Digital Musical Interactions: Performer–system relationships and their perception by spectators

MICHAEL GUREVICH* and A. CAVAN FYANS[†]

Sonic Arts Research Centre, Queen's University Belfast, Belfast BT7 1NN, UK
E-mail: *m.gurevich@qub.ac.uk; [†]afyans01@qub.ac.uk

This article adopts an ecological view of digital musical interactions, considering first the relationship between performers and digital systems, and then spectators' perception of these interactions. We provide evidence that the relationships between performers and digital music systems are not necessarily instrumental in the same was as they are with acoustic systems, and nor should they always strive to be. Furthermore, we report results of a study indicating that spectators may not perceive such interactions in the same way as performances with acoustic musical instruments. We present implications for the design of digital musical interactions, suggesting that designers should embrace the reality that digital systems are malleable and dynamic, and may engage performers and spectators in different modalities, sometimes simultaneously.

1. INTRODUCTION

For as long as music has been made, it has existed in ecosystems. Sound, by definition, modulates the medium through which it travels, and the nature of the medium influences the way sound propagates. The human sensorimotor system is acutely attuned to the acoustic medium; musicians have been shown to adjust the way they play according to environmental conditions (Ueno, Kato and Kawai 2010). The listener, too, is sensitive not only to the acoustic environment (Truax 1994) but to the social context (DeNora 2000) and intentions of the performer (Nakamura 1987; Davidson 1993; Gabrielsson and Juslin 1996; Dahl and Friberg 2007; Broughton and Stevens 2009). An ecological approach to music performance – one that considers the relationships between people, objects and their environments – should therefore not be new or surprising; yet drastic alteration of the musical landscape precipitated by the rapid pace of technological change over the past hundred years has thrust fundamental questions about musical ecologies to the forefront.

Examples of musical ecologies not encountered in the West before the twentieth century abound: music as ambience (Toop 1995), private listening in public spaces (Chambers 1994) and DJing (Miller 2004), to name a few. Each of these raises its own set of ecological challenges to which entire volumes could be (and have been) devoted. This article therefore focuses on one such phenomenon in particular: music performance with interactive digital systems. Of course, the overall system of any performance is an immensely complex tangle of social, cultural, technical, perceptual, symbolic and emotional relationships, and so we narrow our focus to the particularly knotty set of issues that performances with interactive digital systems pose for the spectator. Our central premise is to show that what we know and assume about spectatorship cannot simply be transplanted from the domain of acoustic performances to that of digitally mediated ones and be expected to hold true.

1.1. Digital musical interactions

The term *digital musical instrument* has recently gained traction to describe a digital system that serves roughly the same role as an acoustic instrument in a traditional performance context (e.g. Miranda and Wanderley 2006): a system that translates human mechanical work into sound energy for the benefit of an audience. However, the embedded term *musical instrument* reflects a widespread view that these systems can function as drop-in replacements for acoustic musical instruments, leaving all other aspects of the performance ecology intact. Indeed, it has been argued that programming languages (Wang and Cook 2004; Blackwell and Collins 2005), laptops (Cascone 2002, 2003; Stuart 2003; Fiebrink, Wang and Cook 2007) and mobile phones (Wang, Essl and Penttinen 2008) can act as musical instruments. This article is not intended to be a critique of terminology, nor do we seek to enter into the debate surrounding whether such devices really are or can be musical. Rather, we embrace the reality that an incredible diversity of praxes have emerged incorporating digital systems in music performance, and examine how these may demand reconsideration of traditional notions of spectatorship.

As such, we use the term *digital musical interactions* (DMIs) to signify the broader scope of our inquiry than that suggested by the more fashionable term with which it shares an acronym. With this new definition, we conceive of DMIs to account for the diverse contexts in which we interact with digital music performance systems, the particular ways in which these engagements differ from those with traditional musical instruments, and the impacts these have on the spectator. Furthermore, the term *interaction* suggests an ecological approach where we focus on the relationships between elements of the system, rather than just the properties of the devices (Gurevich and Treviño 2007).

1.2. DMIs, performers and spectators

In music psychology and emotion research, music performance is frequently framed in terms of communication and cognition, where the spectator receives and decodes messages that are encoded by a performer and/or composer (Gabrielsson 1988; Juslin 1997; Cohen 2005; Poepel 2005). Gurevich and Treviño (2007) challenged the applicability of such a model to DMIs primarily because it doesn't account for experimental, indeterminate, improvisatory, open-form or process-based musics that may feature a deliberate and pronounced absence of expressive intent. Furthermore, as with the term *digital musical instrument*, it imposes anachronistic and unnecessary constraints on the role of the digital system in music performance.

Others, particularly those allied with the field of ecological psychology, have challenged the cognitivist/communication model of music on a more fundamental level, arguing that it neglects the primacy of perception in shaping individual experience. Although the encoding and communication of expressive intent may be a convenient analytic construct, it fails to describe the actual experience of playing music (Ingold 2000: 413). Furthermore, music can take on meaning for a spectator independently of anything encoded in the music itself – due to social context, personal experience or environmental conditions. These factors can directly influence and indeed alter our perception, and thus our experience of music. As Clarke (2005: 8) states, 'the experience of musical meaning is fundamentally – though not exclusively – a perceptual experience'. Even Cross (2005), in a chapter on musical communication, acknowledges that the information-theoretic model is not applicable in a variety of musical contexts and argues for the situatedness of musical experience: that musical meaning varies radically according to the specific context and spectator, and therefore cannot be said to exist independently of individual experience.

In the present article, we adopt an ecological perspective by considering the situated relationships between different elements of the overall system that encompasses a digital musical interaction. We first explore how and why the relationship between the performer and system in a DMI may be fundamentally different from the one between a musician and an acoustic musical instrument. As a consequence, we assert that spectators' perceptual experiences of performative interactions with DMIs are not necessarily equivalent to those with acoustic instruments. Although one may argue that these interactions are *conceptually* equivalent, there is little reason to believe that they will be *perceived* as such. We present outcomes from a recent study that compared spectators' perceptions of performances with two DMIs and conclude by discussing implications for the design of digital performance systems.

2. ARE PERFORMER–DMI INTERACTIONS NECESSARILY 'INSTRUMENTAL'?

In this section we discuss what constitutes an instrumental relationship between a musician and an instrument, and identify qualities of DMIs that may lead to other, non-instrumental interactions. Finally, we examine characteristics of non-instrumental interactions.

2.1. Gesture and instrumentality

Performative interactions with both digital and acoustic instruments are frequently described in terms of *gesture*. The musical significance of this term has been widely discussed recently (Cadoz and Wanderley 2000; Hatten 2004; Gritten and King 2006; Godøy and Leman 2010), but there is general agreement that the two important components of gesture are motion and meaning. Describing the human *input* to a DMI as gestural (e.g. Hunt and Kirk 2000; Wanderley and Depalle 2004) implies that there is something meaningful in the performer's action, potentially for both the performer and spectator. In reality, though, many DMIs rely on generic movements that don't convey specific meaning or purpose – through typing, pointing devices or button-presses, but also actions such as ocular motion (Hornof and Sato 2004) – or even on human energetic phenomena that don't necessarily involve motion at all (e.g. Knapp and Lusted 1990; Miranda and Brouse 2005). Although the sonic outcomes of these actions may invoke meaning for spectators, it is difficult to say that the actions themselves are significant or gestural.

Indeed, Cadoz (2009) argues that the notion of 'instrumentality' has been stretched too far when it comes to DMIs. To him, the important criterion for instrumental interactions is that they represent mechanical relationships through processes that respect consistent energetic exchange. This definition accounts for interactions with digital, synthetic, robotic and even virtual systems; it is the representation of mechanical energy exchange between system elements that is crucial. Many DMIs fall outside this domain, even gestural ones such as those involving video tracking of movement in free space, as well as 'in the case where the gesture is performed on a physical device like a key or a joystick', because 'the transduced signals have no physical (ergotic) meaning' (Cadoz 2009: 227).

2.2. Specificity and skill

Norman (1998) distinguishes 'strong specific' devices – those that are suited for a particular user, task and context – from 'weak general' ones, which are unspecialised and generic. QWERTY keyboards are 'weak general' interfaces in that their associated actions are not differentiable, nor are they distinctive; therefore 'it is difficult to meaningfully relate action to function' (Jensen, Buur and Djajadiningrat 2005: 9). However, this lack of specialisation has become a hallmark of many digital devices, leading to a style of interaction that 'shifts the complexity from the motor actions to the decision process of what to do. It is exactly because button pushing is so simple from a motor point of view that learning is shifted almost completely to the cognitive domain' (Djajadiningrat, Matthews and Stienstra 2007: 659). Interaction with such a 'weak general' device may be as skilled and may take as long a period of practice to develop as with a 'strong specific' one, but the engagement is fundamentally cognitive and lacking in bodily skill. Norman (1993) similarly contrasts devices that encourage *reflective engagement* – the generation of abstractions or ideas – with those that promote *experiential engagement* – interaction in the 'here and now'.

These two kinds of engagement correlate to what are known as *intellectual* and *perceptual-motor* skills. If skill relates to the likelihood of reliably achieving a goal, then intellectual skills are those whose goals are symbolic, and perceptual-motor skills are those where the goals are non-symbolic (Rosenbaum, Carlson and Gilmore 2001). Perceptual-motor skills are distinguished in that their outcomes are more immediate and less easily specifiable than those of intellectual skills. Perceptual-motor skills like riding a bicycle or weaving a basket require continuous, near-instantaneous sensorimotor control, but such activities are notoriously difficult to specify, document, teach or learn symbolically (Ingold 2000; Rosenbaum et al. 2001).

The distinction appears particularly salient in order to characterise different modes of interaction with digital systems. 'Weak general' interfaces typically rely on intellectual skills – they allow the efficient specification of symbolic goals. In Cadoz's (1994, 2009) terms, there is no ergotic meaning to the actions involved in operating weak general devices and therefore they will not tend to facilitate instrumental interactions. The performer does not conceive of the interaction in physical or mechanical terms, but rather in symbolic or conceptual terms. In this sense, moving a slider to adjust the density of a cloud of sound grains, pressing a button to activate a ring modulator or playing a glissando by waving my hand in the air are all fundamentally different from bowing a violin or striking a drum. To reiterate, we do not attach a value judgement to this statement. In fact, our conception of a DMI is purposely inclusive to account for this diversity, as there are a large number of successful, widely adopted DMIs which are non-instrumental (e.g. live coding, button-matrix devices, fader boxes, joysticks or motion-tracking systems).

2.3. Non-instrumental interactions

How, then, can we characterise interactions with DMIs that are not instrumental? Pressing (1990) identified two 'other performance models besides the traditional "person physically manipulating instrument"'. The first resembles one we discussed previously: relying on action that is 'either traditionally considered to be involuntary' or 'doesn't directly involve motion in space'. The second involves shaping 'some external ongoing process or its effects', where 'shaping can be simply turning on and off, filtering, or various types of parametric control' (Pressing 1990: 13). Wessel and Wright (2002: 16) discuss one such interaction, described as 'dipping': using continuous controls to manage the volume or other parameters of autonomous generative processes. Schloss (2003: 241) describes a level above the 'note' and 'timbre' levels at which musicians can control DMIs: a 'macroscopic' or 'process' level, analogous to conducting, where the musician may interact with 'algorithmic processes or sequences'.

In an analysis of interactions with a screen-based 'virtual musical instrument', Johnston, Candy and Edmonds (2008) identified two distinct modes beyond an 'instrumental' one in which musicians interacted with the system. The first, along the same lines as the previous examples, was an 'ornamental mode', where musicians surrendered primary control of sound generation to the system and augmented or 'ornamented' the sounds of the system. In the second 'conversational mode', musicians treated the system as a responsive agent, allowing it to have a stake in determining the direction of the music. That the three modes of interaction observed by Johnston et al. (2008) all emerged from musicians interacting with the same system, in different compositional contexts, highlights that DMIs tend to be more malleable and flexible than traditional instruments, and that the nature of the interaction is situated – the system alone cannot account for the diversity of ways in which musicians will engage with it.

In both 'ornamentation' and 'conversation', the system has a measure of autonomy and the performer's contribution is influenced by the system's behaviour. In some DMIs, the agency of the system is explicit: they may be designed to model or mimic human behaviour (Lewis 2000; Pachet 2003; Weinberg and Driscoll 2006). In others, as in Johnston et al.'s (2008) example, the computer models a physics-based dynamic system. In these interactions the systems exhibit substantial dynamic behaviour at the frequencies of human motion. In ecological terms, we can consider that DMIs with low-frequency dynamics that unfold over time *afford* human interaction at these frequencies. The temporal behaviour of DMIs may thus determine the degree to which they afford non-instrumental interactions. By manipulating slowly varying processes, many DMIs allow users to specify temporally remote outcomes in ways that are not possible with acoustic instruments. Temporal remoteness between action and outcome is a primary differentiator of intellectual from perceptual-motor skill (Rosenbaum et al. 2001); the symbolic goals specified by actions involving intellectual skills are less immediate and lacking in the temporal precision required by actions involving perceptual-motor skills.

Live coding is a prime example of a widely practised DMI that involves symbolic specification of temporally remote events. Live-coding musicians typically create processes and algorithms that unfold over time. The musician generates symbolic code that is interpreted and enacted by the computer, possibly in 'real-time', but only after a series of keystrokes or clicks have been assembled to formulate the intended command. Indeed, far from a perceptual-motor skill, live coding has been described as a 'great intellectual challenge' (Collins, McLean, Rohrhuber and Ward 2003: 322).

From this background emerges a relationship between non-instrumental interactions, the specification of symbolic goals, and the temporal behaviour of DMIs. The human voice is an interesting illustration in that it can function both instrumentally and non-instrumentally. In a non-instrumental mode – conversation – the speaker's goals are symbolically specified (words or ideas), and are temporally remote. The speaker is unconcerned with the precise timing of each sound or utterance, nor does he or she pay much attention to the precise sonic qualities of the voice in this mode. Of course, the entire vocal apparatus is continuously, non-consciously adjusted in order to attain the desired goal, but it is not the sound itself that is the goal; rather, it is the expression of thoughts, ideas or words. Although, in a naïve view, one may say that in the instrumental mode – singing – the goals are also symbolically specified in the form a sequence of pitches, once a skilled singer learns the tune, the notes themselves are not the explicit goal. Much like riding a bicycle, singing is a matter of innumerable, continuous, immediate and precise adjustments to attain a desired outcome whose specification is constantly updated.

We provide this example for several reasons. Like Johnston et al.'s (2008) system, the human voice demonstrates that properties of the technology alone do not determine the nature of the relationship a performer will have with it. A particular technology may give rise to a number of different modes of interaction as a product of the entire performance ecology; there are differing social and environmental conditions in which singing and speaking are appropriate or effective means of expression. The kinds of goals that a performer may specify, which may be temporally immediate or remote, symbolic or not, have an equally significant impact on the nature or modality of the interaction – whether it is instrumental or not.

Furthermore, the human voice affords the ability to readily switch from one mode to another. Digital systems allow for much more drastic changes of mode; they may execute radically different responses to the same input depending on the state of the system. Modal interactions require awareness of the state of the system, and in switching from one mode to another the performer must momentarily disengage and cognitively realign himself or herself towards the interaction. Finally, the example of the human voice supports our proposition that expressiveness cannot be determined solely by the properties of a device or technology, nor are certain interactional modalities necessarily more or less expressive than others. Is speaking more or less expressive than singing?

3. PERCEPTION OF PERFORMER–DMI INTERACTIONS

In the case of acoustic music, several studies have investigated the perception of expressive intent by listeners through auditory information (Nakamura 1987; Clarke 1989; Gabrielsson and Juslin 1996). Davidson (1993) demonstrated that kinematic information provides a sufficient basis for judgements of expressive intent, using the point-light display (PLD) technique (Johansson 1973) as a way of representing *ancillary gestures* – physical motions that are not involved in sound production yet convey meaning in some way (Cadoz and Wanderley 2000; Wanderley, Vines, Middleton, McKay and Hatch 2005). Numerous subsequent studies have similarly shown that visual information – performers' motions – can significantly inform the perception of properties such as emotion, tension, phrasing and intensity (Vines, Krumhansl, Wanderley and Levitin 2006; Dahl and Friberg 2007; Broughton and Stevens 2009; Nusseck and Wanderley 2009). However, in many of these studies there were substantially different responses between visual-only and audio-only conditions. Vines et al. (2006) cite several examples in their experiment where the disparity can be attributed to conflicting cues from video and audio; significantly, these visual cues include qualities such as facial expressions that fall outside the realm of ancillary gestures.

It is clear that neither bodily movement nor sound alone tells the whole story with regard to the perception of music performances. Yet, in focusing on ancillary gestures, what all these studies have in common is that none examines the *technical actions* (Leach 1976) – *effective gestures*, in Wanderley et al.'s (2005) terms – involved in the performer–instrument interaction. There is little discussion in the literature of perception of the performers' embouchures, grips or fingering techniques. It is assumed that the spectator implicitly understands the performer–instrument interaction and is able to focus on the performer's expression *through* the instrument, rather than their relationship *with* it. As studies have shown little difference in perceptual assessments by performers and non-performers of the instrument in question (Nusseck and Wanderley 2009), such an assumption seems to rest on the spectator's embodied understanding of the operation of musical instruments. This may be valid in most cases for acoustic instruments where there is a strong causal, mechanical relationship between the controls of the instrument and the resulting sound – an instrumental relationship. Leach (1976: 9) asserts that what may typically be seen as a mundane technical action can in fact be quite expressive: 'The way I prepare the coffee and the instruments which I use in the process give information about my cultural background'. However, this is only true if the spectator actually knows what coffee is and how it can be made. As we have seen, there are many DMIs that employ valid, non-instrumental relationships, so we should not assume that that spectators will experience the same embodied understanding of the performance.

Consequently we focus first on the matters of what spectators actually perceive and understand of performer–system interactions, before jumping ahead to questions of expression or gestural communication. Through a qualitative study, we show that spectators had difficulty understanding the interaction between a performer and a relatively simple DMI, which impacted upon their perceptions of skill, error and instrumentality. However, we found that participants who perceived performances as involving intellectual skill also had distinct aesthetic experiences, indicating that designers of DMIs need to be considerate of spectators' diverse perceptions of performance, including the perception of cognitive skill.

3.1. Investigating the spectator experience of DMIs

As a basis for this investigation, our research is guided by 'five questions' about interactive sensing systems posed by Bellotti, Back, Edwards, Grinter, Henderson and Lopes (2002), reframed from the spectator's perspective. The five questions become:

1. Address: How does the spectator know that the performer is directing communication to the system?
2. Attention: How does the spectator know that the system is responding to the performer?
3. Action: How does the spectator think the user controls the system?
4. Alignment: How does the spectator know that the system is doing the right thing?
5. Accident: How does the spectator know when the performer or the system has made a mistake?

Most of the answers are self-evident in performances with acoustic instruments; with DMIs, we face a much more diverse landscape. We conducted a study that focused primarily on the third and fifth of these questions: spectators' abilities to understand how performers' actions relate to system responses in a DMI, and to perceive errors in their performances. We expected that, among other factors, spectators would need an understanding of how the performer interacted with the system in order to perceive errors. We hypothesised that both would have bearing on spectators' experiences and assessments of the performances, and on their perceptions of skill.

3.1.1. Study design

Twenty-seven participants were selected and divided into three groups. All participants were individually shown a series of videos, which were followed by a structured interview. All participants were shown two solo performances of approximately two minutes' duration: one with the theremin, the other with a device called the Tilt-Synth, which was designed specifically for this study. Each performance followed a similar compositional structure, consisting of three repetitions of a theme interspersed with contrasting improvisations. Prior to the performances, participants across the three groups were shown different introductory videos that provided varying degrees of additional information regarding the operation of the two devices.

The performers were PhD students in computer music with substantial performance experience, but neither had significant expertise with the devices used in the study. In both performances, the repeated themes clearly followed the same pitch/timbre/rhythm contours, but neither performer was able to reproduce the themes exactly. The improvised sections were sonic or timbral explorations, each with a distinct character, but again without clear pitch or metric structure.

The Tilt-Synth, shown in Figure 1, is a self-contained device built from segments of ABS pipe with a speaker located in one end. The synthesis

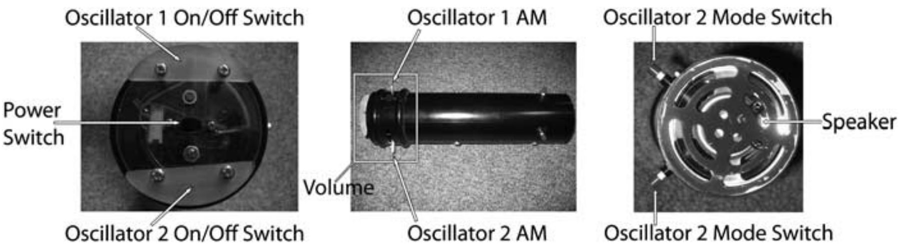

Figure 1. Tilt-Synth.

comes from two PWM outputs of an embedded microcontroller activated by discrete switches on the opposite end. The rotation of one tube segment relative to the other controls the overall volume. A second set of switches toggle between two modes of oscillation. In the pitched mode, a two-axis accelerometer allows the performer to continuously control the pitches of the oscillators through the x–y tilt of the instrument, while two radial sliders control an amplitude modulation effect. In the second mode, tilt controls the bandwidth of a chaotic pitch stream. In this mode the performer has higher-level parametric control of the sound but no access to note-level events.

After the videos had been viewed, the structured interview addressed the following: emotional response, perception of the performer's skill, expectation, mental model of the performer–system interaction, understanding of the performer's intentions, and perception of error. Each interview was videotaped, and transcripts were subsequently analysed following a qualitative data analysis approach informed in part by Grounded Theory (Glaser and Strauss 1967). Some sections were assigned numerical scores by independent coders. See Fyans, Gurevich and Stapleton (2010) for a more detailed description of the study's methodology.

3.2. Findings

We report selected findings that are relevant to the present discussion of the spectator's experience of DMIs, specifically with regard to perceptions of instrumentality, skill and error.

3.2.1. Instrumentality

In describing the performers' interactions with the theremin and the Tilt-Synth, a number of participants used language suggesting they perceived differences in the nature of the devices. By Cadoz's (2009) definition, the theremin does not facilitate strictly instrumental interactions; although we may conceive of a metaphorical one, there is no mechanical relationship, real or virtual, between the performer's motions and the resulting sounds. Nonetheless, significantly more participants ($n = 6$) offered spontaneous positive associations of the theremin to acoustic musical instruments when discussing the performer's skill. Of these, two participants compared the continuous control of pitch to that of a violin. According to one, 'It's not like a piano, where you have a very firm grasp of where your pitches are. It's unguided in a sense, you have to learn where the spots are, kind of like the violin but more difficult.' The other four discussed the coordination required to play both antennae simultaneously in instrumental terms, two of whom made reference to drums: 'It's kind of like rubbing your tummy and patting your head … I guess it's like playing the drums, I imagine its hard to do both things.' In contrast, of the participants ($n = 3$) who referred to musical instruments when discussing the Tilt-Synth, two made negative associations: one suggested it would take far less time to become a virtuoso with the Tilt-Synth than with a violin; another compared its pitch to the imprecision of a honky-tonk piano.

This sense of imprecision or indeterminacy featured prominently in perceptions of the Tilt-Synth. A number of participants expressed that the perceived complexity or the inability to specify definite pitches made the interaction somehow different from that with an acoustic instrument. Many of these described interaction with the Tilt-Synth's *process*, reminiscent of Johnston et al.'s (2008) ornamental mode of interaction. One participant said, 'I think the device is quite complicated in what it can do, and he is playing a lot with the randomness of it … it's not really right or wrong, it's just things happening.' Another suggested that indeterminacy affected its learning curve: 'Since it's generally more noisy, you wouldn't be trying to get discrete pitches out of it so I'd say it's fairly simple to learn … you can just play around with it.' These views reflect a perception of instrumentality with the theremin that was lacking in the Tilt-Synth.

However, this perception wasn't shared among participants who saw the Tilt-Synth as being less autonomous. Eleven participants believed that the number of controls of the Tilt-Synth made it difficult to control and thus to play well. When the device's behaviour was perceived as being under more direct control, the increased complexity relative to the theremin was seen to indicate that the interaction with the Tilt-Synth required greater perceptual-motor skill. When asked to estimate how long the performer had been playing the device, this group's average estimate was significantly greater than that of all other participants – a mean of 634 days versus 266 days. The specific ability to control pitch or individual note duration appeared important. When participants believed that the performer did not control these, they assumed that it was either too difficult to do so, or that it wasn't in the device's nature.

3.2.2. *Perception of error*

It is clear from the preceding comments that a large number of participants did not understand the performer's interaction with the Tilt Synth. This was partly anticipated according to the design of the study (see Fyans et al. 2010). The lack of a strong understanding of the performer–system interaction was associated with an inability to identify errors in the performance. Among all participants, there were only two errors identified in the Tilt-Synth performance versus ten in the theremin performance. (The theremin performer self-identified eight separate errors in his performance; the Tilt-Synth player self-identified five.) Seven of the ten perceived errors in the theremin performance were specified in physical terms, referring to the same instance where the performer's hand struck the pitch antenna. Interestingly, this had no noticeable sonic effect, nor did the demonstration video explicitly mention whether the antennae should be touched or not. Participants were reacting to the physical action itself, indicating an embodied understanding of a correct or incorrect way of interacting with the instrument. The other three perceived errors were specified in musical terms – pitch or rhythmic inaccuracy in the repetition of the theme.

What is most significant regarding error perception is that 25 of 27 participants did not detect any of the Tilt-Synth performer's five self-identified errors. This is not particularly surprising in light of the fact that there was such poor understanding of the interaction overall; a customer who had never seen coffee being made would not be expected to know if a barista were preparing it correctly. Some explicitly attributed their inability to detect errors to this lack of understanding; when asked if he noticed any errors, one participant answered, 'I couldn't say. I don't know the instrument.' But further explanations of *why* participants did not perceive errors were revealing.

Two trends relevant to our discussion emerged in more elaborated responses. The first was a perception that the performance was inherently immune to error. One participant who observed no errors added, 'But I think it's because of the nature of the instrument. I assume it's more improvisatory anyway, that it's not so much a mistake. It's kind of open to the fact that the soundscape could change drastically.' To these participants, it is the affordance of control over a process, of the overall 'soundscape', rather than over immediate events – characteristics of ornamental, non-instrumental interactions – that gives the Tilt-Synth immunity from error. Participants felt that the Tilt-Synth did not afford the kind of accurate control of pitch in particular that was attributed to the theremin; consequently, pitch errors were impossible. Participants also had difficulty associating particular actions with distinct sound events, which appeared to be necessary to perceive error. One participant said that errors were difficult to detect 'because the pitch was less pure, it was less clear'. The same participant went on to indicate that it would require an extreme physical 'slip' (Reason 1990) to commit an error with the Tilt-Synth: 'I think to notice an error with that instrument it would need to be more drastic, you would need to … misplace it or drop it.'

In the second trend, participants discussed relying on physical or gestural cues, or a nonspecific perception of physical 'control' in the absence of an intimate understanding of the interaction. However, there was little agreement on how to interpret such cues. To some, a lack of explicit expressions indicated an absence of errors. One participant said, 'I couldn't really tell from his expression … It seemed like he intended to do everything.' Another noted: 'Looking at the body language there wasn't anything like "oh I goofed!" You see what happens when people are improvising and they screw up.'

Others interpreted the Tilt-Synth performer's physical actions more positively:

> There was something about his quality of playing that he knew what the instrument did. Even though I don't understand fully what the instrument was doing, I could see that his control of the instrument was precise – he wasn't like 'oh what does this do?' You can see that in the person's body when they don't really know what to do with the stuff. I could see he knew what to do with it. There was a physical confidence.

3.2.3. *Perception of skill*

This sense of 'confidence' informed several participants' perceptions of skill as well, but only seemed to be invoked with regard to the Tilt-Synth, where for some there was no other basis to form an assessment. Several participants cited the Tilt-Synth performer's technical knowledge of the device or others like it as another important factor in determining his skill. Five participants spontaneously suggested that the performer had built the device himself (he hadn't) or that it was designed specifically for him (it wasn't). These participants seemed to believe that the performer needs a deep understanding of the operation of the Tilt-Synth in order to develop skill. However, none said anything similar about the theremin.

Interestingly, although many participants had a strong mental model of the gesture–sound relationship of the theremin, very few were able to offer an accurate technical explanation of its operation. But, unlike the Tilt-Synth, none suggested that such an understanding was necessary for their own appreciation of the interaction, nor for the performer to be skilled. This suggests that participants were more attuned to the performer's perceptual-motor relationship with the theremin.

4. DISCUSSION AND IMPLICATIONS FOR DESIGN

Our study shows that many spectators experienced the Tilt-Synth and the theremin differently. Participants focused on the thereminist's perceptual-motor skills or their immediate sonic manifestations. Participants made more direct, positive associations with acoustic musical instruments, and were unconcerned with their own or the performer's intellectual knowledge of the theremin's technology. In contrast, spectators described the perceived indeterminate or process-based nature of the Tilt-Synth as making it 'error-proof'. Unlike with the theremin, participants discussed perceptions of errors and skill with the Tilt-Synth in terms of a technical understanding of the device's operation – either the absence of their own, or the presence of the performer's. Without a mental model of the interaction, some participants were primarily attuned to bodily cues – facial expressions or physical 'confidence'. To be clear, this didn't always appear to negatively impact on spectators' experiences. One participant said, 'I hadn't seen that instrument before. It was kind of hard to understand how it works. But … the movement of it, and the operation of it was very fascinating.'

There is a widespread desire to see DMIs in which performers can develop skill and spectators can appreciate virtuosity (Wessel and Wright 2002; Schloss 2003; Dobrian and Koppelman 2006). Rodger (2010) found that bodily motion is an important basis for the perception of skill with acoustic instruments. Further to this, several authors have indicated that human movement, specifically when it is involved in skilled bodily practice, can be a rich contributor to aesthetic experiences (Jensen et al. 2005; Dobrian and Koppelman 2006; Djajadiningrat et al. 2007). However, as we have seen, not all DMIs demand the kind of perceptual-motor engagement that gives rise to skilled bodily movement in the same way as instrumental interactions. In spite of our best efforts, it seems likely that even if we strive to design DMIs that *are* instrumental, spectators won't always perceive them that way. Digital systems are also sometimes better suited to intellectual or cognitive engagement. Furthermore, non-instrumental performative interaction brings about a new set of intriguing aesthetic possibilities.

But how can we design non-instrumental DMIs with the spectator's perception in mind? We are left with a paradox. Our study showed that spectators' perception is especially attuned to bodily cues and gestures when they don't have access to an understanding of the performer–system relationship or where the musical context doesn't provide a clear sense of expectation; this is more likely in the case of non-instrumental interactions. Yet the attributes spectators identified – facial gestures, physical 'confidence' – are not in any way integral to the interaction. It appears that, for some spectators, all a performer needs to do to be an expert is to not *look* like they are making a mistake; all they need to do to be engaging is to make exaggerated motions, whether they are pertinent to the interaction or not.

Although some spectators appreciated the Tilt-Synth player's confidence and movement, others indicated that they needed more of a basis to understand, and thus to assess, the performance. The question then becomes one of designing for the spectator's perception of cognitive or intellectual performance skills: can we design skilled, non-instrumental DMIs where such information is more readily available to spectators and provides a basis for engaging with the performance?

One interesting example where performers have implicitly acknowledged this need is in live coding. Some musicians advocate projecting the performer's computer display for the audience (Collins 2003: 69), although this practice would seem to require some prior knowledge of the programming environment in order for the spectator to benefit. We have frequently observed instances where live coders not only project their screen images, but also type comments that provide a running 'meta-commentary' on what they are doing. This initiates a form of direct communication with spectators that may help attune their perception – an intellectual engagement that may affect the perceptual experience. Yet the lack of bodily motion in live coding provides no fallback for spectators who are still in the dark.

Our study also provided evidence that making the interaction understood is not the only way for a spectator to appreciate it. Non-instrumental relationships can engender other forms of engagement for spectators. Stuart (2003: 60) suggests that spectators can move beyond perception; that 'the audience has become caught up in the visual spectacle and physical gesture of musical performance', and 'needs to make a shift in their understanding of performance … from a visual focus to that of aural performativity'. One of our study participants seemed to go through this process over the course of the two-minute Tilt-Synth performance:

> I think it was the actual sonic material that I enjoyed rather than his actual performance of it. There's a certain novelty that an electronic instrument that's unpredictable

> has – its difficult for me, it always makes this disconnect when you almost make it do [something], and you almost try and force it to do something and it doesn't cooperate. I sort of have to make a distinction: that you really don't know fully what's going to happen, so I have to step back and appreciate the sonic material for what it is, instead of 'I'm here to *watch* a performance'.

Yet physicality and visual perception provided still a different level of appreciation for some participants, even when they didn't see the interaction as instrumental. As one participant said, 'It was interesting to try and figure out what he was doing with [the Tilt-Synth]. The theremin – you know what a theremin does. This instrument seemed more interesting to me.'

It appears that one of the hallmarks of DMIs is the ability to engage spectators and performers in different ways. Johnston et al.'s (2008) study showed that a particular system could afford multiple modes of interaction; our study showed that spectators of a DMI perceived different modes of interaction, sometimes simultaneously. We suggest that, more than anything, this is what designers of DMIs need to account for. There is no universal experience, not for performers, nor for spectators. Instrumentality is best seen as a continuum; the modalities of engagement – in terms of immediacy of action-outcomes, symbolic or non-symbolic goals, and perceptual-motor or intellectual skills – may be dynamic, nonlinear, overlapping and inconsistent from one performer to the next. Spectators may simultaneously engage with a performance in terms of technical action, bodily movement, facial expressions, soundscape and environmental conditions. Locking a digital musical interaction and its perception by spectators into a single modality of experience may be an unattainable goal. Successful DMIs are more likely to be those that account for this diversity and capitalise on this flexibility that digital devices afford.

REFERENCES

Bellotti, V., Back, M., Edwards, W.K., Grinter, R.E., Henderson, A. and Lopes, C. 2002. Making Sense of Sensing Systems: Five Questions for Designers and Researchers. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press, 415–22.

Blackwell, A. and Collins, N. 2005. The Programming Language as a Musical Instrument. In P. Romero, J. Good, E. Acosta Chaparro and S. Bryant (eds.) *Proceedings of the 17th Workshop of tohe Psychology of Programming Interest Group*, Sussex University: 120–30. http://www.ppig.org/papers/17th-blackwell.pdf.

Broughton, M. and Stevens, C. 2009. 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): 137–53.

Cadoz, C. 1994. Le geste canal de communication homme/machine: La communication 'instrumentale'. *Technique et Science Informatiques* **13**(1): 31–61.

Cadoz, C. 2009. Supra-Instrumental Interactions and Gestures. *Journal of New Music Research* **38**(3): 215–30.

Cadoz, C. and Wanderley, M.M. 2000. Gesture-Music. In M.M. Wanderley and M. Battier (eds.) *Trends in Gestural Control of Music*. Paris: IRCAM – Centre Pompidou.

Cascone, K. 2002. Laptop Music – Counterfeiting Aura in the Age of Infinite Reproduction. *Parachute* **107**: 52–60.

Cascone, K. 2003. Grain, Sequence, System: Three Levels of Reception in the Performance of Laptop Music. *Contemporary Music Review* **22**(4): 101–4.

Chambers, I. 1994. The Aural Walk. In C. Cox and D. Warner (eds.) *Audio Culture: Readings on Modern Music*. New York: Continuum, 2004.

Clarke, E.F. 1989. The Perception of Expressive Timing in Music. *Psychological Research* **51**(1): 2–9.

Clarke, E.F. 2005. *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. Oxford: Oxford University Press.

Cohen, A.J. 2005. Music Cognition: Defining Constraints on Musical Communication. In D. Miell, R.A.R. MacDonald and D.J. Hargreaves (eds.) *Musical Communication*. Oxford: Oxford University Press.

Collins, N. 2003. Generative Music and Laptop Performance. *Contemporary Music Review* **22**(4): 67–79.

Collins, N., McLean, A., Rohrhuber, J. and Ward, A. 2003. Live Coding in Laptop Performance. *Organised Sound* **8**(3): 321–30.

Cross, I. 2005. Music and Meaning, Ambiguity and Evolution. In D. Miell, R.A.R. MacDonald and D.J. Hargreaves (eds.) *Musical Communication*. Oxford: Oxford University Press.

Dahl, S. and Friberg, A. 2007. Visual Perception of Expressiveness in Musicians' Body Movements. *Music Perception* **24**(5): 433–54.

Davidson, J.W. 1993. Visual Perception of Performance Manner in the Movements of Solo Musicians. *Psychology of Music* **21**(2): 103–13.

DeNora, T. 2000. *Music in Everyday Life*. Cambridge: Cambridge University Press.

Djajadiningrat, T., Matthews, B. and Stienstra, M. 2007. Easy Doesn't Do It: Skill and Expression in Tangible Aesthetics. *Personal and Ubiquitous Computing* **11**(8): 657–76.

Dobrian, C. and Koppelman, D. 2006. The 'E' in NIME: Musical Expression with New Computer Interfaces. *Proceedings of the 2006 Conference on New Interfaces for Musical Expression*. Paris: IRCAM – Centre Pompidou, 277–82.

Fiebrink, R., Wang, G. and Cook, P.R. 2007. Don't Forget the Laptop: Using Native Input Capabilities for Expressive Musical Control. *Proceedings of the 2007 Conference on New Interfaces for Musical Expression*. New York: ACM Press, 164–7.

Fyans, A.C., Gurevich, M. and Stapleton, P. 2010. Examining the Spectator Experience. *Proceedings of the 2010 Conference on New Interfaces for Musical Expression*. Sydney, Australia: 451–4.

Gabrielsson, A. 1988. Timing in Music Performance and its Relations to Music Experience. In J.A. Sloboda (ed.) *Generative Processes in Music*. Oxford: Clarendon Press.

Gabrielsson, A. and Juslin, P.N. 1996. Emotional Expression in Music Performance: Between the Performer's Intention and the Listener's Experience. *Psychology of Music* **24**(1): 68–91.

Glaser, B.G. and Strauss, A.L. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. London: Weidenfeld & Nicolson.

Godøy, R.I. and Leman, M. (eds.) 2010. *Musical Gestures: Sound, Movement, and Meaning*. New York: Routledge.

Gritten, A. and King, E. 2006. *Music and Gesture*. Aldershot: Ashgate.

Gurevich, M. and Treviño, J. 2007. Expression and its Discontents: Toward an Ecology of Musical Creation. *Proceedings of the 7th International Conference on New Interfaces for Musical Expression*. New York: ACM Press, 106–11.

Hatten, R.S. 2004. *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert*. Bloomington and Indianapolis: Indiana University Press.

Hornof, A. and Sato, L. 2004. EyeMusic: Making Music with the Eyes. *Proceedings of the 2004 Conference on New Interfaces for Musical Expression*, Hamamatsu, Japan: 3–5.

Hunt, A. and Kirk, R. 2000. Mapping Strategies for Musical Performance. In M.M. Wanderley and M. Battier (eds.) *Trends in Gestural Control of Music*. Paris: IRCAM – Centre Pompidou.

Ingold, T. 2000. *The Perception of the Environment: Essays on Livelihood, Dwelling and Skill*. London: Routledge.

Jensen, M.V., Buur, J. and Djajadiningrat, T. 2005. Designing the User Actions in Tangible Interaction. *Proceedings of the 4th Decennial Conference on Critical Computing: Between Sense and Sensibility*. New York: ACM Press, 9–18.

Johansson, G. 1973. Visual Perception of Biological Motion and a Model for its Analysis. *Perception & Psychophysics* **14**(2): 201–11.

Johnston, A., Candy, L. and Edmonds, E. 2008. Designing and Evaluating Virtual Musical Instruments: Facilitating Conversational User Interaction. *Design Studies* **29**(6): 556–71.

Juslin, P.N. 1997. Emotional Communication in Music Performance: A Functionalist Perspective and Some Data. *Music Perception* **14**(4): 383–418.

Knapp, R.B. and Lusted, H.S. 1990. A Bioelectric Controller for Computer Music Applications. *Computer Music Journal* **14**(1): 42–7.

Leach, E.R. 1976. *Culture and Communication: The Logic by Which Symbols are Connected*. Cambridge: Cambridge University Press.

Lewis, G.E. 2000. Too Many Notes: Computers, Complexity and Culture in *Voyager*. *Leonardo Music Journal* **10**: 33–39.

Miller, P.D. 2004. *Rhythm Science*. Cambridge, MA: The MIT Press.

Miranda, E.R. and Brouse, A. 2005. Interfacing the Brain Directly with Musical Systems: On Developing Systems for Making Music with Brain Signals. *Leonardo* **38**(4): 331–6.

Miranda, E.R. and Wanderley, M.M. 2006. *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*. Middleton, WI: A-R Editions.

Nakamura, T. 1987. The Communication of Dynamics Between Musicians and Listeners Through Musical Performance. *Perception & Psychophysics* **41**(6): 525–33.

Norman, D.A. 1993. *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*. Cambridge, MA: Perseus Books.

Norman, D.A. 1998. *The Invisible Computer: Why Good Products Can Fail, the Personal Computer is so Complex, and Information Appliances are the Solution*. Cambridge, MA: The MIT Press.

Nusseck, M. and Wanderley, M.M. 2009. Music and Motion-How Music-Related Ancillary Body Movements Contribute to the Experience of Music. *Music Perception* **26**(4): 335–53.

Pachet, F. 2003. The Continuator: Musical Interaction with Style. *Journal of New Music Research* **32**(3): 333–41.

Poepel, C. 2005. On Interface Expressivity: A Player-Based Study. *Proceedings of the 2005 Conference on New Interfaces for Musical Expression*, Vancouver, Canada: 228–31.

Pressing, J. 1990. Cybernetic Issues in Interactive Performance Systems. *Computer Music Journal* **14**(1): 12–25.

Reason, J. 1990. *Human Error*. Cambridge: Cambridge University Press.

Rodger, M. 2010. *Musicians' Body Movements in Musical Skill Acquisition*. PhD thesis, Queen's University Belfast.

Rosenbaum, D.A., Carlson, R.A. and Gilmore, R.O. 2001. Acquisition of Intellectual and Perceptual-Motor Skills. *Annual Review of Psychology* **52**(1): 453–70.

Schloss, W.A. 2003. Using Contemporary Technology in Live Performance: The Dilemma of the Performer. *Journal of New Music Research* **32**(3): 239–42.

Stuart, C. 2003. The Object of Performance: Aural Performativity in Contemporary Laptop Music. *Contemporary Music Review* **22**(4): 59–65.

Toop, D. 1995. *Ocean of Sound: Aether Talk, Ambient Sound and Imaginary Worlds*. London: Serpent's Tail.

Truax, B. 1994. The Inner and Outer Complexity of Music. *Perspectives of New Music* **32**(1): 176–93.

Ueno, K., Kato, K. and Kawai, K. 2010. Effect of Room Acoustics on Musicians' Performance. Part I: Experimental Investigation with a Conceptual Model. *Acta Acustica united with Acustica* **96**(3): 505–15.

Vines, B.W., Krumhansl, C.L., Wanderley, M.M. and Levitin, D.J. 2006. Cross-Modal Interactions in the Perception of Musical Performance. *Cognition* **101**(1): 80–113.

Wanderley, M.M. and Depalle, P. 2004. Gestural Control of Sound Synthesis. *Proceedings of the IEEE* **92**(4): 632–44.

Wanderley, M.M., Vines, B.W., Middleton, N., McKay, C. and Hatch, W. 2005. The Musical Significance of Clarinetists … Ancillary Gestures: An Exploration of the Field. *Journal of New Music Research* **34**(1): 97–113.

Wang, G. and Cook, P.R. 2004. On-The-Fly Programming: Using Code as an Expressive Musical Instrument. *Proceedings of the 2004 Conference on New Interfaces for Musical Expression*, Hamamatsu, Japan, 138–43.

Wang, G., Essl, G. and Penttinen, H. 2008. Do Mobile Phones Dream of Electric Orchestras? *Proceedings of the 2008 International Computer Music Conference*. Belfast/San Francisco: ICMA.

Weinberg, G. and Driscoll, S. 2006. Toward Robotic Musicianship. *Computer Music Journal* **30**(4): 28–45.

Wessel, D. and Wright, M. 2002. Problems and Prospects for Intimate Musical Control of Computers. *Computer Music Journal* **26**(3): 11–22.