Liveness Through the Lens of Agency and Causality

Florent Berthaut
University of Bristol
florent@hitmuri.net

David Coyle
University College Dublin
d.coyle@ucd.ie

James Moore
Goldsmiths, University of London
j.moore@gold.ac.uk

Hannah Limerick
University of Bristol
Hannah.Limerick@bristol.ac.uk

ABSTRACT

Liveness is a well-known problem with Digital Musical Instruments (DMIs). When used in performances, DMIs provide less visual information than acoustic instruments, preventing the audience from understanding how the musicians influence the music. In this paper, we look at this issue through the lens of causality. More specifically, we investigate the attribution of causality by an external observer to a performer, relying on the theory of apparent mental causation. We suggest that the perceived causality between a performer's gestures and the musical result is central to liveness. We present a framework for assessing attributed causality and agency to a performer, based on a psychological theory which suggests three criteria for inferred causality. These criteria then provide the basis of an experimental study investigating the effect of visual augmentations on audience's inferred causality. The results provide insights on how the visual component of performances with DMIs impacts the audience's causal inferences about the performer. In particular we show that visual augmentations help highlight the influence of the musician when parts of the music are automated, and help clarify complex mappings between gestures and sounds. Finally we discuss the potential wider implications for assessing liveness in the design of new musical interfaces.

Author Keywords

Digital Musical Instrument, Liveness, Agency, Augmentations, Apparent mental causation

ACM Classification

J.5 [Computer Applications] ARTS AND HUMANITIES —Performing arts (e.g., dance, music) H.5.5 [Information Interfaces and Presentation] Sound and Music Computing J.4 [Computer Applications] SOCIAL AND BEHAVIORAL SCIENCES — Psychology

1. INTRODUCTION

Performers using Digital Musical Instruments (DMIs) provide the audience with a considerably different experience to acoustic instruments and this is largely due to the complexity and diversity of DMIs. A key difference between the two lies in the visual information provided to the audience during a performance. The visual component of a performance using an acoustic instrument is rich with information, for example they can convey emotion to the audience and affect the audience's emotional state [4] [13]. DMIs lack such a rich visual information transfer to the audience, largely due to the absence of a direct physical link between the gestures and the sound. DMIs tend to produce more complex sounds than acoustic instruments this increased complexity makes it more difficult to understand how the instruments are used.

This paper investigates liveness from the perspective of the audience's perceived causality between the performer's action and the music they hear. We draw upon the theory of 'apparent mental causation' [14] which was formulated to account for the attribution of causality of one's own conscious thoughts and their actions. This is known as the experience of self-agency - the experience of initiating and controlling one's actions to influence the external environment - but can also be extrapolated and applied to the experience of attributing agency to another agent. We focus on an examination of whether this theory can be used as a valid framework for identifying causality between a performer using a DMI, their actions, and the observers' perception of the performance. We believe that this inference of causality between the performer and their actions is central to the audience's perceptions of liveness during a performance. Our contribution is two-fold. Firstly, we introduce a psychological framework based on inferred causality for the analysis. Secondly, through the empirical application of visual augmentations in line with this framework we gain insights into audience's perception of causality in performers' use of DMIs.

In Section 2, we describe how related work has shown that causality is an important part of liveness, and that visual augmentations can be used to reveal information to the audience. Section 3 explains the psychological research from which we derive our framework. The framework is presented in detail, illustrating how DMIs lack enough visual information for inferred causality. Later, Section 4 describes the visual augmentations designed to provide this missing information. Additionally, the experimental protocol is outlined. Finally we discuss our findings and the opportunities opened by our approach for the NIME community.

2. RELATED WORK

To our knowledge, no previous research has focused on the study of *inferred causality* in the context of the liveness issue with DMIs. However, some research have speculated about its importance in the perception of liveness.

In [12], the authors discuss the importance of causality in the perception of liveness. However, they investigate the evaluation of liveness through the measurement of spectator's emotions, in particular enjoyment. Nonetheless, in their observations from participants answers, several dimensions that have an effect on liveness perception are strongly linked to the perception of causality, such as visibility of gestures, miming effect, interaction fidelity and conventions on instrumental control.

A mixed-reality system providing visual representation of the mechanisms of DMIs aimed at improving liveness was introduced in [1]. The Rouages system displays the connections between gestures and the sound processes with visual links, so that spectators can see the gestures and corresponding sound process of the instrument. This is an example of visual augmentation aimed at mapping the performers' actions to the resultant sound, however was not developed within a theoretic framework of causality. Because of this, the nature of the gestures, such as continuous or discrete are not visualised, which can lead to confusion when a similar gesture leads to changes of different type in the sound. Also, the system does not represent which particular parameter of a sound process individual gestures control, which would enable it to be distinguished from those controlled through an automated process. Furthermore, it does not account for temporal disruption of musicians' actions, e.g. delay or quantization. Instead it only represents controls immediately applied to the sound.

Intentions and prior knowledge regarding DMIs was investigated in [7]. They show that explaining the instrument to the participant with preliminary videos had a positive impact on their enjoyment of the performance. The better participants understood the instrument and the way it was played, the better they were able to appreciate the performance. The authors suggest that, because it is not always possible to explain an instrument beforehand, interaction design should make the instrument and it's interaction more obvious. [9] echoes that familiarity with DMIs positively impacts the audience's experience of a performance. Familiarity and understanding are driven by an awareness of how musicians use DMIs to produce sounds. The audience have a greater experience of liveness when there is a greater exposure to the causality between actions and their musical outcomes. As pointed out in [3], *"... the onus of justification of liveness is shifted to the causal link between the performer's action and the computer's response"*. Acoustic instruments are based on the physical properties of elements composing the instruments and on the mechanical energy transfer between the musicians' movements and the vibrations produced. For example, when observing a drummer hitting a cymbal, the causality between the gesture and sound can easily be inferred because one has seen the arm movement, the collision with the cymbal and heard the sound. These are causal events that humans observe and experience themselves on a daily basis. However, DMIs involve causal events which are far less familiar. This may be due to the lack of visual information provided to the audience about the parameters of sounds and how they are influenced by the performer during the live performance. This lack of visual information also provides the audience with less familiarity with the instrument in order to infer causality between gestures and sound. Therefore, a disengagement occurs between the performer, their actions, the music and the audience.

In this paper, we look at the liveness issue focusing on the perceived causality between gestures and the resultant sound. In particular, we propose a theoretical framework to analyse how this perception of causality is impaired in DMIs. We also describe visual augmentations that specifically aim at visualizing the causality between gestures and musical results in DMIs.

3. INFERRED CAUSALITY

3.1 Attribution of Agency

"*When a thought appears in consciousness just before an action (priority), is consistent with the action (consistency) and is not accompanied by conspicuous alternative causes of the action (exclusivity), we experience conscious will and ascribe authorship to ourselves for the action[15]*"

The quote above summarises the theory of apparent mental causation. According to the current thinking in psychology, inferences regarding the causal link between a conscious thought and action are central to the attribution of self-agency. Crucially, if the inference about the causality between a conscious thought and an action meets three criteria - priority, consistency and exclusivity - the cause of the action will be ascribed to the self. This theory has been shown experimentally, for example, priming people with thoughts corresponding to these three criteria regarding an action that they did not perform, induced a false sense of self-agency for the actions [14]. It is widely thought that the mechanisms supporting the attribution of agency to the self also support the attribution of agency to others [6] [5]. A common system for attributing intentions to the self and others "*enables us to communicate mental states and thereby share our experiences*" [6]. Here, we extrapolate the criteria of self-agency attribution and apply them to investigate the audience's attribution of agency to a performer.

3.2 Liveness, DMIs and criteria for inferred causality

Related work suggests that issues with visual information performance with DMIs adversely affect the perception of liveness. More specifically, a lack of visual information may impair the audience's ability to make inferences about the performer's intentions, actions and musical outcomes, and thus degrade their experience. For instance, contrary to acoustic instruments, the interfaces used in DMIs do not provide any visual information on the nature of the sounds the instrument generates. The same physical interface can be used to control any parameter of any sound process. The gestures can also be subtle or hidden, and therefore not perceived fully by the audience, but at the same time have a strong impact on the musical result. Finally, with changes in scale and nature between the input and output values, the transfer of energy is very different to that which is visually perceived for acoustic instruments. We hypothesise that the impaired visual information with DMIs may cause disruption in one or more of the 3 criteria for causal inference and thus result in a diminished attribution of agency to the performer.

3.2.1 Priority

In Wegner & Wheatley's original theory of apparent mental causation [14], priority refers to the temporal order of events. If a conscious thought occurs before the action, and within a sufficiently close time window, priority criterion is met, which in turn supports an inference of agency. We suggest that this criterion can be adapted and applied to investigate audience's attribution of agency to a performer. This can be done by focusing on the temporal nature of the performer's actions in relation to the musical outcome.

An example of how priority can be impaired in DMI's is the use of quantization, quite common in instruments which rely strongly on a pulse such as those based on pre-existing or live-recorded loops. In order to keep all loops temporally aligned, it is easier to delay their activation or manipulation until the next beat or bar. Thus, although the outcome is occurring after the action, it may be such a long time after the action that the audience cannot infer causality.

3.2.2 Consistency

Consistency [14] refers to the congruence between the content of intention and the related action. That is, if the action is consistent with the prior thought, an inference of self-agency is supported. Here we adapt this criteria for the purpose of investigating attribution of agency to a performer by considering the consistency between actions and the musical outcome.

Musical performances involving DMIs disrupt this criteria, through discrepancies between the nature and scale of the gestures and those of the expected resulting sound events. For example, instantaneous and continuous excitation gestures, as proposed by Cadoz [2], can be easily understood in acoustic instruments, because the musical result follows the nature of the gesture. This principle rarely holds in the case of DMIs. For example, an instantaneous press on a button can trigger a ramp on the pitch of a sinewave. On the other hand, continuous gestures can be used to trigger discrete changes in the sound, such as jumps between meaningful values of a parameter. Differences can also appear between the scale of gestures and the scale of the resulting musical event. For example sometimes turning a knob will result in a subtle change in the timbre of one sound process, equally, the same gesture can apply a strong change such as a low-pass filter on the whole musical output of the instrument. This issue relates to the difference between input and output complexity of DMIs, as described in [11].

3.2.3 Exclusivity

Exclusivity [14] refers to the number of possible causes for an action. Attributions of self-agency are supported in situations where one's intention is the most likely cause of the action. This can be applied to attribution of agency to another in a performance context in several ways. In multiprocess DMIs, several sound processes generate the sound at the same time, some directly from a musician's gestures, other in a completely automated manner. This results in musical events for which the origin is not clear. Exclusivity issues also arise in laptop orchestras, where it can be difficult for the audience to perceive which performer is responsible for the changes in the resulting music. DMIs can also make use of complex mappings. As described in [10], one gesture can be connected to multiple sound parameters, or multiple gestures can be combined to modify a single parameter. Multiple changes in the sound occur while multiple gestures are being made without a clear connection between a gesture and its effect.

4. EXPERIMENT

So far we have discussed the theory of apparent mental causation which defines three criteria for attribution of agency to themselves. We have explained how this theory could be extrapolated and applied to the attribution of agency to others. Here we are focusing on the audience's attribution of agency to a performer. The related work on liveness highlights perceived causality as an important feature in liveness. Furthermore this theory could be used to test liveness in performance with DMIs. This section presents an empirical study where we applied the adapted framework for apparent mental causation to investigate the effect of visual augmentations on liveness in a performance using a DMI.

Participants were presented with a set of videos which showed a live performance using DMIs. This study intends to investigate whether visual augmentations have an impact on attributed agency. For each of the three criteria for inferred agency, a different DMI was designed. Participants would watch a video of each DMI being performed with two visual conditions - with visual augmentations (WithAug) and without visual augmentations (NoAug) (see Figure 1). Participants were asked to make ratings regarding the influence they believed the performer had over the music they heard (agency ratings). Participants were also asked to rate their confidence for their judgements (confidence rating). We hypothesise that for all the criteria, visual augmentations will provide the missing link between required to make causal inference regarding the performer's gestures and thus attribute agency. Therefore, the presence of WithAug will result in participants providing higher ratings about agency of the performer, compared to NoAug. In addition that the confidence ratings will be higher in the WithAug conditions than in the NoAug condition for all criteria. DMIs involve causal processes that are unfamiliar to a typical audience. By providing the audience with visual augmentations highlighting the musical parameters and how they influence the music, they will become more familiar with the causal processes and thus have higher confidence ratings as discussed above in Section 2.

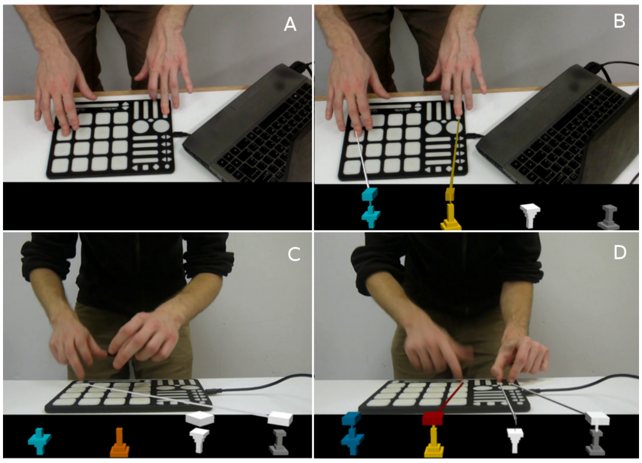

Figure 1: Stills of experiment videos. A) Exclusivity Without Augmentations. B) Exclusivity with Augmentations. C) Priority with Augmentations. D) Consistency with Augmentations.

4.1 Methods and Apparatus

4.1.1 Visual Augmentations

To create the visual augmentations, we first represent sound processes as 3D shapes placed below the DMI, as depicted on Figure 1. We use a set of audiovisual mappings following studies such as [8]. These are: 1) color hue and luminance associated to pitch, 2) shape associated to timbre/sound identity, 3) size associated to loudness, 4) rotation on the Y axis associated to pattern. These mappings make it possible to visually discriminate between the parameters that are being controlled on each sound process and between the sound processes themselves. More importantly, the connections between the gestures and sound processes are represented with specific 3D widgets. One of these widgets is added to each shape representing a sound process, as shown on Figure 1. The widgets are composed of three parts, representing the input, the core and the output of mappings. An input segment is created for each sensor attached to the sound process. It appears when a gesture is sensed and represents the controlled parameter at the sensor value. All the input segments are connected to a core segment that appears when one of them is activated. While the input segments disappear once the gesture is over, this core segment remains visible until the change is applied to the sound. The core segment displays the values of the parameters as they will be set on the sound processes, making changes in scale from the sensor values obvious. The output segment of the widgets goes from the core to the sound process, display the output value and appears only when the value is applied to the parameter of the sound process. Overall, these 3D widgets will therefore compensate for the disruption of the three causality criteria by displaying the connections and parameters between gestures and sound parameters.

4.1.2 DMIs and Videos

For the purpose of the experiment, we designed three DMIs, all with the same interface and sound processes, but which differ in the chosen mappings, so that each of them highlight a disruption of one of the criteria for inferred causality and attribution of agency. We filmed short performances with each DMI from two camera angles front and above, as shown on Figure 1. Two camera angles were chosen here to capture the reality of viewing a musical performance from different viewing perspectives. Therefore there were 3 x 2 x 2 videos in total; criteria x camera angle x augmentation condition.

The QuNeo controller was used as the interface for the DMI because it requires generic gestures, such as instantaneous (hits) and continuous (presses and 1D displacements). On the QuNeo, 16 pressure pads and four touch faders were used. The three DMIs also rely on the same four sound processes, designed to be perceptively separable: One low and one high pitched continuous harmonic sounds composed of sinusoids (respectively between 70Hz and 170Hz, and between 260Hz and 530Hz), with controllable pitch and volume; On low and one high pitched percussive sounds (white noise mixed with sinusoids respectively at 80Hz and 260Hz) with controllable velocity. In addition, four different patterns were composed for each sound process, at a tempo of 120bpm.

For the instrument designed to illustrate the *priority* criteria for inferred causality, our goal was to make it difficult for spectators to determine temporally when a change in the musical pattern has been triggered by the performer's actions. Therefore, the changes in the musical pattern for the sound processes were applied every 8 beats. The second instrument was designed to disrupt the *consistency* criteria for inferred agency, i.e. the spectator's expectations of consistency between action and outcome. We assume that the participants expect discrete gestures, such as striking a pad, to trigger a discrete change, and the same for continuous gestures. Therefore the DMI was designed to match discrete gestures with continuous changes in the music and continuous gestures with discrete changes in the music. The third DMI was designed to disrupt the *exclusivity* criteria by introducing ambiguity over the originating source of the music - either from the live hand gestures of the performer or pre-programmed, automated processes mediated by the computer. This was achieved by the same musical parameters being controlled by both live gestures and automated processes throughout the video. Automated patterns on pitch and volume of the harmonic sounds and the velocity of the percussive sounds were interrupted by the performer using the interface. The presence of the laptop in the video was intended to create the ambiguity between live control and automated processes.

4.1.3 Protocol

The experiment is a repeated measures design where all participants viewed all 12 videos. To control for potential order effect, we counterbalanced the order that videos and criteria were viewed. 17 participants, aged between 18 and 40 years completed the experiment and in return they received a £10 gift voucher.

Participants were asked to sit in front of a 2x1m screen projected onto a white wall. The lighting was kept constant for all participants. We chose a large projection to best re-create a live music performance. A speaker (Behringer MPA40BT-Pro) was used to playback the sound from the performance, again to re-create a live performance.

At the end of each video, two rating boxes appeared on the screen, asking the following questions: 1) *How much was the music influenced by the performer's hand gestures ?* 2) *How confident are you about your judgement* ? Participants then made their ratings for both questions on a scale from 1 to 10. For the agency rating, 1 on this scale represented the lowest rating corresponding to the judgement that the performer had no influence over the music. A rating of 10 on this scale corresponded to the judgement that the music was influenced only by the performer. These ratings were explained to the participant before the study.

4.2 Results

4.2.1 Consistency

A repeated measures 2x2 ANOVA was performed for Camera position X Augmentation condition for both agency ratings and confidence score. There was a significant main effect for the confidence ratings in the augmentation condition $F(1,16) = 4.480$, $p = .05$. A follow up t-test for all the NoAug and WithAug irrespective of camera condition (because camera has no effect on the confidence ratings). WithAug condition has significantly higher confidence ratings than NoAug $t(33) = -2.54$, $p = 0.016$. *Therefore the presence of visual augmentations resulted in a significantly higher confidence rating.*

4.2.2 Exclusivity

A repeated measures 2x2 ANOVA was performed for Camera position X Augmentation condition for both agency ratings and confidence score. There was a significant main effect for the agency ratings in the augmentation condition $F(1,16) = 7.391$, $p = .015$. To further examine the main effect, we performed a follow up t-test for all NoAug and WithAug conditions irrespective of camera position. t-tests revealed that there was a significant difference between agency ratings NoAug and WithAug; $t(33) = -3.19$, $p = 0.003$. *Therefore participants rated agency significantly higher in the WithAug condition than NoAug.*

4.2.3 Priority

A repeated measures 2x2 ANOVA was performed for Camera position X Augmentation condition for both agency ratings and confidence scores however no main effects were found.

4.3 Discussion

The results show that the participants' perception of the performances in our study were significantly altered using visual augmentations in two of the three criteria examined. Participants' perception of the videos of a live performance using the DMI designed to disrupt *consistency* showed a significant improvement of confidence for their agency ratings with the visual augmentations. The videos in the consistency condition were designed to make the gestures inconsistent with participants' expectations regarding actions and their influence over the music. This result suggests that the addition of visual augmentations to represent the causal link between the performer's gestures and the sounds produced, increased participants' confidence in their judgements. This finding in the confidence rating is in line with our hypothesis. Our hypothesis was based on the prior research suggesting that familiarity with the causal processes involved in producing music has a positive effect on audiences' experience of liveness [3]. This may explain why the agency ratings were not altered by the augmentation condition because in reality the influence of the performer over the music was constant for all four videos. We suggest that providing visual information in a context where the outcome of gestures during a performance are inconsistent with standard expectations gives the audience a better understanding of such processes.

The results regarding the *exclusivity* condition showed a significant improvement to the agency ratings with the visual augmentations. This was in line with our hypothesis and suggests that when there are multiple potential causes for the music, the addition of visual information regarding the origin of gestures and their causal influence over the music improves attribution of agency to the performer. Further, that when the amount of influence the musician has over the music is relatively low, visual augmentation assists the audience to distinguish which parts of the music the performer is actually controlling.

There was no difference between the NoAug and WithAug conditions for both confidence ratings and agency ratings in the videos designed for the *priority* criteria. This may be due to an incorrect interpretation of the priority criteria, temporal *order* of events being more central than temporal delay. Temporality is a complex issue in the context of music and a more systematic investigation of this is required.

5. CONCLUSION

In this paper, we used a well-established theory for apparent mental causation and agency to design the visual information provided to the audience which conveyed specific parts of causality between the performer and their actions. By isolating three distinct criteria of inferred causality and then asking participants to rate both their judgements of agency and confidence about these judgements, we are able to better understand the nature and application of visual information during a performance. This framework can help inform the design of new musical interfaces so that they can compensate for known issues in performances In particular we have shown that restoring the perception of causality helps highlighting the influence of the musician in a performance where part of the musical result is automated. We have also shown that causality helps the audience understand the link between musicians' gestures and the resulting changes in the music.

We see two main directions for future research. Firstly, a further investigation into the results presented in this paper and establish thresholds for the point at which the audience's perception changes. For example, further investigating how inconsistent the action and outcome has to be before the audience become less confident about their judgements of agency. Secondly while the framework we have discussed is applied to attributed agency, the psychological research that it relies on focuses on self-agency. A logical follow-up will therefore be to look at the perception of causality by the musicians themselves while they are performing. This framework could then be used to assess both the interaction with and perception of new musical interfaces.

6. ACKNOWLEDGMENTS

This project was partially funded through the Marie Curie FP7 framework (Grant Agreement PIEF-GA-2012-330770).

7. REFERENCES

[1] F. Berthaut, T. Marshall, Mark, S. Subramanian, and M. Hachet. Rouages: Revealing the Mechanisms of Digital Musical Instruments to the Audience. In *Proceedings of NIME*, 2013.

[2] C. Cadoz. *Les nouveaux gestes de la musique*, chapter Musique, geste, technologie, pages 47–92. Éditions Parenthèses, 1999.

[3] J. Croft. Theses on liveness. *Org. Sound*, 12(1):59–66, Apr. 2007.

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

[5] D. C. Dennett. *The intentional stance*. MIT press, 1989.

[6] C. Frith. Attention to action and awareness of other minds. *Consciousness and cognition*, 11(4):481–487, 2002.

[7] A. C. Fyans, M. Gurevich, and P. Stapleton. Examining the spectator experience. In *Proceedings of NIME*, pages 451–454, 2010.

[8] K. Giannakis. A comparative evaluation of auditory-visual mappings for sound visualisation. *Organised Sound*, 11(3):297–307, 2006.

[9] M. Gurevich and A. Cavan Fyans. Digital musical interactions: Performer–system relationships and their perception by spectators. *Organised Sound*, 16(02):166–175, 2011.

[10] A. Hunt and R. Kirk. Mapping strategies for musical performance. *Trends in Gestural Control of Music*, pages 231–258, 2000.

[11] S. Jordà. Interactive music systems for everyone: exploring visual feedback as a way for creating more intuitive, efficient and learnable instruments. In *Proceedings of SMAC*, 2003.

[12] M. Marshall, P. Bennett, M. Fraser, and S. Subramanian. Emotional response as a measure of liveness in new musical instrument performance. In *CHI 2012 Workshop on Exploring HCI Relationship with Liveness*, May 2012.

[13] B. W. Vines, C. L. Krumhansl, M. M. Wanderley, I. M. Dalca, and D. J. Levitin. Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, 118(2):157–170, 2011.

[14] D. Wegner and T. Wheatley. Apparent mental causation: Sources of the experience of will. *American Psychologist*, 54, 1999.

[15] D. M. Wegner. The mind's best trick: how we experience conscious will. *Trends in cognitive sciences*, 7(2):65–69, 2003.