Was that me? Exploring the Effects of Error in Gestural Digital Musical Instruments

Dom Brown
Holosphere VR
Birmingham, UK
dom@dom-brown.co.uk

Chris Nash
CT Lab, University of the West of England, Bristol, UK
Chris.Nash@uwe.ac.uk

Thomas J. Mitchell
CT Lab, University of the West of England, Bristol, UK
tom.mitchell@uwe.ac.uk

ABSTRACT

Traditional Western musical instruments have evolved to be robust and predictable, responding consistently to the same player actions with the same musical response. Consequently, errors occurring in a performance scenario are typically attributed to the performer and thus a hallmark of musical accomplishment is a flawless musical rendition. Digital musical instruments often increase the potential for a second type of error as a result of technological failure within one or more components of the instrument. Gestural instruments using machine learning can be particularly susceptible to these types of error as recognition accuracy often falls short of 100%, making errors a familiar feature of gestural music performances. In this paper we refer to these technology-related errors as *system errors*, which can be difficult for players and audiences to disambiguate from *performer errors*. We conduct a pilot study in which participants repeat a note selection task in the presence of simulated system errors. The results suggest that, for the gestural music system under study, controlled increases in system error correspond to an increase in the occurrence and severity of performer error. Furthermore, we find the system errors reduce a performer's sense of control and result in the instrument being perceived as less accurate and less responsive.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**; **Sound-based input / output**; Auditory feedback; Activity centered design.

KEYWORDS

virtual reality, augmented reality, sonification, game audio, spatial audio, sonic interaction design, musicology, sound art

ACM Reference Format:
Dom Brown, Chris Nash, and Thomas J. Mitchell. 2020. Was that me? Exploring the Effects of Error in Gestural Digital Musical Instruments.

1 INTRODUCTION

An indicator of musical accomplishment and virtuosity in Western classical music is the ability to perform a faithful rendition of complex musical scores with fluidity and expression. The development of musical skill is sometimes viewed as a journey on which players learn from their mistakes, and failure is recognised as the "consequence of striving to succeed" [26]. Gurevich et al describe success as "the inverse of error" [25] and while the assessment of musical skill is certainly complex and subjective, an important factor is the frequency and severity of performer-related errors [1, 13].

In performances featuring digital musical instruments (or DMIs), the evaluation of skill can be further complicated, as the connection between performer action and auditory response is not always obvious [15, 40]. This unfamiliarity regarding cause and effect is often unavoidable, especially when DMIs introduce new affordances and performance technique. Additionally, new instruments and interfaces are often handmade prototypes and, compared with commercially engineered instruments, their designs are typically less refined and their performers are understandably less practised [12]. Consequently, it can be challenging for spectators to recognise a performer's accomplishments [20] and it may not be obvious when mistakes have been made [4].

Live music performances with new and experimental instrument prototypes are often disrupted by errors, particularly when gestures and body movement are the primary control modality [5]. Despite the widespread impact of these problems, and the clear importance of error and failure in music practice more widely, the examination of error in DMI performance is still a relatively unexplored area of research. In previous work, the spectator experience has been studied, demonstrating that audience perception of performance error does not correspond to a reduction in enjoyment [4]. O'Modhrain has previously argued that "the most important stake-holder in the process of designing and building a DMI is the performer," suggesting that DMI evaluation should also focus on the performer's perception of their instrument. Consequently, this paper begins to explore the effect of error on a player's perception of, and ability to perform with DMIs. In particular, the focus here is on technology-related errors that cause an instrument to incorrectly interpret a player's actions, i.e., when a player precisely expresses their intention, but the instrument fails to respond correctly. This problem is particularly prevalent in gestural music systems, using motion capture and machine learning, which can introduce an unavoidable degree of indeterminacy. We refer to these technology-related errors as *system errors* and present the results of a pilot study that has been designed to reveal the impact of these types of errors on a player's perception of, and ability to perform with, a dataglove-based gestural DMI. In Section 2 we review the literature relating to musical indeterminacy, error and failure, before presenting the system error study and results in Section 3. In Section 4 we conclude with a summary of findings identifying potential areas for future work.

2 PERFORMER AND SYSTEM ERROR

In the Western classical tradition, performers are often framed as transparent mediators of composer intent, with little tolerance for error [23]. Consequently, traditional acoustic musical instruments have evolved to be robust and predictable, responding to the same player actions with a consistent musical response. With the exception of mechanical errors such as broken strings and reeds, unintended errors are typically attributed to the player. We refer to these types of performer-related errors as *performer errors.*

In music performance, errors are certainly not always regarded negatively, rather they are embraced as an inevitable consequence of exploratory performance technique that can open creative opportunities for musical progress [5]. As Scott Adams famously stated "creativity is allowing yourself to make mistakes, art is knowing which ones to keep" [2]. Similarly, indeterminacy and chance also play an important role in contemporary music, for example, Cage used an *abstract notation* in *Concert for Piano and Orchestra* to provide space for performer improvisation and Xenakis used stochastic processes to compose *Pithoprakta.* Notions of indeterminacy have also been a longstanding feature of performances with DMIs; for example, Chadabe's *Echos* [9] and Dahlstedt's *Pencil Fields* [10] both incorporate elements of indeterminacy, where unexpected changes are viewed as creative input from the instrument. Hazzard et al have recently studied the aesthetic opportunity of errors and contributed a set of response strategies for live musical performance [26].

In these examples, indeterminacy and error are intentional and measured; however, performances with experimental, prototype DMIs are frequently disrupted by unintended errors owing to some form of technical failure within the instrument, or what we refer to as *system errors.* These prototype instruments are often hand made and rarely subjected to rigorous testing; consequently, the potential for system errors can be high, and may be traced to a multitude of sources including:

(1) mechanical failure of tactile components or enclosures
(2) electrical failure linked with sensors/batteries/soldering/noise
(3) firmware or software bugs
(4) communication problems causing latency/interference

Machine Learning (ML) techniques are becoming increasingly used in DMI and music interaction design [29, 41, 44], particularly in gestural interaction systems, which are noted for their susceptibility to error [3, 5, 37]. These errors are a consequence of the challenges recognition systems face when distinguishing between different classes of similar gestures [35]. ML is valuable in this context because it allows researchers to move away from explicitly defining the relationships between multivariate sensor readings and the human actions that cause them; instead, these connections can be "taught by example" [19, 30, 34]. A number of libraries and applications have been developed to enable this approach including the Gesture Recognition Toolkit [21] and Wekinator [16]. In the context of gestural music interaction, ML enables end-users to easily manipulate sound using unique and complex freehand movements and gestures. While this approach enables gestures to be abstracted away from their corresponding data streams, it can make precise control in performance scenarios challenging [18]. ML algorithms are not by themselves a source of indeterminacy but they are susceptible to misclassification errors and unpredictable predictions as a result of noise in the gesture capture system and/or limitations of the training data. Consequently, recognition accuracy always falls short of 100% [42]. While this imprecision and unpredictability can be embraced as a means of creative expression [17], unintended system errors can become an inevitable feature of live performance. In our previous work, misclassification errors have been shown to be a hindrance to both novice and expert users of gestural music systems [6].

What are the effects of system errors? Do they make instruments harder to play and do they increase the likelihood of performer errors? And how do system errors affect a performer's perception of their instrument? In the following sections we present a study that begins to explore the impact of system error when using a dataglove-based gestural music interaction system. Participants are asked to repeatedly practise a musical task as deliberate system errors are introduced that are designed to simulate system errors that can occur in gestural DMIs. Three metrics that are linked with musical skill are measured: movement smoothness [22, 24, 39], timing error [13, 14] and note selection error [13]. In addition, the participants' perceptions of the instrument are also captured in terms of accuracy, responsiveness and control. It is hypothesised that the presence of system error will negatively impact the participants' perception of their instrument and increase the prevalence of performer errors.

3 SYSTEM ERROR PILOT STUDY

We conducted a within-group pilot study with five participants to explore the effects of system error, 2 male, 3 female, aged 22–35 (M = 29, SD = 4.6). All participants were right handed and were assessed for musical ability using the Gold-MSI metric [36] with general sophistication recorded in the range 57–76 (M = 66.8, SD = 6.24). Participants were paid a nominal fee of £10 and asked to complete a musical task using a dataglove-based gestural DMI that has been described in previous work [28, 31, 33]. The system is shown in figure 2 and comprises an inertial measurement unit (on the wrist) combined with eight flex sensors (one on the thumb and little finger, two on the proximal and distal joints of the remaining fingers). Participants had no prior experience using the system and for each task they were asked to perform 20 repetitions of a short piece of music (Figure 1) in time to a 60 BPM metronome. For the first four repetitions, the reference melody was played alongside the metronome. The task was repeated six times with intentional system error introduced at 0%, 1%, 2%, 5%, 10% and 20% respectively. To mitigate carryover effects, the order in which participants were exposed to the different error rates was randomised. Errors were introduced when participants selected to play a note, where, for example, at 0% error, no deliberate errors were introduced and at 20%, one in five notes was played incorrectly by choosing a different note at random from the score (Figure 1).

Figure 1: The musical score for the task showing the corresponding direction for each note

3.1 Mapping Strategy

In DMIs, the permutations of how a player's actions generate a musical response are limited by the physical affordances of a system's interface and the constraints of the underlying audio processes.

While some mid-air performance systems have drawn from existing gestural disciplines such as Soundpainting [43], typically connections between action and sound are either defined by the system developer [27] or by end-users [7, 11]. With no established convention for connecting gestures to music, mappings were chosen to satisfy the following criteria:

(1) mappings should be drawn from the practice of existing gestural music performers
(2) mappings should be chosen that minimise the occurrence of genuine system errors to avoid the conflation of real and simulated errors

In this study, participants were asked to complete a simple note selection task and the chosen mapping strategy was based upon a simple "point and grab" metaphor that was designed, and is widely used by practising gestural music performers [6]. Note on and off events are triggered by performing a grab (fist) posture within one of five directional lobes: up, down, left, right and forwards, with each direction mapped to a note in the task's musical score (see Figure 1). Directional lobes were defined using the *Segmented Orientation* method defined in [33].

To further reduce the occurrence of genuine system errors, the machine learning components of the gestural system, used for posture recognition, were disabled. Instead, the participant's grab posture was engaged when a threshold on the middle finger proximal flex sensor was exceeded. Additionally, the system was connected to the logging equipment over a wired connection, to mitigate errors associated with wireless interference [32]. One glove was used, with all participants completing the task with their preferred right-hand.

3.2 Experimental Setup

Figure 2 shows the experimental setup and Figure 3 shows the glove and motion capture markers, which were placed on the participant's back, lower arm and upper arm. The task and glove mapping strategy were explained to the participants, and they were then given as long as they needed to familiarise themselves with the system, which amounted to a few minutes at most. Participants were provided with a score, denoting the order of the point-and-grab directions required to play the piece. A pair of Genelec 8030CL monitors provided auditory feedback, with the notes played by the participant panned to the left monitor and the metronome and reference track panned to the right.

3.3 Collected Data

Movement data were recorded using a Vicon T40s motion capture system along with all data streams from the accelerometers and flex sensors on the dataglove.

Music-related data was collected in MIDI format and included the note on the score, the note selected by the performer and the note emitted by the system. This enabled the disambiguation of performer errors from system errors. For example, if the musician intended to play C3, but the system introduced an error and played D3, both the performer's intended note and the system's erroneous note were recorded. These data were logged alongside a reference performance, which represented a precisely quantised rendition of the piece.

Timestamp	Label	MIDI output
0	REFERENCE	Note on C3
-132	USER_PERFORMED	Note on C3
-132	SYSTEM_CORRECT	Note on C3
1000	REFERENCE	Note on D3
1011	USER_PERFORMED	Note on E3*
1011	SYSTEM_CORRECT	Note on E3
2000	REFERENCE	Note on E3
1958	USER_PERFORMED	Note on E3
1958	SYSTEM_CORRECT	Note on E3
3000	REFERENCE	Note on C4
2991	USER_PERFORMED	Note on C4
2991	SYSTEM_ERROR	Note on G3**
4000	REFERENCE	Note on D3
4053	USER_PERFORMED	Note on G3*
4053	SYSTEM_CORRECT	Note on G3
5000	REFERENCE	Note on E3
5016	USER_PERFORMED	Note on E3
5016	SYSTEM_CORRECT	Note on E3
6000	REFERENCE	Note on G3
5919	USER_PERFORMED	Note on G3
5919	SYSTEM_CORRECT	Note on G3
7000	REFERENCE	Note on C3
7101	USER_PERFORMED	Note on E3**
7101	SYSTEM_ERROR	Note on D3*

Table 1: Example of one trial of MIDI data. *Examples of performer error. **Examples of system error.

In addition to capturing quantitative movement and performance data at each error level, qualitative measurements around the participant's subjective experience were recorded after each test condition. This survey involved a series of Likert scale questions concerning perceived responsiveness, control and accuracy:

(1) On a scale from 0 − 10, where 0 is not responsive at all and 10 is completely responsive, how responsive would you say the gloves were to your actions?
(2) On a scale from 0 − 10, where 0 is no control at all and 10 is complete control, how much control did you feel you had?
(3) On a scale from 0 − 10, where 0 is not at all accurate and 10 is completely accurate, how accurately did you feel the gloves were responding to your actions?

3.4 Movement Smoothness

Movement smoothness was calculated using the method described by Caramiaux [8]: an integrated squared jerk value was measured along each axis of movement for the inter-onset intervals between performed note onsets. For each trial, an overall movement smoothness score was recorded as the cumulative jerk values for each inter-onset interval (i.e. the jerk value between each note). Higher values represent a larger jerk-cost, and thus less-smooth movement.

To allow meaningful averages to be calculated for each condition, median normalisation was applied to each participant's movement smoothness scores.

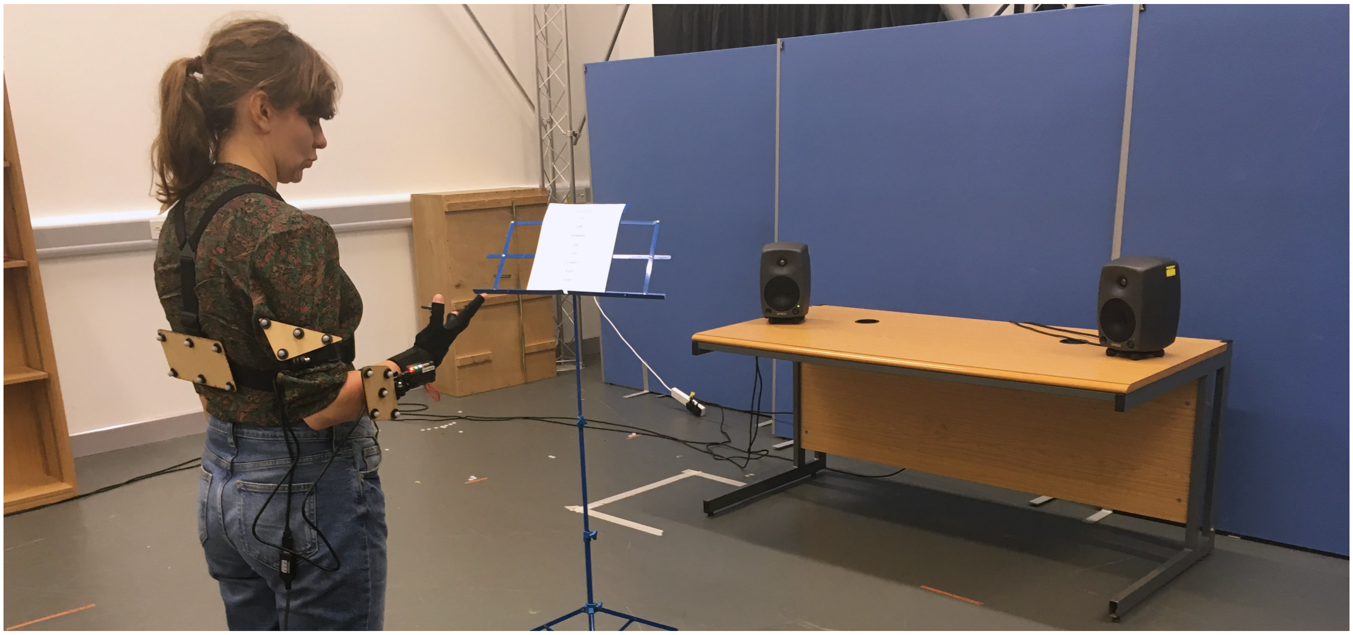

Figure 2: Experimental Setup

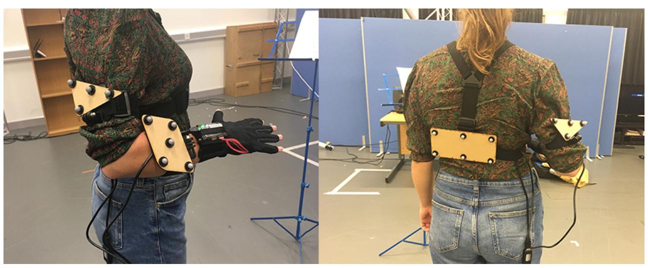

Figure 3: Motion capture marker placement

3.5 Note Selection Error

To examine participant note selection error, the user performed Note On information was compared against the reference note in the score. For example, in the sample provided in Table 1, the participant performed three wrong notes: the second (E3 instead of D3), fifth (G3 instead of D3) and eighth (E3 instead of C3).

3.6 Timing Error

Performer timing error was also examined. A timing error score was assigned to each performed trial by measuring the cumulative difference between the user performed note onset and the corresponding time stamp in the reference MIDI data (see Table 1). In cases where the participant failed to perform a note for a corresponding reference note, an error of one second (the length of one beat/note in the trial) was recorded.

3.7 Results

User perception results for the participants are shown in Figure 4 along with movement smoothness, note selection error and note timing error in Figure 5, where whiskers indicate $1.5 \times$ interquartile range. The user perception responses (Figure 4) indicated a trend in all three aspects of perception towards more positive responses with smaller amounts of system error, with a significant reduction in perceived accuracy, responsiveness and control, once any system error had been introduced.

A Shapiro-Wilk test of normality indicated that the majority of the quantitative results were not normally distributed. Consequently, the data were analysed using non-parametric Friedman rank sum tests and post hoc pairwise Wilcoxon signed rank tests with Bonferroni adjustment. Movement smoothness results were median-normalised by scaling each participant's scores by their median across all conditions. The Friedman test indicated no significant differences between the 0, 1, 2, 5, 10 and 20% error conditions, $\chi^2(5) = 10.6$, $p = 0.059$. However, the Friedman test did indicate a significant deference between observed note timing error at the different error conditions, $\chi^2(5) = 100.549$, $p < 0.001$. Post hoc tests revealed that, when compared with the control condition (0%, $Mdn = 41$), timing errors were greater when error rates were set to 5% ($Mdn = 56$, $T = 352$, $r = 0.36$, $p < 0.001$); 10% ($Mdn = 54$, $T = 352$, $r = 0.19$, $p < 0.01$) and 20% ($Mdn = 69$, $T = 352$, $r = 0.36$, $p < 0.001$). The Friedman test also indicated a significant difference between observed note selection error for the different error conditions, $\chi^2(5) = 19.7$, $p < 0.01$. However, the post hoc tests revealed no significant difference between the control condition (0%, $Mdn = 3$) and the other error rate conditions. However, a significant difference was measured when comparing the results of the 2% error condition ($Mdn = 3$) with the 5% error condition ($Mdn = 4$, $T = 39$, $r = 0.48$, $p < 0.05$) and, similarly, when comparing the 2% and 20% conditions ($Mdn = 5$, $T = 39$, $r = 0.48$, $p < 0.05$).

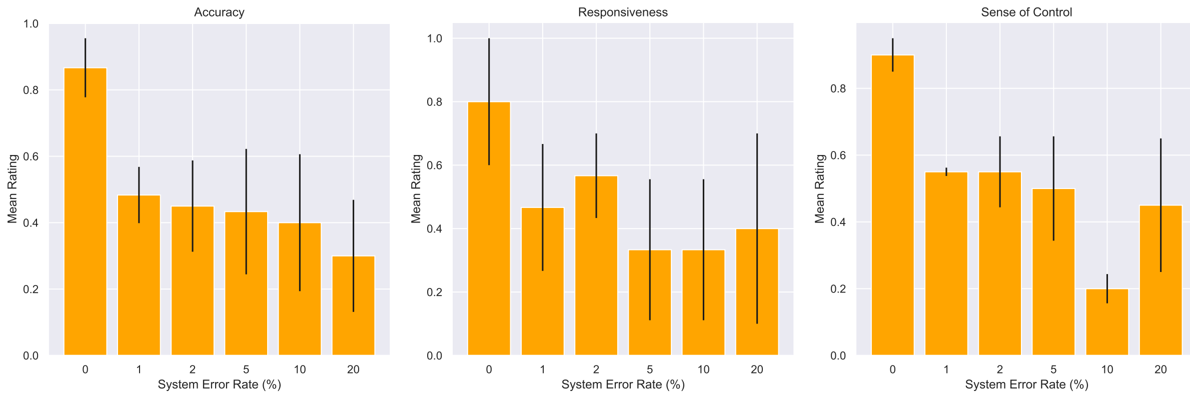

Figure 4: User perception results

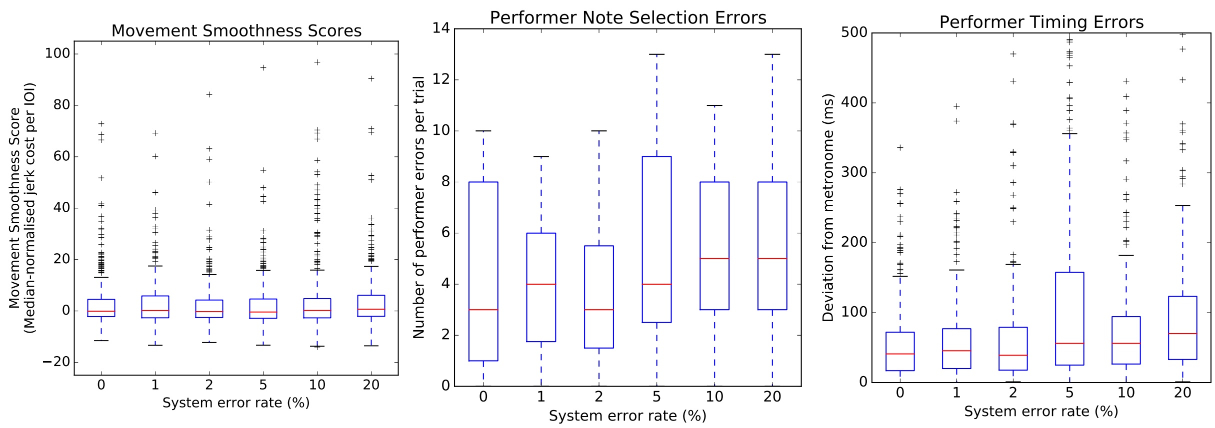

Figure 5: Movement smoothness, note selection error and timing error results

3.8 Discussion and Limitations

The results of this pilot study indicate that system errors can have a negative impact on a both player's perception of, and their ability to perform with, a gestural DMI. The quantitative results suggest that system errors increase the occurrence of performer note selection errors and the severity of timing errors. It appears that when a system erroneously interprets a player's actions, the player is more likely to make a mistake. This observation is to some extent axiomatic: the importance of accurate feedback for musical control is well-established and reducing feedback quality invites uncertainty and error. A clear limitation of this study is the small number of participants which may explain why some of the between group comparisons did not show significant differences. Consequently, results should be treated with some caution. There may also be other factors at play, for example at the highest error rate (20%) several participants were well aware that the system was misinterpreting their actions. It might be the case that less obvious errors occurring at lower rates (i.e. 5%) might be more disruptive as their source is unknown. Perhaps more interestingly, the results of the qualitative study indicate that in the presence of any error, the system was perceived to be significantly less accurate, less responsive and led to a reduced sense of control. The most significant drop in these perceptual factors occurs between system error rates of 0% and 1%, an effect that largely plateaus as the error rate is further increased. This finding suggests that performers could have a very low tolerance for error. If these result are indicative of DMIs more widely this could present an important consideration for the designers of prototype instruments and interfaces that exhibit unintended system errors.

4 CONCLUSION

Musical performances with DMIs are frequently disrupted by unintended errors that can be traced to two possible sources: mistakes made by the performer and technology-related errors made by the instrument. In this paper we defined these two types of error as *performer error* and *system error* respectively, before exploring the impact of the latter in a pilot study with five participants. Participants performed a note selection task using a gestural DMI and their performance and perceptions were analysed as controlled rates of system error were introduced. The errors were designed to resemble genuine system errors that can occur in gestural music systems.

Although the number of participants was small, the study suggests that as system errors increase, the prevalence of performer note selection errors and the severity of timing errors also increase. That is, players are more likely to make mistakes when the system on which they rely is itself prone to error. Additionally, the survey indicated that even 1% of system error has a significant negative effect on a player's sense of control and their perception of the system's accuracy and responsiveness.

The study also indicated that the introduction of error had no measurable effect on the average movement smoothness of a performer's actions, a metric which has been used in previous work as an indicator of skill acquisition. The within group study design may to some extent explain this finding, as each participant completed the task at every rate of system error. Consequently, the randomised order makes it difficult to separate and compare the skill acquired within each trial from the skill carried over between trials. Additionally, the study time might have been too short to capture any meaningful development of skill.

While further investigation is required, these results could highlight important implications for the DMI research community: there appears to be an important link between the unreliability of instruments that we create and the error-proneness and confidence of the performers that use them. DMIs integrating potentially unreliable technologies, materials and mechanisms can result in a high and in some cases inevitable risk of system error, which can disrupt musical performances for both players and audiences. This issue is of particular relevance to systems incorporating machine learning algorithms, that are susceptible to inevitable system errors as a result of noise or insufficient training data. If system error is an unavoidable feature of modern DMIs, further work that attempts to understand how these errors can be accommodated and managed in music performance scenarios will be helpful, building on the initial work of Hazzard et al [26].

The pilot study helps to provide useful parameters for future research and we intend to conduct two further studies with a larger number of participants, performing for longer periods with a reduced set of error conditions (e.g. 0%, 5% and 10%). Firstly, a study identical to the one presented here that examines note selection and timing error. Second, a between group study in which participants within each group perform at a fixed error rate to identify the effects of system error on skill acquisition and movement smoothness.

5 ETHICAL STANDARDS

This study was approved by the the University of the West of England Ethics Committee. All participants provided informed consent.

ACKNOWLEDGMENTS

DB would like to acknowledge support from University of the West of England and thank the Bristol Robotics Lab, Mehdi Sobhani in particular for assisting with the motion capture technology. TJM would like to acknowledge support from the Bristol and Bath Creative Cluster (AHRC) and thank members of the CT Lab, Chagall van den Berg, Rachel Freire, Imogen Heap, Seb Madgwick, Hannah Perner-Wilson, Kelly Snook, Adam Stark and the rest of the Mi.Mu gloves team.

REFERENCES

[1] ABRSM. 2019. Grade 8 Piano Syllabus 2019 and 2020. https://gb.abrsm.org/fileadmin/user_upload/PDFs/Grade_8_Piano_Syllabus_2019___2020.pdf Last Accessed: 9th August 2019.

[2] Scott Adams. 1996. *The Dilbert Principle.* HarperBusiness, New York, New York, NY, USA and London, UK.

[3] Ahmed Sabbir Arif, Wolfgang Stuerzlinger, Euclides Jose de Mendonca Filho, and Alec Gordynski. 2014. Error Behaviours in an Unreliable In-air Gesture Recognizer. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, Toronto, ON, Canada, 1603–1608.

[4] S. M. Astrid Bin, Fabio Morreale, Nick Bryan-Kinns, and Andrew P. McPherson. 2017. In-the-moment and Beyond: Combining Post-hoc and Real-Time Data for the Study of Audience Perception of Electronic Music Performance. In *Human-Computer Interaction - INTERACT 2017*. Springer International Publishing, 263–281.

[5] Steve Benford. 2010. Performing musical interaction: Lessons from the study of extended theatrical performances. *Computer Music Journal* 34, 4 (2010), 49–61.

[6] Dom Brown, Chris Nash, and Tom Mitchell. 2018. Simple Mappings, Expressive Movement: A qualitative investigation into the end-user mapping design of experienced mid-air musicians. *Digital Creativity* 29, 2–3 (August 2018), 129–148.

[7] Dom Brown, Chris Nash, and Tom Mitchell. 2018. Understanding User-Defined Mapping Design for Mid-Air Musical Performance. In *Proceedings of the 5th International Conference on Movement and Computing (MOCO)*. ACM, Genoa, Italy, 27:1–27:8.

[8] Baptiste Caramiaux, Frédéric Bevilacqua, Marcelo M Wanderley, and Caroline Palmer. 2018. Dissociable effects of practice variability on learning motor and timing skills. *PloS one* 13, 3 (2018).

[9] Joel Chadabe. 1977. Some Reflections on the Nature of the Landscape within Which Computer Music Systems are Designed. *Computer Music Journal* 1, 5 (1977).

[10] Palle Dahlstedt. 2012. Pencil Fields: An Expressive Low-Tech Performance Interface for Analog Synthesis. In *Proceedings of the International Conference on New Interfaces for Musical Expression.* https://doi.org/10.5281/zenodo.1178241

[11] Balandino Di Donato, Jamie Bullock, and Atau Tanaka. 2018. Myo Mapper: a Myo armband to OSC mapper. In *Proceedings of New Interfaces for Musical Expression (NIME) 2018*. Blacksburg, VA, USA, 138–143.

[12] Christopher Dobrian and Daniel Koppelman. 2006. The 'E' in NIME: Musical Expression with New Computer Interfaces. In *Proceedings of New Interfaces for Musical Expression (NIME) 2006*. Paris, France, 277–282.

[13] Carolyn Drake and Caroline Palmer. 2000. Skill acquisition in music performance: relations between planning and temporal control. *Cognition* 74, 1 (2000), 1–32.

[14] Robert A Duke, Carla Davis Cash, and Sarah E Allen. 2011. Focus of attention affects performance of motor skills in music. *Journal of Research in Music Education* 59, 1 (2011), 44–55.

[15] Sidney Fels, Ashley Gadd, and Axel Mulder. 2002. Mapping Transparency Through Metaphor: Towards More Expressive Musical Instruments. *Organised Sound* 7, 2 (2002), 109–126.

[16] Rebecca Fiebrink and Perry R Cook. 2010. The Wekinator: a system for real-time, interactive machine learning in music. In *Proceedings of The Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010)*. Utrecht, The Netherlands.

[17] Rebecca Fiebrink and Laetitia Sonami. 2020. Reflections on Eight Years of Instrument Creation with Machine Learning. In *Proceedings of New Interfaces for Musical Expression (NIME)*. Birminham, UK.

[18] Rebecca Fiebrink, Daniel Trueman, N Cameron Britt, Michelle Nagai, Konrad Kaczmarek, Michael Early, MR Daniel, Anne Hege, and Perry R Cook. 2010. Toward Understanding Human-Computer Interaction In Composing The Instrument. In *Proceedings of the International Computer Music Conference (ICMC) 2010*. New York, NY, USA.

[19] Rebecca Fiebrink, Dan Trueman, and Perry R Cook. 2009. A Meta-Instrument for Interactive, On-the-Fly Machine Learning. In *Proceedings of New Interfaces for Musical Expression (NIME) 2009*. Pittsburgh, PA, USA, 280–285.

[20] A Cavan Fyans, Michael Gurevich, and Paul Stapleton. 2009. Where Did It All Go Wrong? A Model of Error From the Spectator's Perspective.. In *Proceedings of New Interfaces for Musical Expression (NIME) 2009*. Pittsburgh, PA, USA, 171–172.

[21] Nicholas Gillian and Joseph A Paradiso. 2014. The Gesture Recognition Toolkit. *The Journal of Machine Learning Research* 15, 1 (2014), 3483–3487.

[22] Werner Goebl and Caroline Palmer. 2013. Temporal control and hand movement efficiency in skilled music performance. *PloS one* 8, 1 (2013).

[23] Lydia Goehr. 1989. Being True to the Work. *The Journal of Aesthetics and Art Criticism* 47, 1 (1989), 55–67.

[24] Víctor Evaristo González-Sánchez, Sofia Dahl, Johannes Lunde Hatfield, and Rolf Inge Godøy. 2019. Characterizing movement fluency in musical performance: Toward a generic measure for technology enhanced learning. *Frontiers in Psychology* 10 (2019), 84:1–84:20.

[25] Michael Gurevich, Paul Stapleton, and Peter Bennet. 2009. Designing for Style in New Musical Interactions. In *Proceedings of New Interfaces for Musical Expression (NIME) 2009*. Pittsburgh, PA, USA, 213–217.

[26] Adrian Hazzard, Chris Greenhalgh, Maria Kallionpää, Steve Benford, Anne Veinberg, Zubin Kanga, and Andrew Mcpherson. [n.d.]. Failing with style: designing

for aesthetic failure in interactive performance. https://doi.org/10.1145/3290605.3300260

[27] Sergi Jordà. 2005. *Digital Lutherie: Crafting musical computers for new musics' performance and improvisation*. Ph.D. Dissertation. Universitat Pompeu Fabra.

[28] Mi.Mu Ltd. 2018. Mi.Mu Gloves. http://www.mimugloves.com. http://www.mimugloves.com Last Accessed: 15th July 2019.

[29] McLean J Macionis and Ajay Kapur. 2018. Sansa: A Modified Sansula for Extended Compositional Techniques Using Machine Learning. In *Proceedings of New Interfaces for Musical Expression (NIME) 2018*. Virginia Tech, Blacksburg, Virginia, USA, 78–81. http://www.nime.org/proceedings/2018/nime2018_paper0018.pdf

[30] David Merrill and Joseph A Paradiso. 2005. Personalization, expressivity, and learnability of an implicit mapping strategy for physical interfaces. In *Proceedings of the CHI Conference on Human Factors in Computing Systems, Extended Abstracts*. 2152–2161.

[31] Thomas Mitchell and Imogen Heap. 2011. SoundGrasp: A Gestural Interface for the Performance of Live Music. In *Proceedings of New Interfaces for Musical Expression (NIME) 2011*. Oslo, Norway, 465—468.

[32] Thomas Mitchell, Sebastian Madgwick, Simon Rankine, Geoffrey Hilton, Adrian Freed, and Andrew Nix. 2014. Making the Most of Wi-Fi: Optimisations for Robust Wireless Live Music Performance. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 251–256. https://doi.org/10.5281/zenodo.1178875

[33] Thomas J. Mitchell, Sebastian. Madgwick, and Imogen. Heap. 2012. Musical Interaction with Hand Posture and Orientation: A Toolbox of Gestural Control Mechanisms. In *Proceedings of New Interfaces for Musical Expression (NIME) 2012*. Ann Arbor, MI, USA.

[34] Paul Modler, Tony Myatt, and Michael Saup. 2003. An Experimental Set of Hand Gestures for Expressive Control of Musical Parameters in Realtime. In *Proceedings of New interfaces for Musical Expression (NIME) 2003*. Montreal, Canada, 146–150.

[35] Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson. 2010. Understanding Users' Preferences for Surface Gestures. In *Proceedings of Graphics Interface 2010 (GI '10)*. Toronto, ON, Canada, 261–268.

[36] Daniel Müllensiefen, Bruno Gingras, Jason Musil, and Lauren Stewart. 2014. The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PloS one* 9, 2 (2014), 1–23.

[37] Donald A Norman and Jakob Nielsen. 2010. Gestural Interfaces: a step backward in usability. *interactions* 17, 5 (2010), 46–49.

[38] Sile O'Modhrain. 2011. A Framework for the Evaluation of Digital Musical Instruments. *Computer Music Journal* 35, 1 (2011), 28–42.

[39] Caroline Palmer, Christine Carter, Erik Koopmans, and Janeen D Loehr. 2007. Movement, planning, and music: Motion coordinates of skilled performance. In *Proceedings of the International Conference on Music Communication Science*. University of New South Wales Sydney, NSW, 119–122.

[40] W. Andrew Schloss. 2003. Using Contemporary Technology in Live Performance: The Dilemma of the Performer. *Journal of New Music Research* 32, 3 (2003), 239–242. https://doi.org/10.1076/jnmr.32.3.239.16866 arXiv:https://www.tandfonline.com/doi/pdf/10.1076/jnmr.32.3.239.16866

[41] Hugo Scurto, Frédéric Bevilacqua, and Jules Françoise. 2017. Shaping and Exploring Interactive Motion-Sound Mappings Using Online Clustering Techniques. In *Proceedings of New Interfaces for Musical Expression (NIME) 2017*. Copenhagen, Denmark, 410–415.

[42] Martin Shepperd, Yuchen Guo, Ning Li, Mahir Arzoky, Andrea Capiluppi, Steve Counsell, Giuseppe Destefanis, Stephen Swift, Allan Tucker, and Leila Yousefi. 2019. The Prevalence of Errors in Machine Learning Experiments. *arXiv* 1909.04436 (2019).

[43] Doug Van Nort. 2018. Conducting the in-between: improvisation and intersubjective engagement in soundpainted electro-acoustic ensemble performance. *Digital Creativity* 29, 1 (2018), 68–81. https://doi.org/10.1080/14626268.2018.1423997 arXiv:https://doi.org/10.1080/14626268.2018.1423997

[44] Richard Vogl and Peter Knees. 2017. An Intelligent Drum Machine for Electronic Dance Music Production and Performance. In *Proceedings of New Interfaces for Musical Expression (NIME) 2017*. Copenhagen, Denmark, 251–256. http://www.nime.org/proceedings/2017/nime2017_paper0047.pdf