Measuring Impact of Social Presence Through Gesture Analysis in Musical Performances

Simon Schaerlaeken, Donald Glowinski and Didier Grandjean
Neuroscience of Emotion and Affective Dynamics Lab, Switzerland

ABSTRACT

Immersive Virtual Environments combined with motion capture systems have been used as experimental set-ups for studying the influence of the presence of an audience on musicians' performances. This study highlights that musicians playing with different expressive manners move differently, through increasing their kinetic energy and body twisting. These factors are increased or decreased by the presence of an audience, depending on the difficulty of the task. Such behavior fits Zajonc's theory of social facilitation.

KEYWORDS

Immersive Virtual Environments, Music, Gesture, Movement, Expressive Gesture, Audience, Stress

ACM Reference format:

. 2016. Measuring Impact of Social Presence Through Gesture Analysis in Musical Performances. In *Proceedings of 4th Internalional Conference on Movement Computing, London, UK, Mai 2017 (MOCO'17),* 5 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

An important aspect of a musician's training involves learning to cope with stress during performances. However, with the exception of live performances and peers sessions, few tools are available for musicians to train with. In addition, researchers interested in the effect of audiences on musicians and their body language are left with uncontrolled and highly variable situations when they study concerts. On a practical side, recording physiological and motion capture data during the performance can hinder musicians during the concert, thus impacting the quality of their playing, making concerts a challenging environment for scientific research.

Immersive Virtual Environments (IVEs) have been used by researchers to control for these complex parameters [23]. IVEs allow researchers to create realistic virtual environments with unlimited configurations that can adapt in real time to users' behavior. It allows for researchers to control for environmental parameters such as audience, space, and light. This environment combined with a motion capture set-up was used in this study to precisely record expressive musical gesture and explore the possible underlying behavioral mechanisms impacted by the presence of an audience in the context of different levels of expressive intents.

MOCO'17, London, UK

DOI: 10.1145/nnnnnnn.nnnnnnn

1.1 Controlling environmental variables

Musicians develop unique abilities to allow them to adapt their behavior to different social and environmental contexts [23]. Consequently it is necessary to control for many parameters when recording musical performances, e.g. sounds, lights, the presence of other musicians or audience. Thus IVEs represent a key methological tool for psychological research as it can provide greater experimental control, more precise measurement, ease of replication across participants, and high ecological validity, making it extremely attractive for researchers [2, 20]. Similar methods have already been used in research for treating phobias such as fear of flying[18], arachnophobia [3], and for patients suffering from post-traumatic stress disorder [8]. Few studies have considered using virtual environments in music to study performance anxiety [1, 16]. In a recent study by Williamon et al [23], musicians were invited to enter IVEs to train their ability to cope with the pressure of performing live and rated the system as useful for developing their performance skills. This study demonstrates that simulated environments are able to offer a realistic experience of performance contexts.

1.2 Music performance and body expressive movements

Musicians make bodily gestures and move during performances according to i) communicate their expressive intentions, and ii) to express their feelings without intend to communicate them [5]. Musical gestures are mainly made to produce sounds but are also used by the musician as means to convey or express emotion (see review Expressive Gesture [10, 17]. Gestures contribute to communicate information to the audience as well as the other musicians. Davidson [7] demonstrated that certain perceptual elements of a musician's gesture are sufficient for the audience to identify a musician's expressive intend. He suggested the use of three level of expressiveness to be able to study the link between expressive gesture and musical performance: (1) without expression, labelled as "deadpan"; (2) with normal expression/concert-like, labelled as "projected"; and (3) with exaggerated expression, labelled as "exaggerated". Some body parts have been reported to convey the most expressive information, specifically head, shoulders, arms, and torso [15, 19, 21, 22]. Some motion features have also been associated with expressive motion such as the kinetic energy. It helps understand broad, unrefined body reaction and gives a first glimpse of the behavioral components of expressive gestures. It might help understand how musicians cope with the audience [11–13]. For example, musician facing an audience and playing in exaggerated expressive manners could be affected by the amount of supplementary stress caused by the difficulty of the task.

2 METHOD

Four violonists participated in this experiment (3 females, mean age = 22). They were asked to play 30 second long interpretations of *la partita* in r minor from *la Sarabande* composed by Bach. The part of the musical score to be played was carefully selected to correspond with complete musical phrases. The excerpts were interpreted according to three selected expressive intentions: deadpan, projected, and exaggerated [7]. Musicians performed inside an Immersive Virtual Environment (IVE) with the use of a system of three screens, seven TITAN QUAD 3D projectors, and stereo glasses presenting seamless and perspectively coherent 3D images. Two different virtual environments were created for this experiment: a room filled with an audience behaving naturally and attending the concert, and the same room without the audience (FIGURE 1). The different combinations of conditions were performed in a pseudo randomised order by each musician. 6 excerpts were recoded for each of the 4 musicians, adding up to a total of 24 excerpts.

Figure 1: Motion capture data recorded in front of the two possible virtual environments: *left*, the empty room and *right*, the room with the audience.

Motion capture data was recorded for every piece using a Vicon optical motion tracking system composed of eight Bonita 3 cameras. A total of 26 markers were used, covering selected body parts based on recent literature on the analysis of music performance, i.e. the head, arms, and torso [15, 19, 21, 22].

Drawing upon the studies by [6, 9], we considered two types of expressive body features: the kinetic energy and the Body Twist Index (BTI). The former helps understand broad, unrefined body reactions and gives an glimpse of the behavioral components of expressive gesture. The latter captures body shape related information, i.e. the relative displacement of body limbs one with respect to another. In the case of violin player, this second feature is critical since the upper and lower part of the body are dissociated. Violinists tend to twist their body more while playing compared to cello players for example. We hypothesized that these two features can help modeling critical changes in musician's expressive responses to audience presence.

Both features were calculated using our own MATLAB© toolbox (soon in open access). The kinetic energy was computed for every marker of the motion capture data and then averaged across markers and over time (FIGURE 2). The Body Twist Index was represented by the average angle between the pelvis and a perpendicular line to the shoulders, considering only the top quantile (above quantile 75%) of the data for each excerpt (FIGURE 3). Both features were z-scored per musician.

Linear mixed models were used in this study for two reasons: they incorporate random effects and they allow us for handling correlated data and unequal variance [14]. Comparing model was done using Chi-squared testing. All p-values were corrected using FDR.

3 RESULTS

3.1 Main effect of the expressive intent

The first model computed was estimating the z-scored value of the two features for each expressive intention: deadpan, projected and exaggerated expressive manners. For both features, a model containing the expressive intent level as fixed effect with the musician playing as random effect outperfomed significantly a model with only the random effect ($\chi^2_{kineticenergy}(2, N = 24) = 54.889, p = 1.205e^{-12}, AIC = -28.750, BIC = -22.860, R^2m = 0.90, R^2c = 0.90$ & $\chi^2_{BTI}(2, N = 24) = 16.053, p = 3.267e^{-4}, AIC = 2.521, BIC = 8.4112, R^2m = 0.50, R^2c = 0.50$)[1].

For both features, any increase in the musician's expressiveness is represented by an increment of these features' value (except between the projected and exaggerated expressive manners in the case of Body Twist Index)(FIGURE 4). All contrasts comparing two expressive intentions at the time were significant ($p < 0.001$).

3.2 Interaction effect of the expressive intent and the presence of an audience

The second model estimated the z-score values of both features using the expressiveness and the interaction between the expressiveness and the presence of an audience as fixed effects, with the musician playing as a random effect. Only the model using the Body Twist Index was significantly different when compared to the previous model using the main effect, indicating an interaction effect between our two factors ($\chi^2_{kineticenergy}(3, N = 24) = 4.8495, p = 0.183, AIC = -27.599, BIC = -18.175, R^2m = 0.92, R^2c = 0.92$ & $\chi^2_{BTI}(3, N = 24) = 9.3372, p = 0.025, AIC = -0.81625, BIC = 8.6082, R^2m = 0.66, R^2c = 0.66$).

While the value of both features increases with stronger expressiveness, the impact of the presence of an audience appears to be expressive-dependant (FIGURE 5). In the case of deadpan expressive manners, the presence of the audience is characterized by an increase in both features (even though only the difference in BTI is significant ($p < 0.01$)). In opposition, during the exaggerated expressive manners, we noticed a trend of both features decreasing when a virtual public is present during the performance.

4 DISCUSSION

In this study, we highlighted the changes in motion features associated with different expressive manners. These findings are consistent with literature showing that the three expressive styles are characterized by different body movements and values of kinetic energy [4, 7]. The kinetic energy gradually increased from deadpan, to projected, to exaggerated expressive manners.

The relationship between body parts was impacted by the different levels of expressiveness as well. The Body Twist Index increased

[1]We report the effect sizes according to the approach of Nakagawa & Schielzeth (2013), implemented in the fiMuMInfi R package.

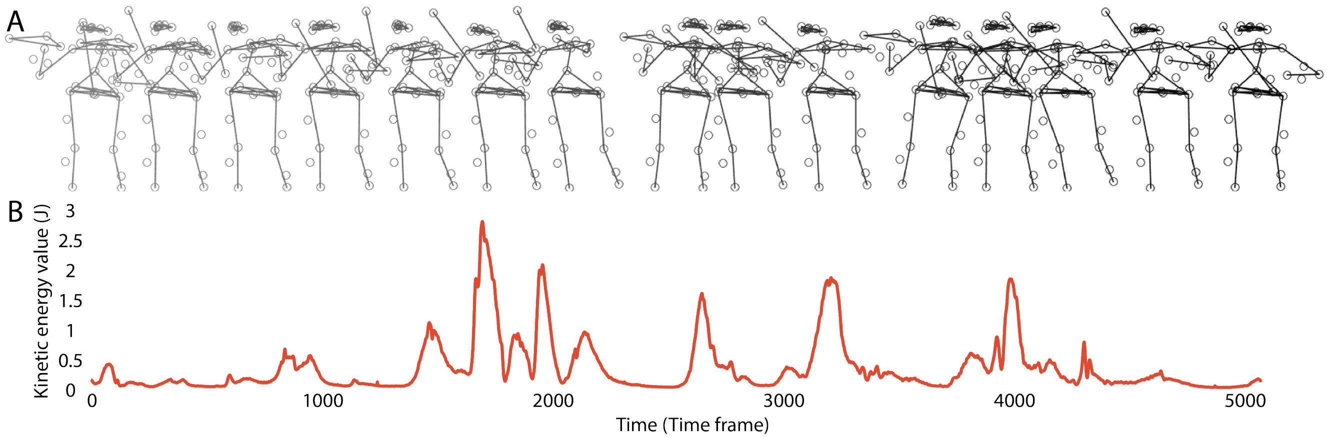

Figure 2: A, Chronograph of the motion capture data for one excerpt played in front of an audience while exaggerating expressive intent. B, Kinetic energy associated with movements of all the markers for the excerpt depicted in A.

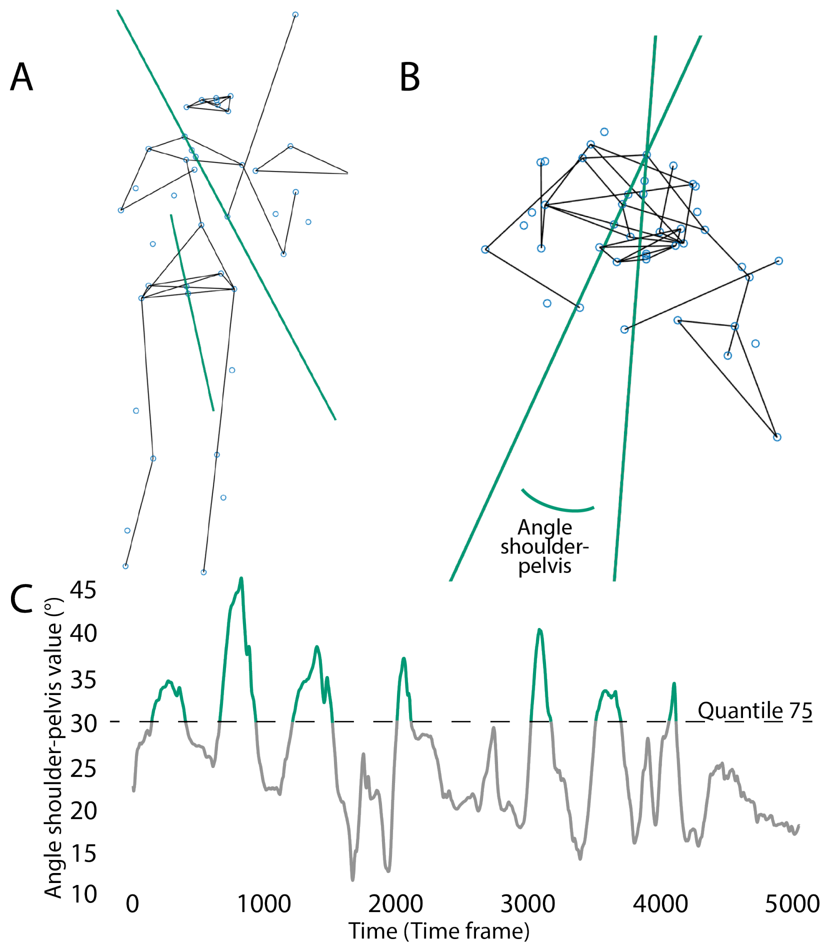

Figure 3: A, 3D view of the motion capture and the line associated with the pelvis and shoulder. B, Transversal view of the motion capture data and the angle computed between the line from the pelvis and shoulders. C, Computation of Body Twist Index. The angle between the aforementioned lines is computed over the duration of the excerpt. The Body Twist Index consists in the average of all values comprised in the top quantile (above the quantile 75 value)

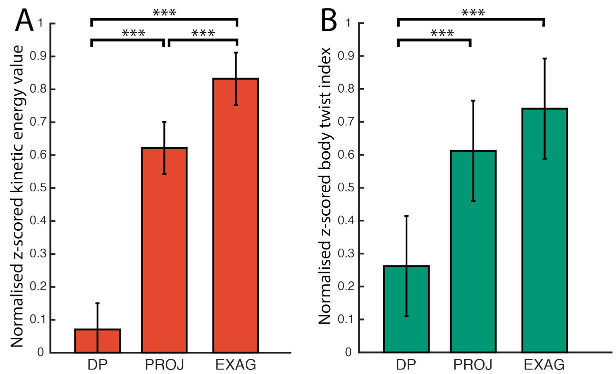

Figure 4: Impact of the different levels of expressiveness on body features (deadpan: DP, projected: PROJ, and exaggerated: EXAG). A, Kinetic Energy; B, Body Twist Index (*: p<0.001).**

with every category of expressive intent, going from deadpan to exaggerated.

When considering the impact of the presence of an audience on the body features, the data showed an interaction effect on the features. When the task was easy, such as during the task of deadpan expressive manners, when musicians were asked to read the music score and play it with no expressive intend ('like a robot'), the presence of a virtual public seemed to increase their musical expressiveness (higher kinetic energy and higher BTI). In the opposite case, when musicians were asked to exaggerate their movements in order to create the most expressive representation of the piece, the task was unusual for musicians. The presence of the audience, in that case, seemed to impact the musicians' performance in the opposite direction by decreasing both the kinetic energy and the BTI. Such context-dependant effects fit Zajonc's theory of social facilitation [24]. This theory states that when people perform simple or very familiar tasks, they tend to do it better in front of others than when alone. Whereas when performing complex or unusual tasks, people tend to perform worse in front of an audience than when alone. The effect of a virtual audience is thus dependant on the difficulty of the task, either easy in the case of deadpan expressive manners or difficult when exaggerating.

An alternative explanation for this phenomenon could be linked to habits. Musicians that are used to play in front of a crowd will affine their movements in a certain way to make them comfortable yet expressive. In this study, the presence of an audience seems to push musicians to use this set of habitual movements. When playing with no expressive intent, the musician could feel the need to still express some emotions for the crowd to enjoy the performance more. In opposition, exaggerating expressive intents in front of a crowd could put musicians in an uncomfortable position where their play seems a bit less authentic. To counterbalance this effect, musicians will then naturally try to be less expressive.

In this study, we computed two motion features using the motion capture data of four different musicians recorded in an Interactive Virtual Environment. While the impact of expressive intent on the kinetic energy is well known in literature, our results highlighted the importance of body shape related information, characterised here by the Body Twist Index. Broad body features such as the kinetic energy and relationships between different body parts were impacted by both the expressive intent and the presence of an audience in a context-dependant manner based on the difficulty of the task.

Our findings suggest that tools such as IVEs can be adapted for training musicians coping with stress related to live performances. The system developed in this study could be easily adapted to context-sensitivity in real-time and could provide feedbacks to the musician. It could help musicians understand the impact of stress on their performance allowing them to develop coping mechanisms for musical performance anxiety.

Future perspectives for this work include combining sound analysis to body features, therefore creating a multi-modal model. We also aim to record how each excerpt is perceived in terms of emotional intensity as well as authenticity. We expect the expressive intent to have a significant effect on the emotional intensity. Finally, the authenticity should be rated lower when the musicians are exaggerating their gestures since the performance appears to be less realistic.

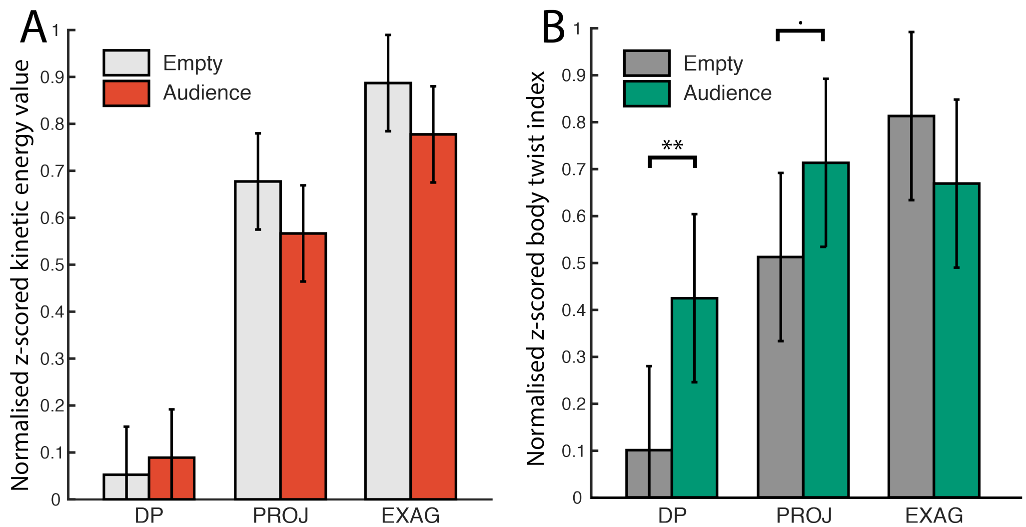

Figure 5: Impact of the interaction of the expressiveness and the presence of an audience on body features (deadpan: DP, projected: PROJ, and exaggerated: EXAG). A, Kinetic Energy; B, Body Twist Index (: p<0.01, . : p<0.1).**

REFERENCES

[1] Josiane Bissonnette, Francis Dubé, Martin D Provencher, and Maria T. Moreno Sala. 2016. Evolution of music performance anxiety and quality of performance during virtual reality exposure training. *Virtual Reality* 20, 1 (mar 2016), 71–81. DOI:http://dx.doi.org/10.1007/s10055-016-0283-y

[2] Jim Blascovich, Jack Loomis, Andrew C Beall, Kimberly R Swinth, Crystal L Hoyt, N Bailenson, and Jeremy N Bailenson. 2012. Psychological Inquiry : An International Journal for the Advancement of Psychological Theory TARGET ARTICLE : Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology Immersive Virtual Environment Technology as a Methodologi. *Psychological Inquiry* 7965, February (2012), 103–124. DOI:http://dx.doi.org/10.1207/S15327965PLI1302

[3] Stéphane Bouchard, Sophie Côté, Julie St-Jacques, Genevì Eve Robillard, and Patrice Renaud. 2006. Effectiveness of virtual reality exposure in the treatment of arachnophobia using 3D games. *Technology and Health Care* 14 (2006), 19–27.

[4] Antonio Camurri, Ingrid Lagerlöf, and Gualtiero Volpe. 2003. Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *International journal of human-computer studies* 59, 1 (2003), 213–225.

[5] Sofia Dahl and Anders Friberg. 2007. Visual Perception of Expressiveness in Musicians' Body Movements. *Music Perception: An Interdisciplinary Journal* 24, 5 (jun 2007), 433–454. DOI:http://dx.doi.org/10.1525/mp.2007.24.5.433

[6] Floriane Dardard, Giorgio Gnecco, and Donald Glowinski. 2016. Automatic classification of leading interactions in a string quartet. *ACM Transactions on Interactive Intelligent Systems* 6, 1 (2016), 1–27. DOI:http://dx.doi.org/10.1145/2818739

[7] Jane W Davidson. 2005. *Bodily communication in musical performance.* Oxford University Press Oxford, Great Clarendon St, Oxford OX2 6DP, UK, 215–237.

[8] JoAnn Difede, Judith Cukor, Nimali Jayasinghe, Ivy Patt, Sharon Jedel, Lisa Spielman, Cezar Giosan, and Hunter G Hoffman. 2007. Virtual reality exposure therapy for the treatment of posttraumatic stress disorder following September 11, 2001. *Journal of Clinical Psychiatry* 68, 11 (2007), 1639.

[9] Donald Glowinski, Naëm Baron, Kanika Shirole, Sélim Yahia Coll, Lina Chaabi, Tamara Ott, Marc-andré Rappaz, and Didier Grandjean. 2015. Evaluating music performance and context-sensitivity with Immersive Virtual Environments. *EAI Endorsed Transactions on Creative Technologies* 2, 2 (2015), 1–10. DOI:http://dx.doi.org/10.4108/ct.2.2e3

[10] Donald Glowinski, Maurizio Mancini, Roddy Cowie, Antonio Camurri, Carlo Chiorri, and Cian Doherty. 2013. The movements made by performers in a skilled quartet: A distinctive pattern, and the function that it serves. *Frontiers in Psychology* 4, NOV (2013), 1–9. DOI:http://dx.doi.org/10.3389/fpsyg.2013.00841

[11] Patrik N Juslin and Petri Laukka. 2003. Communication of emotions in vocal expression and music performance: different channels, same code? *Psychological bulletin* 129, 5 (2003), 770–814. DOI:http://dx.doi.org/10.1037/0033-2909.129.5.770

[12] Marc Leman. 2008. *Embodied music cognition and mediation technology.* Mit Press, 1 Duchess St, Marylebone, London W1W 6AN, UK.

[13] Matei Mancas, R B Madhkour, and Dominique De Beul. 2011. Kinact: a saliency-based social game. *Proceedings of the 7th …* 4, 3 (2011), 65–71. http://www.uninova.pt/

[14] Robert A. McLean, William L Sanders, and Walter W Stroup. 1991. A Unified Approach to Mixed Linear Models. *The American Statistician* 45, 1 (feb 1991), 54. DOI:http://dx.doi.org/10.2307/2685241

[15] Manfred Nusseck and Marcelo M. Wanderley. 2009. Music and Motionfi!?How Music-Related Ancillary Body Movements Contribute to the Experience of Music. *Music Perception: An Interdisciplinary Journal* 26, 4 (apr 2009), 335–353. DOI:http://dx.doi.org/10.1525/mp.2009.26.4.335

[16] Evelyn K Orman. 2004. Effect of virtual reality graded exposure on anxiety levels of performing musicians: a case study. *Journal of music therapy* 41, 1 (2004), 70–78.

[17] Caroline Palmer and Diana Deutsch. 2013. Music performance: Movement and coordination. *The psychology of music* 3 (2013), 405–422.

[18] B O Rothbaum, L Hodges, S Smith, J H Lee, and L Price. 2000. A controlled study of virtual reality exposure therapy for the fear of flying. *Journal of consulting and clinical psychology* 68, 6 (2000), 1020–1026. DOI:http://dx.doi.org/10.1037/0022-006x.68.6.1020

[19] Mamiko Sakata and Sayaka Wakamiya. 2009. An Analysis of Body Movement on Music Expressivity Using Motion Capture. In *Intelligent Information Hiding and Multimedia Signal Processing, 2009. IIH-MSP'09. Fifth International Conference on.* IEEE, IEEE, 445 Hoes Lane Piscataway, NJ 08854-4141 USA, 1172–1176.

[20] Maria V Sanchez-Vives and Mel Slater. 2005. From presence to consciousness through virtual reality. *Nature reviews. Neuroscience* 6, 4 (2005), 332–339. DOI:http://dx.doi.org/10.1038/nrn1651

[21] M. R. Thompson and G. Luck. 2012. Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae* 16, 1 (2012), 19–40. DOI:http://dx.doi.org/10.1177/1029864911423457

[22] Anemone G. W. Van Zijl and Geoff Luck. 2013. Moved through music: The effect of experienced emotions on performers' movement characteristics. *Psychology of Music* 41, 2 (2013), 175–197. DOI:http://dx.doi.org/10.1177/0305735612458334

[23] Aaron Williamon, Lisa Aufegger, and Hubert Eiholzer. 2014. Simulating and stimulating performance: Introducing distributed simulation to enhance musical learning and performance. *Frontiers in Psychology* 5, FEB (2014), 1–9. DOI:http://dx.doi.org/10.3389/fpsyg.2014.00025

[24] Robert B. Zajonc. 1965. Social Facilitation. *Science* 149 (1965), 269–274. DOI:http://dx.doi.org/10.1126/science.149.3681.269