View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Queen Mary Research Online

Design for auditory imagery: altering instruments to explore performer fluency

Andrea Guidi
Centre For Digital Music
Queen Mary University of London, UK
a.guidi@qmul.ac.uk

Fabio Morreale
Creative Arts And Industries
University of Auckland, NZ
f.morreale@auckland.ac.nz

Andrew McPherson
Centre For Digital Music
Queen Mary University of London, UK
a.mcpherson@qmul.ac.uk

ABSTRACT

In NIME design, thorough attention has been devoted to feedback modalities, including auditory, visual and haptic feedback. How the performer executes the gestures to achieve a sound on an instrument, by contrast, appears to be less examined. Previous research showed that auditory imagery, or the ability to hear or recreate sounds in the mind even when no audible sound is present, is essential to the sensorimotor control involved in playing an instrument. In this paper, we enquire whether auditory imagery can also help to support skill transfer between musical instruments resulting in possible implications for new instrument design. To answer this question, we performed two experimental studies on pitch accuracy and fluency where professional violinists were asked to play a modified violin. Results showed altered or even possibly irrelevant auditory feedback on a modified violin does not appear to be a significant impediment to performance. However, performers need to have coherent imagery of what they want to do, and the sonic outcome needs to be coupled to the motor program to achieve it. This finding shows that the design paradigm should be shifted from a direct feedback model of instrumental playing toward a model where imagery guides the playing process. This result is in agreement with recent research on skilled sensorimotor control that highlights the value of feedforward anticipation in embodied musical performance. It is also of primary importance for the design of new instruments: new sounds that cannot easily be imagined and that are not coupled to a motor program are not likely to be easily performed on the instrument.

Author Keywords

Performance Studies, Augmented Instruments, Auditory Imagery, Motor Programs, Skill

CCS Concepts

• **Applied computing** → **Performing arts;** *Sound and music computing;* • **Human-centered computing** → *User centered design;*

1. INTRODUCTION

Learning to play new musical interfaces can pose the challenge of a long period of training even for expert musicians. Such a limitation may constitute one of the constraints to the development of expertise and performative practices for new instruments. Expertise is sometimes discussed in NIME in the context of instruments allowing expressivity [21], facilitating the development of virtuosity [3], or within the reports of players who used particular instruments over an extended period of time [15]. Existing expertise is also discussed as an element to preserve to shorten the training needed to master an instrument [3, 13, 18]. However, the role of expertise in guiding design decisions is still relatively under-explored.

This paper specifically looks at questions of using expert musicians' existing skills to embody a new or modified instrument. The focus is placed on the type of expertise that allow players to hear or recreate sounds in the mind even when no audible sound is present [4]. Such capability is generally referred to as *auditory imagery* and is a fundamental component of music execution where players do not consciously think about the instrument [8].

A performer study was designed to include twelve professional violinists, who engaged in a series of musical tasks involving auditory imagery to assess to what extent expert musicians can preserve their expertise when playing unfamiliar sounds, and under which circumstances expert musicians' auditory imagery could support performance fluency of new sounds.

2. BACKGROUND

Auditory imagery and auditory feedback can be described as two core components of instrumental performance [19, 22, 24]. Auditory imagery relates to the ability to hear or recreate sounds in the mind even when no audible sound is present [4] and is part of a feedforward process where musicians generate expectations about the sound they want to execute [23]. Auditory feedback may be defined as the sound produced in response to musicians' physical actions on a musical instrument. Such feedback allows players to tune their performing and achieve subtle adjustments with respect to intonation and articulation. This process particularly characterises traditional musical executions, where the musician has an embodied relationship with the instrument [14]. During the performance, musicians can subconsciously formulate auditory expectations of what the sound will be [1], perform detailed sensory-motor tasks on their instrument to achieve it [8], and eventually adjust their playing by comparing those expectations against the auditory feedback coming from the instrument [8]. Auditory images also function as a bridge between perception and action [2], leading to activation in motor planning areas that serves to guide planning movement [5].

Auditory feedback has been extensively investigated with different strategies, including altering the sound coming from the instrument [16, 17, 6]. This study partially builds

Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copyright remains with the author(s).

NIME'20, July 21-25, 2020, Royal Birmingham Conservatoire, Birmingham City University, Birmingham, United Kingdom.

upon the research presented in [12]. Specifically, it replicates part of the methodologies used to study musicians' reactions to altered auditory feedback and builds upon some of its research findings. In [12], a selected group of professional violinists were asked to engage in a series of musical tasks involving a violin featuring reversed strings. Hence, the connection between violinists' auditory expectations (auditory imagery) and the motor programs to produce the intended sounds was disrupted. Given the reordering of the strings, a gesture that would have usually produced a particular note, resulted in a different note and thus an unfamiliar auditory feedback.

During the study, violinists were asked to play a movement of a Bach partita. Subsequently they were asked to play the same music notation with the same violin as if the strings were not reversed, ignoring the auditory feedback. Players performances were analysed concerning duration and pitch accuracy. Findings suggest that participants were able to play more accurately and fluently when they ignored the auditory feedback coming from the violin and instead focused on their auditory expectations and motor programs to meet them. The study presented in this paper builds upon the results of the mentioned paper to establish auditory imagery as a critical component to support skill transfer between musical instruments.

3. METHODOLOGY

Auditory imagery is part of a process that unfolds subconsciously where actions are accomplished without conscious thought. Hence posing direct questions to musicians could potentially lead to biased information. James, in his Principles of Psychology, writes "We are then aware of nothing between the conception and the execution. All sorts of neuromuscular processes come between... but we know absolutely nothing of them. We think the act, and it is done" [7]. Existing literature presents how auditory imagery can be queried using the measurement of physiological responses and performance-based tasks [1]. In the study presented in this paper, musicians were asked to perform musical assignments following a pre-defined protocol. Corrective effects, and mistakes were considered as data to test a series of research hypotheses.

Raw data from the study can be found at this link.

4. PERFORMER STUDY

This study was designed to examine expert musicians' ability to use auditory imagery to play in tune in the presence of familiar and unfamiliar sounds.

The term *familiar material* describes musical notes and music passages situated in the Western tonal music tradition. These sounds, which are generally referred to as semitones, are familiar for their being coherent with the type of studies and training followed by the selected violinists. For these music events, familiarity reflects both in the auditory imagery developed while studying the instrument and by the auditory feedback produced by playing it. The term *unfamiliar pitch material* refers to quarter tones. Quarter tones were chosen for being rare in the Western traditional music repertoire. Hence, a type of pitch material for which the selected violinists did not have a developed auditory imagery.

The study was divided into two parts. Tasks involving semitones and quarter tones are situated in the first part and refers to the following research question *how accurately can violinists play pitch material for which they do not have a developed auditory imagery? If the actual target sound is unfamiliar, how does that differ in terms of accuracy and response time to being able to play ordinary chromatic music on a violin?* Our hypothesis is that even expert musicians, when confronted with pitch material sitting outside their auditory imagery, will perform with lower accuracy. If this hypothesis was confirmed, we could highlight a fundamental limitation of making a new instrument. If a particular musical feature does not feed into a kind of culturally defined structure of how people think they hear music, then they struggle to play.

Table 1: **The three conditions in the Quarter-tone performance Study Section**

Notation A	Audio Playback	Notation B
12 single notes	24 single notes	8 single notes
12 intervals	24 intervals	12 intervals

In the second part of the study, *unfamiliar pitch material* defines the notes coming from a re-tuned violin. In that case, unfamiliarity reflects in the unexpected auditory feedback coming from the violin. Executing the gesture to play a specific note, because of the re-tuning, resulted in the instrument producing a different note. This modification challenged the expectations of the violinists and therefore, their auditory imagery. Tasks involving the re-tuned violin referred to the research question *to what extent does performers' pitch accuracy and fluency deteriorate in the presence of an unfamiliar auditory feedback and mismatched auditory imagery? Does their performance improve if the task allows participants to ignore the sound coming from the violin and to focus on their internal representation of sound (i.e. their auditory imagery)?*

4.1 Participants

Auditory imagery is an ability well-developed by players with at least seven years of training [25]. For this reason, only musicians holding an ABRSM's Grade 7 [1] and above were selected.

Twelve participants in an age range between 22 and 46 years old were selected with an open call sent through music schools. Each violinist filled a questionnaire before beginning the study indicating the years of training, the type of repertoire that characterised their study and their performing, and their familiarity with string instruments retuning. All accepted violinists were trained in the Western music repertoire, were unfamiliar with quarter tones and with string instruments re-tuning (scordatura).

4.2 Quarter-tone performance

The first part of the study tests the presence of already-existing auditory imagery for pitches outside the Western traditional music repertoire. Violinists engaged in a series of tasks involving quarter tones. The execution was evaluated regarding accuracy in pitch and time response comparing semitones and quarter tones. The underlying assumption is that participant's auditory imagery of quarter tones is vaguer than for semitones. Therefore, it could be possible to see pitch accuracy be lower during tasks involving quarter tones compared to semitones.

4.2.1 Procedure

Violinists were asked to play semitones and quarter tones notes with their personal violin in three separate conditions: notation A, audio playback, notation B. Violinists were asked to avoid using vibrato in their executions. The

---

[1] ABRSM (Associated Board of the Royal Schools of Music) is an accredited board awarding exams and diploma qualifications in music within the UK

audio section was used to discriminate notation effects coming from reading the unfamiliar symbols describing quarter tones. Quarter tones were notated using the symbols ♭↓ ♮↓ ♯↓. Table 1 present the number of stimuli for each section. In each section single notes were shown on a monitor one at a time, then intervals were shown one at a time with the same modality. Every single note and every melodic interval was presented only once to each participant to prevent biases resulting from playing the same stimulus twice (i.e. potential learning effects). The first note of melodic intervals (which was always a semitone [2]) served to provide a context to the second note (the actual stimulus which could be a semitone or a quarter tone). The reason for having two notation sections (one at the beginning and one at the end) was to test whether having the quarter tones presented as audio would improve pitch accuracy when going back to playing from notation.

The audio coming from the violin was recorded with a DPA microphone and analysed to measure pitch deviation for each played note. The monophonic pitch tracking algorithm pYin [10] was used to measure pitch accuracy of participants' executions. Each note recording was segmented into start (100 ms following the attack phase), mean (the time interval following the start the note up to the 100 ms before the end of the note), and end (the last 100ms before the violin bow detached from the strings of the violin).To measure the pitch deviation we calculated the expected frequency $f_E(n)$ as

$$f_e(n) = m_0 2^{d_n/12}$$

where m_0 is the average frequency of the first note (the lowest G), and d_n is the difference in semitones between the note n and the lowest G. We then calculated the accuracy of the played note at different moments by calculating, for each note, the error in cents for the beginning aB_n and the end of the note aE_n and for the mean frequency aM_n

$$aB_n = 1200 * log_2(mB_n/fE_n)$$

$$aE_n = 1200 * log_2(mE_n/fE_n)$$

$$aM_n = 1200 * log_2(mM_n/fE_n)$$

Wilcoxon signed-rank tests were conducted to assess the significance of the comparison between absolute values of the pitch deviation for semitones and quarter tones in the three considered phases of the note (start, mean and end) for notation A plus audio playback plus notation B data, notation A plus notation B data, audio playback data. We also measured the time interval between each single note being presented as a notation and the start of its execution. Specifically, we used a light sensor which measured a blinking box on the display each time a new notation stimulus was presented. The recording of the sensor was synchronous with the audio recording of the violin. Both signals were recorded using Bela [11], a low-latency platform for sensor and audio processing. The significance of the comparison between time response for semitones and quarter tones was assessed with Wilcoxon signed-rank test across the two notation sections (A, and B). For this paper, we are only presenting the data coming from the analysis of the single notes. Data from the analysis of the intervals will be part of a future publication.

4.2.2 Semitones versus quarter tones

Data from notation A, audio playback, and notation B sections revealed that semitones were performed with better pitch accuracy and with a shorter time response than quarter tones.

Pitch was more accurate at the beginning of semitones with a mean pitch deviation of 26.91 vs 42.13 cents. This result was statistically significant with p $\approx$ 5.34E-14 and a moderate effect size 0.23. Pitch was also more accurate for semitones in the mean of performers' playing with a mean pitch deviation of 18.09 vs 32.73 cents (p $\approx$ 7.8E-23 and effect size 0.30), and at the end of the note with 17.20 vs 31.70 cents (p $\approx$ 2.25E-23 and effect size 0.30).

In the notation sections A plus B semitones were performed sooner than quarter tones with a mean time response of 1.52 vs 1.90 seconds. Statistical significance for this comparison is p $\approx$ 6.11E-09, effect size 0.39.

Table 2 presents a summary of data from this section.

4.2.3 Notation A plus Notation B versus audio

Data from Notation A section plus data from the Notation B sections shows that semitones were played more accurately than quarter tones at the beginning of the notes with a mean pitch deviation of 26.38 vs 40.92 cents, and p $\approx$ 3.12E-06. However effect size was small at 0.20. Pitch was also more accurate for semitones in the mean, and the end of performers' executions with a mean pitch deviation of 22.60 vs 40.04 cents (p $\approx$ 8.18E-11, effect size 0.28), and 22.1 vs 39.519 cents (p $\approx$ 2.00E-10, effect size 0.28). In the audio section, semitones were still played more accurately than quarter tones with a mean pitch deviation of 27.43 vs 43.32 cents (p $\approx$ 2.37E-09, effect size 0.26) at the start, 13.64 vs 25.52 cents (p $\approx$ 4.14E-14, effect size 0.33) in the mean, and 12.37 vs 23.99 cents (p $\approx$ 2.52E-15, effect size 0.34) at the end of the executions.

In the beginning of the notes, the intonation for semitones and quarter tones was more accurate in the notation section. By contrast, during the mean and the end of the notes, intonation for semitones and quarter tones was more accurate in the audio section.

4.3 Re-tuned Violin

In part 2, participants were asked to perform a series of music passages with a re-tuned violin. Because of the violin re-tuning, imagining a certain note and playing it corresponded to the sound of a different note. The violin was re-tuned to disrupt the relation between auditory expectations and performed gesture. A potential outcome of such a disruption is impaired fluency, a state where it is not possible to play something at tempo or with proper rhythm or intonation because it is necessary to pay conscious attention to each action.

Three different conditions were designed to answer to the following question: *when auditory-imagery and motor programs are decoupled with a re-tuned instrument, which are the circumstances that allow faster/more accurate executions?* Each condition was explained to participants beforehand. In concert notation musicians played musical excerpts notated for violin (concert notation). This condition tested whether inducing players to think of a new connection between their auditory imagery and their motor programs to achieve the right sound result in a loss of pitch accuracy and fluency. In the transposed notation condition musicians played musical excerpts whose notated pitches are altered to account for the different tuning of each violin

---

[2] more specifically, a note from the non microtonal 12-TET scale

Table 2: **mean Pitch Deviation and mean Time Response for Semitones and Quarter Tones considering data from notation A plus audio playback plus notation B sections**

	Semitones	Quarter Tones	P value
Start of the note: pitch deviation	26.91 cents	42.13 cents	≈ 5.34E-14
Mean of the note: pitch deviation	18.09 cents	32.73 cents	≈ 7.80E-23
End of the note: pitch deviation	17.20 cents	31.70 cents	≈ 2.25E-23
Time Response	1.52 seconds	1.90 seconds	≈ 6.11E-09

string. The notation indicates the pitch that would sound on a normally-tuned violin, but the sounding pitch will be different. This condition was designed to allow musicians to ignore the auditory feedback coming from the violin and to focus on the notation. The condition tested whether participants could perform fluently in presence of a notation that shows the right sensory-motor actions while producing a unfamiliar auditory feedback. By using notation, we designed a direct link to the performer's motor programs and auditory-imagery to assess if pitch accuracy and fluency improves. In the audio playback condition, musicians replicated a music passage reproduced by a speaker. This condition allowed to discriminate the influence of notation in the previous two sections, and was designed to assess whether playing by ear on a re-tuned instrument is at least as hard as by playing with concert notation.

4.3.1 Procedure

In this part of the study, violinists were asked to perform short musical passages (between two and seven seconds of duration) on a re-tuned violin. Musicians were asked to avoid using vibrato in their executions. Music passages were selected among unknown compositions from the Western baroque music repertoire. Each music passage was performed by participants twice. The intent was to present short melodies that were familiar to the violinist's cultural space but that they would have not seen before. In this second part, violinists used a re-tuned violin made available by the researchers. A violin is usually tuned from the lower string to the upper string as follows: G3, D4, A4, E5. Starting from the lower string up to the higher string the violin was tuned in fourths as follows: A3, D4, G4, C5.

Violinists were asked to perform music passages in three different conditions. In the first condition (concert notation), eight music passages were notated as they are. In the second condition (transposed notation) eight music passages were notated as they sound on the re-tuned violin (the notation was transposed to match the violin re-tuning). In the third condition (audio playback), eight melodies were reproduced through speakers. Notations included instructions on tempo.

The order of the transposed notation, concert notation and audio playback sections was randomised for each participant. Participants were asked to play only in the first position. It was critical to have violinists to move between strings to maximise the disruption produced by the re-tuning and observe its effects. Each music passage was presented only once to avoid biases resulting from having already played a melody in a previous section. Violinists were asked to perform each melody twice.

The audio coming from the violin was recorded with a DPA microphone. The software Tony [9] was used to automatically segment players' recordings for each passage into notes and provide frequency estimates. Pitch deviation was computed in three steps. First we identified the closest pitch on the chromatic scale for each played note (using a custom algorithm), then we measured the average frequency of the note (using the monophonic pitch tracking algorithm pYin [10]), and finally we applied the following formula

$$pitchDev_n = 1200 * log_2(averageF_n/expectedF_n)$$

Wilcoxon signed-rank tests were conducted to assess the significance of the comparison between absolute values of the pitch deviation for concert notation and transposed notation, then for concert notation and audio playback and finally for transposed notation and audio playback. For this test, and the following Wilcoxon tests, the resulting p value and effect size were reported to facilitate the interpretation of the data.

The length of each execution was divided by the expected length of the music passage (calculated by converting its notation to an audio file using the software Muse Score [20]). The resulting ratios highlight the relation between the duration of the performances and the duration of the music passages.

Wilcoxon signed-rank test [23] was conducted to assess the significance of the comparison between duration ratios in the concert notation and in the transposed notation conditions.

The number of notes in the performance recording of each music passage was compared to the number of notes forming the melody to estimate the number of extra notes played. The term *extra* identifies notes that are extraneous to the notation to be executed (i.e. repeated notes or wrong notes performed in the attempt of finding the right finger position on the re-tuned violin). The resulting data should be taken as approximate as music passages were segmented using an automatic music transcription algorithm.

Wilcoxon signed-rank test was conducted to assess the significance of the comparison between the number of extra notes in the concert notation and in the transposed notation conditions.

In the audio section, participants struggled to identify and repeat music passages. Most of the passages were not performed entirely, and the duration of the execution was randomly shorter or longer according to the difficulty of the passage (which could be quantified in number of notes to remember, tempo in bpm). Therefore, it was decided to exclude the audio section from the extra notes and performances duration data analysis.

4.3.2 Transposed notation versus concert notation

Performers' pitch accuracy in the transposed notation condition compared to the concert notation condition presented a deviation mean of 14.91 vs 17.30 cents with $p \approx 0.001$. However, effect size was small at 0.039.

Performances duration was closer to the music passages duration in the transposed notation condition compared to the concert notation condition with a mean of 3.69 vs 6.46 seconds, $p \approx$ 1.75E-16, and effect size 0.40.

Executions of transposed notation were characterised by less extra-notes compared to concert notation. Specifically 49 versus 770 notes with $p \approx$ 1.75E-16, effect size 0.30.

A summary of data presented in this section is showed in table 3.

Table 3: **mean Pitch Deviation, Performance Duration, and Number of Extra Notes for Transposed Notation and Concert Notation**

	Transposed Notation	Concert Notation	P value
Pitch Deviation	14.91 cents	17.3 cents	$\approx$ 0.001
Performance Duration	3.69 seconds	6.46 seconds	$\approx$ 1.75E-16
Extra Notes	49	770	$\approx$ 9.11E-10

4.3.3 Transposed notation versus audio playback

Performers' pitch accuracy was better in the transposed notation condition compared to the audio playback condition with a pitch deviation mean of 14.91 vs 17.42 cents with p $\approx$ 1.17E-11, small effect size at 0.09.

4.3.4 Concert notation versus audio playback

Performers' pitch accuracy in the concert notation condition compared to the audio playback condition presented a deviation mean of 17.03 vs 17.42 cents (p $\approx$ 4.32E-05, small effect size at effect size 0.05).

5. DISCUSSION

The first part of the study suggests that if a musician does not have a familiar mental image of the sound that is intended to play, then it does not matter whether the instrument is familiar or not. During the study, the violinists were playing an instrument that they know very well. Nonetheless, they struggled with their pitch accuracy, and time response in playing quarter tones both in the notation and audio sections. We propose that quarter tones were not played as well as semitones as participants lack a coherent sonic imagination coupled to the needed motor programs.

In the second part of the study, our results suggest that the transposed notation led to more fluent executions. Music passages had a duration and number of notes more coherent with the indications presented in the notation compared to both concert notation and audio playback performances. It is proposed that transposed notation worked so well because it let musicians use their familiar auditory imagery, which, in turn, is connected to the actions needed to perform the notation. Even if the resulting sound was not the one they imagined to be, the fluency of the performance improved. In the transposed notation scenario, the imagery was deliberately incorrect; the notation did not produce the sounds that it specifies. Nonetheless, it worked better to use a notation that *specified the correct motor actions* and that *produced* a "wrong" but familiar auditory result (transposed notation) than to have a notation *specifying* a "correct" but unfamiliar auditory feedback (concert notation). Hence, these results suggest that having access to an existing auditory imagery and motor programs is more critical than perceiving the correct auditory feedback.

5.1 Implications for instrument design

In this study, we saw that playing quarter tones was a challenging task even with a traditional violin with no modifications. In this condition, having the right mapping precision, and auditory feedback did not support musicians. When the auditory imagery of a musician struggles to identify a particular sound event, then the motor program to perform it is also unavailable. The resulting accuracy of the performance diminishes (quarter tones and concert notation). When the auditory feedback is unfamiliar, but musicians can access their auditory imagery, and its related motor programs, the accuracy of the performance (in terms of pitch precision and fluency) is better preserved (transposed notation). Additional examples that demonstrate these principles are prepared piano, and MIDI keyboards. A piece for these instruments can still be notated as a piano piece, and players will not have troubles when playing it even in presence of unfamiliar auditory feedback. Possibly because they can imagine the traditional music space where the keyboard and the notation sit, and they can play the instrument whatever sound it produces.

In NIME discussions, there is often-times a focus on the ability to produce any given sound using mapping strategies and technological solutions. We propose that a lack of auditory imagery constitutes a mental limitation rather than a technical constraint in the design process. Hence, the design lens should shift from a technology-focused view to include a more human-based perspective to address such a limitation. Taking a technocentric approach could only account for questions concerning topics like mapping, precision, and degree of freedom. However, it would be insufficient to only consider human-based aspects like ergonomics of the interface, and useful feedback (audio, tactile, visual). In this is paper we proposed a further level of attention, which accounts for the sensory-motor link in music performance, and the feedforward mechanism that describes it. We propose this process as an essential element to account for in the design of an interface that is meant to produce new sounds. If the kind of performance that is aimed to enable is the one afforded by a traditional instrument (where musicians have an embodied relationship with it), then it almost does not matter what the interface is. If players cannot imagine the sounds, then they can not play them.

This approach poses a fundamental difference from asking if an instrument can afford to play microtones, or is the instrument ergonomically sensible to the hands of the player. Including the feedforward process in the design process leads to the following design question: is there an intrinsic link between a particular interface that is being designed, and the existing imagination of performers that will play the interface plus the execution techniques they already own? If the answer is no, then the instrument would run in a fundamental limitation which is mental rather then technological. If the goal as a designer is to let people play unfamiliar sounds, then this human factor needs to be taken into account.

6. CONCLUSIONS

The results of this study challenge the notion of playing as a feedback loop where performers think what they want to do, they play it, and then they evaluate the result based on the feedback to correct the performance. This study suggests that this feedback loop it is not critical for the fluent execution of a performance. Feedback may nonetheless be needed for correction and refinement (i.e. the ends of each note in the first part of the study were typically more in tune than the beginning which we wouldn't expect to happen without feedback). Musicians can still play when the imagery is not entirely aligned to the auditory feedback (as in the transposed notation condition) as long as their anticipation leads them to the right motor program. Musicians can potentially substitute unfamiliar imagery if they can leverage a notation system that it is connected to their existing auditory imagery.

To support the kind of performance that is guided by auditory imagination (i.e. execution with music scores), there is a need to translate from the imagery of sound to an action that is congruent and appropriate on the instrument. If musicians do not have that connection (i.e. because the instrument is totally unfamiliar), then they will also lack the skills needed to perform with the instrument or its augmentation. Establishing that connection is a goal that provides space for more research questions like *to what extent is it possible to use existing auditory imagery to play unfamiliar sounds? For which musical aspects other than this approach would it be valid?*

In this study, we focused on pitch and fluency. However, we could also consider other musical aspects like articulation. Finally, when participants where playing transposed notation, were they truly completely ignoring their auditory feedback? Which aspects were still taken into account (spectral envelope, articulation)? These questions constitute material for future research.

As shown with concert notation, while thinking and reasoning, the quality of musicians' performance deteriorated. Design strategies that assure that performance remain automatic are needed and could rely on auditory imagery.

7. ACKNOWLEDGMENTS

This research is supported by EPSRC under the grants EP/L01632X/1 (Centre for Doctoral Training in Media and Arts Technology) and EP/N005112/1 (Design for Virtuosity).

8. ETHICAL STANDARDS

Violinists received an information sheet explaining the nature and demands of the research. They also signed a consent form to take part in the study.

9. REFERENCES

[1] T. Clark, A. Williamon, and A. Aksentijevic. Musical imagery and imagination: The function, measurement, and application of imagery skills for performance. *Musical imaginations: Multidisciplinary perspectives on creativity, performance, and perception*, pages 351–365, 2012.

[2] I. D. Colley, P. E. Keller, and A. R. Halpern. Working memory and auditory imagery predict sensorimotor synchronization with expressively timed music. *The Quarterly Journal of Experimental Psychology*, (just-accepted):1–49, 2017.

[3] C. Dobrian and D. Koppelman. The 'E' in NIME: Musical expression with new computer interfaces. In *NIME*, volume 6, pages 277–282, 2006.

[4] R. I. Godøy. Gestural imagery in the service of musical imagery. In *International Gesture Workshop*, pages 55–62. Springer, 2003.

[5] E. B. Greenspon, P. Q. Pfordresher, and A. R. Halpern. Pitch imitation ability in mental transformations of melodies. *Music Perception: An Interdisciplinary Journal*, 34(5):585–604, 2017.

[6] H. Hafke-Dys, A. Preis, and D. Trojan. Violinists' perceptions of and motor reactions to fundamental frequency shifts introduced in auditory feedback. *Acta Acustica united with Acustica*, 102(1):155–158, 2016.

[7] W. James. of William James's experimental psychology. *Reflections on the Principles of Psychology: William James After A Century*, page 33, 2013.

[8] P. E. Keller. Mental imagery in music performance: underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, 1252(1):206–213, 2012.

[9] M. Mauch, C. Cannam, R. Bittner, G. Fazekas, J. Salamon, J. Dai, J. Bello, and S. Dixon. Computer-aided Melody Note Transcription Using the Tony Software: Accuracy and Efficiency. In *Proceedings of the First International Conference on Technologies for Music Notation and Representation*, 2015.

[10] M. Mauch and S. Dixon. pYin: A fundamental frequency estimator using probabilistic threshold distributions. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 659–663. IEEE, 2014.

[11] A. McPherson and V. Zappi. An environment for submillisecond-latency audio and sensor processing on BeagleBone Black. In *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015.

[12] F. Morreale, J. Armitage, and A. McPherson. Effect of instrument structure alterations on violin performance. *Frontiers in psychology*, 9:2436, 2018.

[13] F. Morreale, A. Guidi, A. McPherson, et al. Magpick: an augmented guitar pick for nuanced control. New Interfaces for Musical Expression, 2019.

[14] A. Paivio. Cognitive and motivational functions of imagery in human performance. *Canadian journal of applied sport sciences. Journal canadien des sciences appliquées au sport*, 10(4):22S–28S, 1985.

[15] C. Palacio-Quintin et al. The hyper-flute. In *NIME*, volume 3, pages 206–207, 2003.

[16] P. Q. Pfordresher. Auditory feedback in music performance: the role of melodic structure and musical skill. *Journal of Experimental Psychology: Human Perception and Performance*, 31(6):1331, 2005.

[17] P. Q. Pfordresher and C. Palmer. Effects of hearing the past, present, or future during music performance. *Perception & Psychophysics*, 68(3):362–376, 2006.

[18] C. Poepel. On interface expressivity: A player based study. In *NIME*, volume 5, pages 228–231, 2005.

[19] M. H. Ruiz, H.-C. Jabusch, and E. Altenmüller. Fast feedforward error-detection mechanisms in highly skilled music performance. In *Proceedings of the International Symposium on Performance Science*, pages 187–197. Citeseer, 2009.

[20] M. Watson. Musescore. *Journal of the Musical Arts in Africa*, 15(1-2):143–147, 2018.

[21] D. Wessel and M. Wright. Problems and prospects for intimate musical control of computers. *Computer music journal*, 26(3):11–22, 2002.

[22] D. M. Wolpert, Z. Ghahramani, and M. I. Jordan. An internal model for sensorimotor integration. *Science*, 269(5232):1880–1882, 1995.

[23] R. Woolson. Wilcoxon signed-rank test. *Wiley encyclopedia of clinical trials*, pages 1–3, 2007.

[24] R. J. Zatorre, J. L. Chen, and V. B. Penhune. When the brain plays music: auditory–motor interactions in music perception and production. *Nature reviews neuroscience*, 8(7):547–558, 2007.

[25] R. J. Zatorre, A. R. Halpern, and M. Bouffard. Mental reversal of imagined melodies: a role for the posterior parietal cortex. *Journal of cognitive neuroscience*, 22(4):775–789, 2010.