphones to hear the sound. A *Tobii T120* display-integrated eye tracker was used to collect gaze data from the AV and VO groups. Figure 1 depicts this setup.

All devices were connected to the same PC, running Windows 7. Slider values were sampled every 50ms and saved to a text file using a *MATLAB* script. *Tobii Pro Studio* software was used to present the stimuli and capture eye tracking data. MIDI and eye tracking data both included millisecond-level timestamps from the operating system, which were used to synchronize the two data streams.

3. RESULTS

To investigate our research questions, we conducted the following analyses on the data collected from the participants:

Visual inspection and correlation analysis. We inspected and compared the mean tension judgment curves for all three groups, relating the data to the music and the performers' gestures where relevant. We then computed the Spearman's ρ values between the means of each presenta-

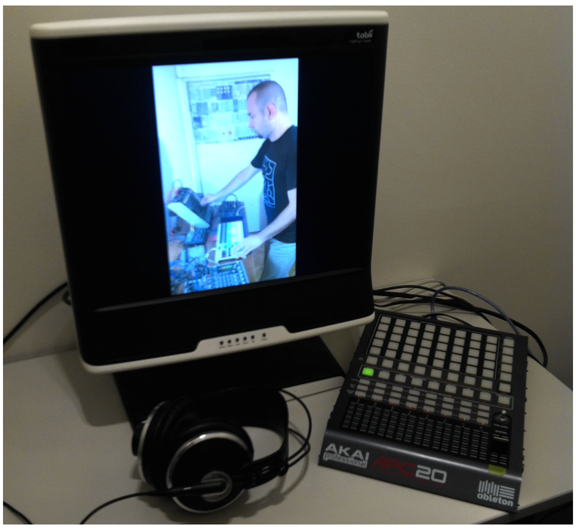


Figure 1: The experiment setup: MIDI controller, display, and head- phones connected to a PC. The frame shown on the screen is part of the stimulus used for the experiment.

Table 1: Spearman correlation values between group mean tension judgments from the present study with a live-sequenced techno performance, with results from the 2006 study with two clarinet performances [18] included for comparison.

	Live-sequenced Techno			Clarinet "Performer R" [18]			Clarinet "Performer W" [18]		
	AO	**VO**	**AV**	AO	VO	AV	AO	VO	AV
AO	**1**	**0.83****	**0.90****	1	0.13**	0.75**	1	0.15**	0.91**
VO	–	**1**	**0.83****	–	1	0.32**	–	1	0.17**
AV	–	–	**1**	–	–	1	–		1

**$P < 0.01$, two-tailed.
AO, VO, AV: Audio-only, visual-only, and audio-visual conditions.

tion condition. This statistic is preferable to the Pearson correlation r for data with serial correlation [17, 18].

Cumulative gaze heatmaps. We generated heatmaps from the eye tracking data throughout the AV and VO conditions. We examined the heatmaps to see what visual aspects of the performance—i.e. effective gestures vs. ancillary gestures—captured spectators' attention. Throughout the stimulus, effective gestures appear as hand movments that manipulate the musical devices, while ancillary gestures largely correspond to the performer's head and torso movements.

Functional data analysis (FDA) and segments of interest (SOIs). FDA comprises statistical approaches for analyzing data that result from continuous underlying processes [14, 15]. In FDA, discrete data are modeled as continuous functions, whereupon analyses are performed. We fitted functional curves to the data and performed functional significance testing to identify *when* the differences between AV and AO conditions are significant; i.e. we isolate the effect of visual information by comparing the baseline AV condition to the case where visual information is removed. Thus we identify SOIs within the stimulus. The SOIs indicate where the experiences of tension conveyed via vision and sound are significantly different. We generated gaze heatmaps for the SOIs and a 2-second onset before each SOI, to see where spectators attend during these segments.

Tension data were truncated on both ends to remove the first 5 seconds and a small duration at the end. The regions between 5,000ms and 269,000ms were used for all analyses.

(a) AV (b) VO

Figure 2: Cumulative gaze heatmaps for the AV and VO conditions show the visual features that spectators attended throughout the performance. The white horizontal lines on both panels represent the linear classifiers we used to quantify the attention to effective vs. ancillary gestures.

This is to remove occasional errors that arise during the initialization of the sampling device and evade small timing inconsistencies between software. For our purposes, errors from both sources may be neglected after truncation.

3.1 Visual Inspection and Correlation Analysis

The mean tension judgments for the three presentation conditions are shown in Figure 3 (top panel). Overall, their appearances are very similar. There are pronounced similarities between the contours of the AV and AO curves throughout. The VO curve diverges from the other two and sits a little higher in terms of magnitude during the second and third minutes of the performance, but retains a similar contour overall. This initial inspection corroborates Vines et al.'s conclusion that "sound" plays "the dominant role in determining the contour and trajectory of the emotional experience for those who could both hear and see" the performance [18].

The first and last minute of the performance comprise many improvised, idiosyncratic and sudden variations to rhythm and timbre. In response to these rapid events, tension increases for the AV and AO conditions are pronounced during the first and last minutes of the performance. Elsewhere, the performance mostly consists of more subtle variations on loops, and the tension changes appear to be more moderate.

Table 1 shows the Spearman correlation values between condition pairs, which range from 0.83 to 0.90 for our experiment. All three conditions appear to be highly and positively correlated with each other, with very high statistical significance. The mean AV - AO judgments are the most highly correlated, which is consistent with the findings for a clarinet performance. Overall, compared to clarinet performances, the correlations between the experiences of tension originating from hearing and/or sight appear to be much higher for the live-sequenced electronic music performance.

3.2 Cumulative Gaze Heatmaps

Figure 2 shows the heatmaps generated from gaze data from the AV and VO groups. A preliminary inspection reveals that, compared to the AV condition, the performer's head has been a more frequent target of attention for VO participants. In contrast, the AV group has attended more frequently to the music-making devices on the table and the performer's effective gestures. This effect can be quantified simply using a horizontal linear classifier at 450px from the top edge of the image, shown on Figure 2 as a white line. We have calculated that, for the AV condition, only 9% of the gaze points from all participants reside above the 450px line, compared to 22% for the VO condition. We may presume this is due to a perceived cause-effect relationship between the sound and the devices in the AV condition—without sound, attention diverts more easily to movement. Finally, it is interesting that the performer's torso movements do not appear to be an area where gaze points have accumulated considerably, compared to the musical devices (where effective gestures occur) and the head.

3.3 Functional Data Analysis

Using Ramsay's FDA package for *MATLAB*[3], we modeled the data from all three groups as functional objects using 600 6th order B-splines, and a a smoothing value $\lambda = 0.1$. The root mean square (RMS) error for all functional models did not exceed 4% of the full range of tension values. The data and the fitted functional curves for the AV tension judgments are shown in the accompanying video. While there are visible erroneous high-frequency oscillations in the functional models at the edges of where tension values plateau, these are negligible for the purposes of the subsequent analysis (per [18]).

We compared the tension jugdments from the AO and AV groups using a functional significance test to determine *when* the difference between the two groups is significant; i.e. we isolated the effects of visual information. A t-value was calculated as a function of time over the duration of the performance. When the calculated t-value rises above the pointwise critical level, the difference between AV and AO is deemed to be significant. Figure 3 shows the results of the functional t-test, aligned with the mean tension judgments from the three groups.

We identified four SOIs where the difference between the AV baseline condition and the AO condition are significant at a 85% confidence level:

- *SOI1:* 2:51 - 2:54 (11:40 - 11:43 in the YouTube video)
- *SOI2:* 3:30 - 3:33 (12:19 - 12:22 in the YouTube video)
- *SOI3:* 3:54 - 4:03 (12:43 - 12:52 in the YouTube video)
- *SOI4:* 4:15 - 4:21 (13:04 - 13:10 in the YouTube video)

At the more conventional 95% and 90% confidence levels, the test does not detect any statistically significant differences between the conditions; which is consistent with our previous result that the correlation between the AV and AO tension judgments is high. While we concede that a 85% confidence level is low for most purposes, it does offer clues for selecting SOIs from the stimulus and eye tracking data for further qualitative analysis.

We can examine auditory and visual components of the stimulus along with (AV) gaze heatmaps for the durations of these SOIs (Figure 4) to gain insights on what causes tension judgments to diverge between the two groups. We also included data from 2-second durations that precede the SOIs, to account for possible delays between stimulus presentation and participants' physical responses [18].

SOI1. During this segment, KiNK first adjusts knobs on the bass synthesizer with one hand for subtle timbral changes. In the final second, his hand switches to the "roller" pad on the BeatStep Pro controller for dramatic effect. His head and torso movements consist of subtle rhythmic bobbing, which halts abruptly and reverses direction when he touches the "roller" pad.

SOI2. Here, a prolonged "roll" effect is applied to the percussion using the BeatStep Pro. Keeping his left hand on the "roller" pad, KiNK reaches for the drum synthesizer to adjust the timbre of the "rolling" hi-hat sound for a few seconds, and then manipulates the "roller" pad to release the hi-hat and "roll" a synthesized tom-tom note. His posture is erect and stable, with very minimal rhythmic movement.

SOI3. For the duration of SOI3, KiNK presses buttons on the bass synthesizer to idiosyncratically adjust the octave of the notes playing on the device. Notably, his posture appears somewhat stiff, small rhythmic head oscillations, and his attention is focused on the bass synthesizer.

SOI4. Similar to SOI2, this segment consists of the artist improvising on the "roller" pad, repeating percussion sounds while manipulating their timbre using knobs on the drum synthesizer. His head and body does not move very much during this gesture.

The stimulus and participants' gaze behavior shows interesting similarities across all SOIs. What is remarkable about these segments of the performance is that three of them correspond to points where the performer is utilizing the "roller" function on the *Beatstep Pro* controller to rapidy repeat a note or percussion hit. Effectively, we may say that these are moments when the controller is being "played" like an instrument, rather than being used to construct or manipulate loops. Another interesting similarity is that the performer's head and body movements are quite stiff during the SOIs. Throughout most of the performance, the opposite is the case: KiNK's head and body movements often follow unsubtle oscillations that mirror the "pulse" of the music. Contrarily, looking at the gaze heatmaps, we observe that participants have been looking more commonly at the devices being manipulated rather than KiNK's gestures. Spectators' attention seems to be focused more on the device(s) being manipulated during each segment, as well as the blinking lights on the sequencing controller. However, it is likely that spectators respond to ancillary gestures detected via peripheral vision.

4. DISCUSSION

We reported on an experiment we conducted to investigate how *watching* a live-sequenced music performance, compared to merely hearing the music, contributes to the spectator's experience of *tension*. Here, we return to our research questions and re-iterate how our results serve to illuminate them.

(RQ1) What are similarities and differences between the experiences of tension elicited by audio, video, and both?

Tension judgments arising from hearing, seeing, and both hearing and seeing a live-sequenced techno performance have remarkable similarities, both statistically and in terms of appearance. That the similarities are more pronounced between the AV and AO conditions corroborates Vines et al.'s results for the same task with classical clarinet performances as stimuli [18]. However, the similarity of tension experiences between vision alone and the other two conditions is stronger for live-sequenced techno. This can be attributed to a discernible "pulse" in the music—and the coordination of performers' body movements to this rhythm—which can provide spectators with visual cues that correspond to the musical structure [18]. Our results can thus serve to confirm that in the context of live-sequenced techno music, *the experience of tension elicited by hearing the music and watching a performer appear to be similar.*

When the performance is both seen and heard, *sound appears to be dominant in influencing the experience of tension.* A comparison of the pairwise correlations supports this, as do the visual appearances of the curves. Moreover, in this regard, we have shortly debriefed our experiment participants following the completion of the task, asking them to "name some of the ways in which tension was conveyed" to them during the performance. We observed that, in the case of the AV group, properties related to sound (e.g. arrangement density, high frequency content, loudness, melodies and harmony...) were cited much more fre-

[3]functionaldata.org

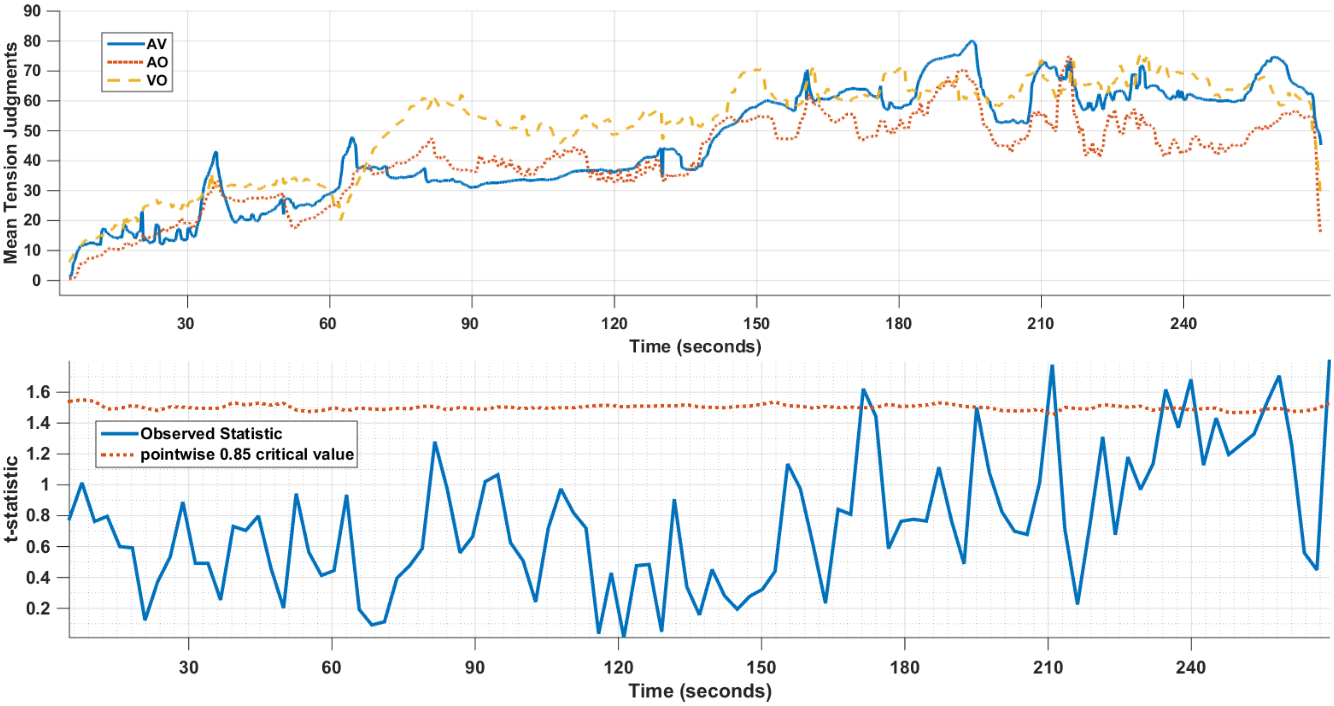


Figure 3: Mean tension judgments for the three presentation groups (top), and the functional significance test (bottom).

quently than visual aspects (e.g. effective gestures, head movements, lights on devices...).

The visual appearances of the group mean tension curves suggests that *watching a performer can serve to both escalate and dampen the experience of tension.* During portions of the performance where the artist frequently manipulates notes and timbres, the three mean curves become similar, and magnitude of the AV tension curve is consistently higher than the AO curve. During portions where the performance comprises more subtle manipulations on a looping groove, the VO tension curve rises above the other two cases, and the AV curve falls below the AO mean. One interpretation of this could be that the experience of tension is more consistent across modalities when the effects of effective gestures are pronounced; but when obvious gestures have subtle results, the combination of sound and image dampens spectators' experiences of tension.

(RQ2) What are the roles of the performer's effective and ancillary gestures in conveying tension?

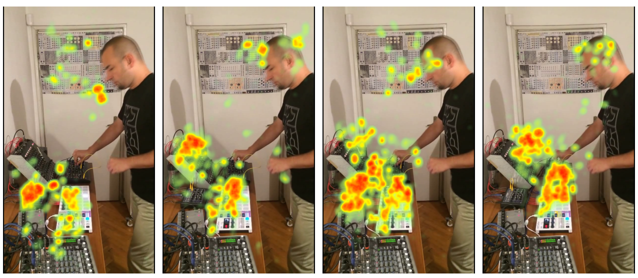
For the AV group, who experienced the stimulus as one would normally watch a performance, the gaze data shows that *effective gestures and musical devices become a more frequent center of attention.* When sound is removed, the tendency to focus on the performer's head movements increases. The presence of sound appears to draw attention to the apparent source of the sound, while larger movements capture spectators' attention when sound is absent.

SOIs where tension judgments from the AV and AO groups are significantly different offer two interesting insights regarding the role of effective and ancillary gestures. First, a lack of head and body movement, in contrast to the performer's rhythmic bobbing throughout much of the video, appears in all four SOIs. On the other hand, during three of the four SOIs, we observe a "roll" effect being "played" like an instrument, which can be considered distinct from the "non-instrumental" performance style that pervades the rest of the piece. We may say that *the experiences of tension evoked by hearing and normally "watching" tend to converge during live-sequencing with loops, and diverge when switching to an "instrumental" approach.*

What is the significance of these results?

In the video we showed to our participants, KiNK interacts with his equipment to make music using knobs, sliders, and buttons. Such interactions are often considered distinct from musical intruments; they have been described as "non-instrumental," and even the use of the term "gesture" in this context can be disputed [12]. However, despite such distinctions, our results imply that *"non-instrumental" interfaces comprising knobs and buttons can support expressive gestures that convey emotional intensity.* Contrary to our initial expectations, the experiences of emotional intensity arising from hearing the music and seeing it being performed on these devices are remarkably similar across all three modalities.

While similar phenomena have been previously investigated in other musical contexts [4, 6, 18], to our knowledge, our work is the first that quantitatively explores the role of visual and auditory modalities in spectator experience for the "non-instrumental" live-sequencing context. One of our novel findings is that the similarities in the experience of tension persist across all three modalities in the context



SOI1 SOI2 SOI3 SOI4

Figure 4: Gaze heatmaps show what the AV group attended to during SOIs where statistically significant differences in tension judgments exist between AV and AO groups.

of live-sequenced techno, compared to previous studies of instrumental interactions where disparities between visual-only and audio-only conditions have been found. We believe our results can be relevant in provoking designers of live-sequencing devices to consider the role of human gesture in performances while developing new concepts.

Performances of live-sequenced music out in the wild often feature stage designs and manners of performing that conceal effective gestures and much of the performers' body movements, perhaps aiming for a "magical" or "secretive" tone [16]. Conversely, our results support that *revealing performers' instruments and gestures can serve to modulate the spectator experience*, and even devices built out of commodity transducers (knobs, buttons, etc.) can still be designed to afford expressivity [7] and communicate emotion to spectators. For instance, the form and layouts of "non-instrumental" controllers, as well how these devices are arranged on stage, could be designed to consider performers' gestures and to enhance the spectator experience.

5. CONCLUSION

We have conducted an experiment to investigate how spectators' experiences of tension in a live-sequenced electronic music performance differs or agrees across sensory modalities, and to investigate the role of effective and ancillary gestures in conveying tension via vision. To our knowledge, for this context, this is the first quantitative exploration of the roles of audition, vision, and performer gestures in influencing the spectator experience. Our findings indicate that the experiences of tension conveyed via sight, hearing, and both are very similar in this context; even more so compared to previous studies of "instrumental" contexts. They also support that the performer's ancillary gestures—e.g. rhythmic head and torso movements—contribute to the experience of tension when they can be seen. However, we observe that spectators' attention remains on effective gestures for the most part.

These findings are significant in that they expose the significance of gesture in "non-instrumental" performance contexts (e.g. live-sequencing) as medium for communicating expressive intent. This can prompt designers of user interfaces for such contexts to consider how gestures and interactions can contribute to expression in performance in developing new concepts. Moreover, the methods we demonstrate in this paper (e.g. continuous judgment elicitation, eye-tracking, and functional analysis) can be used to further explore this field, and they can be adapted to study other domains as well.

Most spectators may not understand the mappings between gestures and sonic events in live-sequencing and NIME performances, but our results show that their perception of the essence of the performance—the emotional intensity—is consistent across hearing and sight. Continuing to systematically explore this performance context where expression does not seem to be hindered by a lack of comprehension can yield interesting knowledge informing NIME development, as well as other design contexts.

6. ACKNOWLEDGMENTS

We thank the anonymous participants in our experiment, our colleagues and reviewers who provided feedback, Dr. Daniel Levitin, Dr. James Ramsay, Dr. İstem Köymen Keser, İdil Bostan, and Damla Çay for their contributions.

7. REFERENCES

[1] J. Barbosa, F. Calegario, V. Teichrieb, G. Ramalho, and P. McGlynn. Considering audience's view towards an evaluation methodology for digital musical instruments. In *Proc. NIME 2012*.

[2] B. Bengler. *The audio mixer as creative tool in musical composition and performance*. Institut fur Elektronische Musik und Akustik / Universitat für Musik und darstellende Kunst Graz, 2011.

[3] A. Bergsland and T. Åse. The 'e' in nime: Musical expression with new computer interfaces. In *Proc. NIME '06*, 2006.

[4] M. Broughton and C. Stevens. Music, movement and marimba: an investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music*, 37(2):137–153, 2009.

[5] K. Caryl. What is [live]? – the new generation of live techno. *Attack Magazine*, 2013. Retrieved January 10, 2015 from attackmagazine.com/features/long-read/what-is-live-new-generation-live-techno/.

[6] S. Dahl and A. Friberg. Visual perception of expressiveness in musicians' body movements. *Music Perception*, 24(4):433–454, 2007.

[7] C. Dobrian and D. Koppelman. The 'e' in nime: Musical expression with new computer interfaces. In *Proc. NIME '06*, 2006.

[8] S. Fels, A. Gadd, and A. Mulder. Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, 7:109–126, 8 2002.

[9] A. C. Fyans, M. Gurevich, and P. Stapleton. Examining the spectator experience. In *Proc. NIME 2010*.

[10] A. C. Fyans, M. Gurevich, and P. Stapleton. Where did it all go wrong? a model of error from the spectator's perspective. In *Proc. NIME 2009*.

[11] A. C. Fyans, M. Gurevich, and P. Stapleton. Spectator understanding of error in performance. In *CHI EA '09*, 2009.

[12] M. Gurevich and A. Cavan Fyans. Digital musical interactions: Performer–system relationships and their perception by spectators. *Organised Sound*, 16:166–175, 6 2011.

[13] C.-H. Lai and T. Bovermann. Audience experience in sound performance. In *Proc. NIME 2013*.

[14] D. J. Levitin, R. L. Nuzzo, B. W. Vines, and J. Ramsay. Introduction to functional data analysis. *Canadian Psychology*, 48(3):135–155, 2007.

[15] J. Ramsay and B. W. Silverman. *Functional Data Analysis*. Springer, 2005.

[16] S. Reeves, S. Benford, C. O'Malley, and M. Fraser. Designing the spectator experience. In *Proc. CHI '05*, 2005.

[17] E. Schubert. Correlation analysis of continuous emotional response to music: Correcting for the effects of serial correlation. *Musicae Scientiae*, 5(1 suppl):213–236, 2002.

[18] B. W. Vines, C. L. Krumhansl, M. M. Wanderley, and D. J. Levitin. Cross-modal interactions in the perception of musical performance. *Cognition*, 101(1):80–113, 2006.

[19] M. M. Wanderley, B. W. Vines, N. Middleton, C. McKay, and W. Hatch. The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, 34(1):97–113, 2005.