Motor Variability in Complex Gesture Learning: Effects of Movement Sonification and Musical Background

WANYU LIU, MICHELLE AGNES MAGALHAES, STMS, IRCAM, CNRS, Sorbonne Université, France
WENDY E. MACKAY, MICHEL BEAUDOUIN-LAFON, Université Paris-Saclay, CNRS, Inria, France
FRÉDÉRIC BEVILACQUA, STMS, IRCAM, CNRS, Sorbonne Université, France

With the increasing interest in movement sonification and expressive gesture-based interaction, it is important to understand which factors contribute to movement learning and how. We explore the effects of movement sonification and users' musical background on motor variability in complex gesture learning. We contribute an empirical study in which musicians and non-musicians learn two gesture sequences over three days, with and without movement sonification. Results show the interlaced interaction effects of these factors and how they unfold in the 3-day learning process. For gesture 1 that is fast and dynamic with a direct "action-sound" sonification, movement sonification induces higher variability for both musicians and non-musicians on day 1. While musicians reduce this variability to a similar level as no auditory feedback condition on day 2 and day 3, non-musicians remain to have significantly higher variability. Across 3 days, musicians also have significantly lower variability than non-musicians. For gesture 2 that is slow and smooth with an "action-music" metaphor, there are virtually no effects. Based on these findings, we recommend future studies to take into account participants' musical background; consider longitudinal study to examine these effects on complex gestures; and be aware when interpreting the results given a specific design of gesture and sound.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; *Sound-based input / output*.

Additional Key Words and Phrases: motor variability, complex gesture learning, auditory feedback, movement sonification, musical background

ACM Reference Format:
Wanyu Liu, Michelle Agnes Magalhaes, Wendy E. Mackay, Michel Beaudouin-Lafon, and Frédéric Bevilacqua. 2022. Motor Variability in Complex Gesture Learning: Effects of Movement Sonification and Musical Background. *ACM Transactions on Applied Perception* 19, 1, Article 2 (January 2022), 22 pages. https://doi.org/10.1145/3482967

1 INTRODUCTION

While simple strokes on touch interfaces have become ubiquitous in Human-Computer Interaction (HCI), the use of more complex gestures, which has been the goal of a range of researchers for many years [22], remains challenging. Recently, several groups proposed to extend simple commands to interaction modalities that convey expressiveness in mobile interaction [3], in mid-air movements [2], and in whole-body interaction in games [49], among many others. These endeavors intersect with research on interactive systems in music and dance (e.g. [12, 42, 59]). We believe that these two perspectives, HCI and music technology are particularly fruitful: on one hand, music experts bring insights on designing movement-based interactive systems, and on the other hand, interactive systems bring novel tools for music pedagogy and performance [7, 48].

Authors' addresses: Wanyu Liu, Michelle Agnes Magalhaes, AbbyWanyu.Liu@ircam.fr,michelleagness@gmail.com, STMS, IRCAM, CNRS, Sorbonne Université, Paris, France, F-75004; Wendy E. Mackay, Michel Beaudouin-Lafon, Université Paris-Saclay, CNRS, Inria, Orsay, France, F-91400, mackay@lri.fr,mbl@lri.fr; Frédéric Bevilacqua, Frederic.Bevilacqua@ircam.fr, STMS, IRCAM, CNRS, Sorbonne Université, Paris, France, F-75004.

In movement-based music systems (sometimes also referred to as embodied music systems) [8, 80], gestures and movements are generally considered to convey various information layers simultaneously, allowing for example to trigger specific sounds, to modulate them and to communicate the performer's intention [38]. This requires performer to practice movement to reach such a conscious control of their action, thus implying sensorimotor learning [43]. Supposedly, practice makes perfect. However, when someone tries to perform the same movement twice, the two movements will never be completely identical [54]. This is commonly referred to as motor variability – the variance of movements generated by an individual under the same task conditions. Variability can come from errors in a given movement as a reflection of inconsistency in a motor skill (e.g. [19, 76]), or come from the fact that human movement is inherently noisy as a neuromascular system [84]. Consequently, motor variability is generally significantly reduced in the course of learning but is never reduced to zero. In fact, several studies have observed that motor variability exists even in highly skilled performers as a key signature of adaptability [72]. For instance, experienced musicians actively value and control motor variability to express individual styles and convey subtle differences in emotion [32].

Many factors, individually or together, affect motor variability in movement learning. These factors include task complexity, learning schedule [14], feedback mechanism [74] and user profile [73]. For example, a complex task that involves more degrees of freedom results in more variability than a simpler task [87]; experts tend to have more precise control of movement than novices [73]; visual, auditory and haptic feedback also facilitate movement learning in different ways, therefore influence motor variability differently.

Specifically, sonification has been shown to facilitate motor learning [27] and to influence movement execution [9]. We refer the term *movement sonification*[1] to a broad set of auditory feedback techniques, from direct transformation of real-time motion data to sound to techniques that allow to control or modify sound or musical contents rendered in real-time [5]. The effectiveness of movement sonification on facilitating learning, however, has been mostly explored for simple tasks [74] and is difficult to be generalized to more complex gestures. Importantly, the interest of movement sonification lies not only on the possible feedback provided to the users on their postures or movements, but also lies in motivational aspects as reported by Nikmaramv et al. [56]. Nevertheless, the conditions and contexts favoring movement sonification remain largely incomplete. One study [37], with a limited statistical analysis, postulates that auditory feedback is beneficial to expert users but not to novice users. Furthermore, Effenberg and Mechling mention that the efficiency of auditory feedback depends on musical abilities [26]. For instance, Neuhoff and Wayand show that higher musical abilities afford a better pitch discrimination [52].

Our goal is to investigate how music practices and sonification influence motor variability in movement learning. We focus on complex gesture sequences similar to the ones applied in movement-based music systems. Typically, such gestures cannot simply be reduced as a mere sequence of units due to gestural co-articulation [6], i.e. boundaries of each units tend to blur. Therefore, it is necessary to consider motor variability over the whole sequence, occurring over several learning sessions. As reported in the review of Sigrist et al. [74], these factors have not been studied together before.

We first review related work, and then describe the design of two gesture sequences and their respective movement-sound relationships. Next, we present an experiment in which 24 musicians and non-musicians learn these two gesture sequences over three days, with and without sonification. We present the results, and discuss the effects of movement sonification and musical background on motor variability. We conclude with directions for future research.

[1]We also include here techniques sometimes referred to as *musification.*

2 RELATED WORK

We review studies of movement learning and motor variability, movement sonification as augmented feedback and complex gesture learning.

2.1 Movement Learning and Motor Variability

Although learning a motor skill feels like a unitary experience, psychologists and cognitive scientists who study motor skill learning break the process into a number of interacting components. One well-known example is Fitts' three-stage model [29]: the *cognitive stage* where a new learner is trying to understand what to do [70] involves information processing and conveyance [1]; the *associative stage*, characterized by less verbal information, smaller gains in performance and conscious performance with adjustments; and the *automatic stage* where motor performance becomes largely automatic, with minimal cognitive processing demands. In general, the cognitive approach models sensorimotor learning as building an internal motor program that is activated when performing. Other approaches consider interacting with the environment fundamental to learn and execute movements.

The traditional view of motor control is made up of an invariant part, the task (e.g. reaching a target that has certain size at a certain distance [28]), and the parameters that allow the task to be adapted to a given context [68]. This approach considers variability in a given movement pattern to be the result of error and therefore movement learning involves gradually reducing motor variability: movement patterns for skilled performers are invariant. While many studies in movement science demonstrate the gradual reduction of variability over time (e.g. [14, 73]), the invariance of movement for experts has been challenged. From a biomechanical standpoint, the human neuromuscular system is inherently *noisy*, thus it is impossible to obverse two completely identical movements [84], even for skilled experts. A more recent view on motor variability in the field of sports and dance, considers it beneficial to develop more adaptive and skilled performance [36]: Dhawale et al. pointed out that variability is beneficial for motor learning [20]; Braun et al. demonstrated that variability allows structure learning [10]. Wu et al. also showed that temporal structure of motor variability can predict motor learning ability [85].

Motor variability can further be divided into intra-individual variability and inter-individual variability. Intra-individual variability is the variability in behavior, or in any signal, of a single individual measured across multiple time points. In contrast, inter-individual variability is defined as the variability in behavior or in any signal exhibited by multiple individuals at a single time point [82]. In this article, we focus on intra-individual variability to explore how movement learning is subject to internally generated noise and how movement sonification and musical background modulate it.

2.2 Sonification as Augmented Feedback

An important question for movement learning is whether and how feedback can be designed to facilitate the learning process. Feedback, in a general sense, is information made available through action, which can be either *intrinsic* or *extrinsic*. While 'intrinsic' feedback is constrained by familiar laws of cause-and-effect [46, 47], 'extrinsic' feedback is generated by an additional system to – and does not arise naturally from – the immediate task [69]. It is often described as 'augmented' feedback. Using sound as augmented feedback, in particular movement sonification in contrast to an auditory alarm[2], has been widely explored in fundamental studies of motor learning, and in specific applications: sports, medicine, pedagogy and performing arts. Movement sonification can be seen as a subfield of sonification research and interactive sound design [31] that explores how sound and music can carry information to users, change perception (from either a first- or third-person perceptive) [11, 26, 51, 55] and more generally increase motivation for certain tasks. From a technological

[2]An auditory alarm often refers to a sound without any kind of modulation that is played as soon as, and as long as, the related movement variable exceeds a predefined threshold [74].

point of view, movement sonification shares methods and technologies with communities[3] focusing on music and performing arts where bodily movements and gestures represent an important theme, but not movement learning [5].

In sports, several activities have been sonified, such as rowing [24, 58] and cycling [45] among others. Schmitz et al. sonified the pressure of water against the hands in breast-stroke swimming and showed that sonification improves activity performance [71]. Movement sonification has also been experimented in clinical settings that involve physical therapy [34]. For example, post-stroke patients can practice simple movements with sonic feedback, such as reaching for an object with their impaired arm [61, 63]. Parkinson's patients find it easier to coordinate movements in rhythmic tasks such as walking with music [62, 67]. Nikmaram et al. [57] also pointed out the possible motivational aspect of sonification for patients. The use of sonification has also been found beneficial for chronic pain patients [53], for dysgraphia rehabilitation [18], and more generally for facilitating handwriting [16, 17].

Although movement sonification has gained growing interest in recent years, its effects on movement learning remains underexplored [74] compared to the use of visual feedback alone. In particular, its efficacy varies significantly on the users' background and skills. In sports, movement sonification is often designed to improve elite or high performance of already well-trained athletes [66]. Rehabilitation programs use rhythmic auditory cues as a means to enhance auditory-motor synchronization and promote sustained functional changes to movement [50, 81]. Hummel et al. explored how experts and novices could benefit from sonified movement on the German wheel task and speculated that novices could not benefit from auditory feedback as they have no idea of the correct movement sonification [37]. The interpretability and effectiveness of sonification also depend on the users, particularly with respect to age, gender, skill level and music abilities [26]. However, these differences have not been systematically evaluated.

2.3 Complex Gesture Learning

The effects of feedback depend on movement complexity. Similar to [74], we use the definition from Wulf and Shea:"We will judge tasks to be complex if they generally cannot be mastered in a single session, have several degrees of freedom, and perhaps tend to be ecologically valid. Tasks will be judged as simple if they have only one degree of freedom, can be mastered in a single practice session, and appear to be artificial" [87, p.186]. For instance, Fitts' aimed movement [28] is a simple task as it involves only one degree of freedom. So are many rehabilitation tasks where patients move their upper limb along one axis. More complex movements include 3D gesture [30], karate [88], ski carving [39] and rowing [64], all of which have been sonified (often using simple movement-to-sound mappings such as audio energy and pitch [23]) but have rarely been evaluated for multi-session learning.

Wulf and Shea [87] show many examples to illustrate that one should be careful in transferring results found in studies on simple tasks to complex tasks, with primarily visual feedback. For instance, concurrent visual feedback was shown to be rather unfavorable to learning simple motor tasks (e.g. simple aiming movements [77]) as it leads to a dependency on the feedback. However, it is generally considered to be useful for more complex tasks as it reduces cognitive load at the early learning stage and attracts an external focus of attention [87]. In general, the majority of movement learning studies focus on visual, haptic and vestibular sensory inputs and rarely include the auditory modality [5]. One exception is Francoise et al. [30], who compared 3d gesture learning with and without auditory feedback among 12 participants (no background screening). With a preliminary experiment using four rather simple gestures (e.g. up and down movement, a circle, etc.), they showed that auditory feedback reduces motor variability in comparison to no auditory feedback. Nevertheless, the effects of movement sonification on complex movement learning remain little explored.

[3]For example, see the NIME community (New Interfaces for Musical Expression): https://www.nime.org.

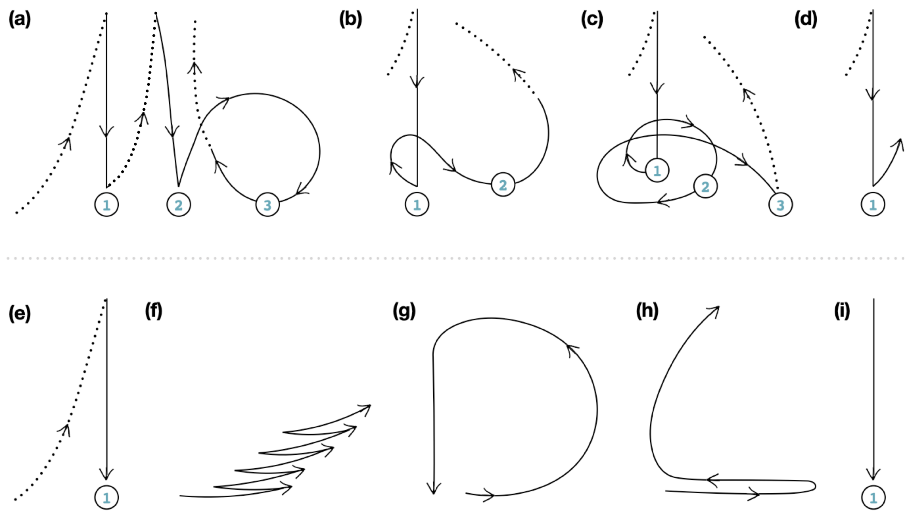

Fig. 1. The two gestures used in the study. (a)–(d): gesture 1, which is more rhythmic and (e)–(i): gesture 2, which is performed at constant speed. Dashed line represents gesture preparation.

2.4 Summary

Previous literature provides little understanding of how motor variability evolves in complex gesture learning, in particular how it is modulated by movement sonification and how it is affected by user's musical background. We argue that it is important to evaluate these factors given the growing interest in more expressive gestural interaction in music, sports and rehabilitation.

3 GESTURE AND SOUND DESIGN

We design two complex gestures as well as their respective relationships with sound, which will be used in the study reported in Section 4.

3.1 Gesture Design

The two gestures have been designed to exhibit very different characteristics. They were created with the help of the second author, who is a professional composer. These gestural phrases are similar to 'musical gestures' with typically impulsive, sustained and iterative actions [38]. Precisely, they are composed from such specific short units that, once combined, must be performed as a single continuous long gestural phrase. Gesture 1 (9-second long) is more rhythmic with a variety of strokes and specific spatial patterns, similar to conducting gestures. The vertical strokes are contrasted and followed by horizontal inwards and outwards movements, reminiscent of an "infinity" gesture (Fig. 1 (a)–(d)). By contrast, gesture 2 (16-second long) is slow, continuous and smooth. It requires users to control the movement speed (Fig. 1 (e)–(i)). We iterated on gesture 2 after a pilot study with two participants and added a percussion movement at the beginning as well as at the end to clearly mark the duration of the gesture. Note that both gestures are performed symmetrically with both hands. Fig. 1 shows only the movement of the right arm.

3.2 Sound Design

We chose a sonification approach based on different action-sound metaphors that are generally understandable by a wide audience [8, 13]. Note that the goal of our study is to explore how sonification *modulates* motor variability in movement learning, rather than *helps* the learning process. Therefore, the sound is not designed nor tested to reinforce movement learning.

We used two different approaches to associate these two gestures with sounds in order to further differentiate how they will be perceived. The sonification of gesture 1 is based on an approach called "action-sound" metaphor that has been described in [8]. This takes advantage of relationships we acquire in our everyday interaction with objects. For example, percussion sounds are typically associated with actions such as hitting or tapping, while continuous friction sound can be associated with actions such as scrubbing. In gesture 1, the gesture strokes trigger percussive sounds (samples of the Berimbau instrument) whose audio energy varies with the stroke intensity, defined as the norm of the three-dimensional acceleration[4]. The continuous movements are associated with a rain-stick sound whose audio energy varies with velocity, and whose pitch increases with hand distance to the body in order to emphasize the inward/outward movement. These associations are common in physical production of sounds. For example, audio energy varies with velocity friction such as bowing and the pitch variation is associated with the vertical or horizontal axes [41]. The sound sequence directly follows the gesture sequences. This small 'composition', was directly inspired by spontaneous vocalizations used to describe the gesture sequence, which is a method that has been proposed to initiate the sonification process [41].

Purposely contrasting gesture 1, the sonification of gesture 2 is based on an "action-music" metaphor. It is built on a melody, using the metaphor of playing a musical box – faster movement results in faster music. Performing the continuous circle gestures controls the progress of the melody by controlling the playback of the soundfile. The initial and final gesture strokes trigger a specific chord from the musical box.

We recorded the performance (movement and sound) of the composer as a baseline, which we call "template" gestures. We then use the instruction videos (accompanying material) as reference for participants to learn the two gestures.

4 EXPERIMENT

Our goal is to investigate how complex gesture learning is modulated by movement sonification and how users' musical abilities contribute to it. Since learning takes place over time [40], we designed an experiment over three days in one week: Monday, Wednesday and Friday, differing from previous studies where participants complete a series of trials in a single day.

We examine how motor variability evolves over the three days, according to the presence or absence of auditory feedback. Since musicians are trained to reproduce movement and to generate sound through movement, we compare trained musicians with non-musicians for this task. We formulate the following three hypotheses:

H1: Motor variability decreases over the three days.

H2: Movement sonification reduces motor variability.

H3: Movement sonification induces higher motor variability for non-musicians than for musicians.

The first hypothesis builds on years of research in movement science that motor variability reduces over the course of learning (e.g. [14, 73]). The second hypothesis is based on the limited literature for rather simple tasks (e.g. [30, 37]) that auditory feedback reduces motor variability. Another interesting aspect is related to *attentional focus* [86]. Research on focus of attention has consistently demonstrated that an external focus (i.e. on the movement effect) enhances motor performance and learning relative to an internal focus (i.e. on body movements). Since musicians develop a unique skill to both pay attention to their gesture and consciously listen to changes in the sound, the effect of their movement, we hypothesize that auditory feedback modulates motor

[4]In this article, we use the word *dynamics* to refer to this movement feature.

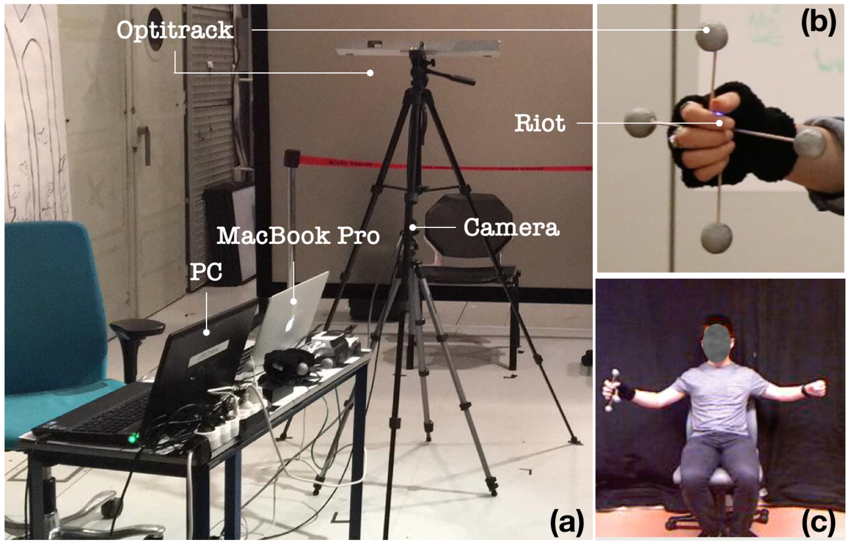

Fig. 2. (a) Experiment setup includes an Optitrack motion capture sensor and PC, a camera and a MacBook Pro that records audio, video and motion capture data; (b) the glove that participants wear, with 4 markers for tracking and a Bitalino-R-IoT sensor; (c) a participant wearing the glove and sitting on a chair in front of the camera and Optitrack sensors.

variability better among musicians than among non-musicians. Note that we do not make hypotheses on the performance of musicians vs. non-musicians when auditory feedback is absent.

4.1 Participants

Twenty-four participants (15 women, 9 men), age 18 to 32 (mean = 25, σ = 3) were recruited from public announcements at a university and at a music school. In the *muscians* group, twelve play at least one acoustic instrument regularly and have an average of 15 years (σ = 2.7) of practice. Five play the piano, two play the violin, and one each plays the harp, guitar, saxophone, flute and contrabass. We acknowledge that different instruments might affect how they learn movement, but compared to non-musicians they are all highly trained in mapping movement to sound and vice versa. For example, previous research [33] demonstrates that piano playing involves complex sensorimotor learning. In the *non-musician* group, none of the twelve participants have experience with a musical instrument [5] nor in dance. At the end of the experiment, all participants received 30 euros as compensation. The study is approved by ethics committee Insead-Sorbonne University under the protocol ID: June 2019/4 - Ref 201905.

4.2 Apparatus

Participants were seated on a chair wearing a custom-made glove in front of an Optitrack V120 Trio motion capture sensor [6] and a video camera (Fig. 2(a)). Four markers, positioned 90 degrees apart on the glove to avoid occlusion, track the participant's 3D hand position, and a Bitalino-R-IoT 3D inertial motion unit [7] was attached to the glove to track acceleration, angular velocity and orientation (Fig. 2(b)). The Optitrack camera was connected

[5] One participant learned drums for one and a half year but she has not played for 10 years.
[6] https://optitrack.com/products/v120-trio/
[7] https://www.stms-lab.fr/shop/product/r-iot/

No.	*Question*
Q1	My gesture is close to that of the video
Q2	My gesture is similar to that of the previous block
Q3	I feel like the gesture is easier to perform across blocks
Q4	My gesture is similar to that of the previous day

Table 1. 5-point Likert-scale questionnaire used in the experiment. For each question, 1 means strongly disagree and 5 means strongly agree.

to a PC that streamed data to a MacBook Pro computer (15-inch, 2.5GHz, MacOS 10.14) via a VRPN client. A custom-made program implemented with Cycling 74 Max/MSP was used to record movement data, audio and video. In the condition with sonic feedback, a loudspeaker was placed in front of the chair to provide real-time movement sonification. Movement data was sampled at 100Hz, video was captured at 30fps and audio was recorded at 48kHz, stored in separate files. Participants watched instruction videos and filled out questionnaires on an iPad.

4.3 Procedure

We used a $[2 \times 2]$ between-participants design with two independent variables: USER GROUP (musicians and non-musicians) and FEEDBACK (with or without auditory feedback). Each USER GROUP was split in two sub-groups to assign the FEEDBACK condition. We used the two gestures described in Section 3. Half the participants started with gesture 1 and half started with gesture 2 (Fig. 1).

The experiment consisted of 3 sessions on Monday, Wednesday and Friday respectively. All participants started on Monday and finished on Friday. In the first session, participants received information about the equipment and the procedure and were instructed to imitate the gestures in terms of shape, dynamics and timing. In each session, they followed the procedure below for each gesture:

(1) Watch the video for the first time for 3 minutes at most;
(2) Block 1: Practice the gesture 5 times;
(3) Watch the video 5 times at most;
(4) Block 2: Practice the gesture 5 times;
(5) Watch the video 5 times at most;
(6) Block 3: Practice the gesture 5 times.

In the condition without auditory feedback, the instruction video did not have sound and participants completed the 3 sessions without any feedback. In the condition with auditory feedback, the video had sound and participants heard the sound generated by their movement in real time. In both conditions, participants were allowed to scroll the video back and forth in step 1 and were free to imitate the gesture while watching the video in each step. They were also allowed to skip the video if they felt that they already mastered the gestures. After each block, participants filled out a Likert-scale questionnaire to provide qualitative feedback on their performance relatively to the video and to the previous block. Table 1 lists the questions and Fig. 3 summarizes the procedure.

After the 3 sessions, we conducted an interview with participants about their learning strategy and experience. Session 1 and 2 lasted about 30 minutes and session 3 lasted about 45 minutes.

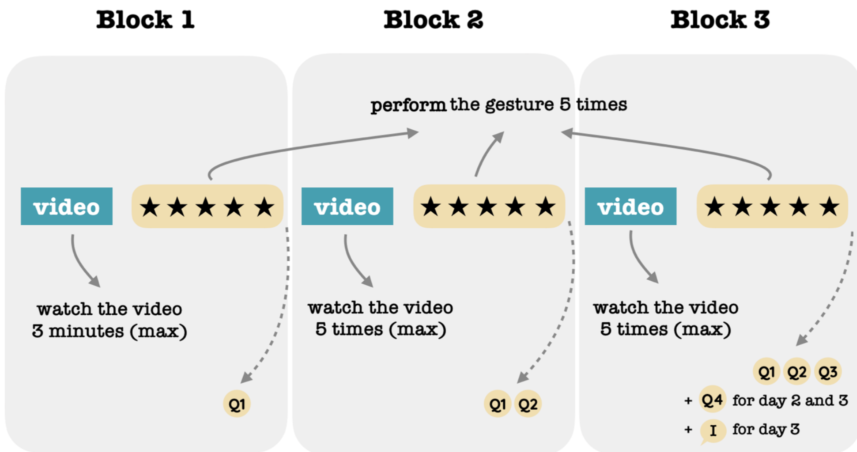

Fig. 3. The procedure for each session includes 3 blocks where participants iterate on watching instruction video and performing gestures. Q1-4 correspond to the questions in Table 1 and *I* represents the interview we conducted on day 3.

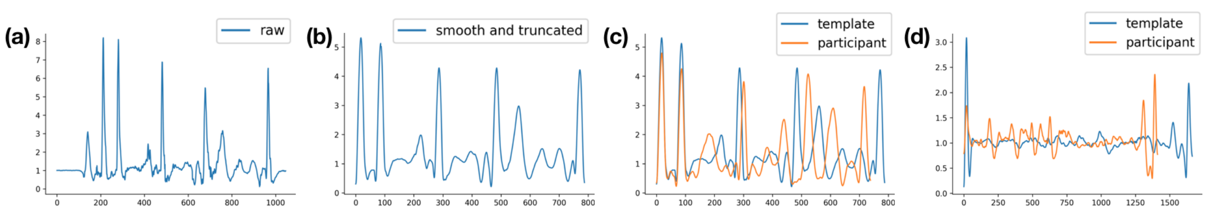

Fig. 4. (a) Gesture dynamics as a function of time for template gesture 1; (b) smoothed and truncated data of the same sample; (c) comparison between template and participant P5 on gesture 1; (4) comparison between template and participant P14 on gesture 2 after smoothing and truncating.

4.4 Data Collection

For each trial, the program collected the movement data including 3D position, acceleration, angular velocity and orientation, as well as video and audio, if any. We collected 3 Blocks × 5 Replications × 2 Gestures × 3 Sessions × 24 Participants = 2160 trials. The data was collected at the Insead-Sorbonne University Behavioral Lab.

5 RESULTS

5.1 Pre-Processing

We first pre-processed collected data before running statistical analysis. Since sampled data was noisy (Fig. 4 (a)), we first applied a Savitzky-Golay filter [65] (window size 31 and polyorder 3) to smooth it. In addition, sampled data inevitably included hand movements before and after the trial that were not part of the gesture, so we truncated the sampled data using the first and the last percussion movement as reference (Fig. 4 (b)), calculated as *dynamics* using the norm of the three-dimensional acceleration data from R-IoT. Algorithm 1 (see Appendix) describes the steps to locate the true start and end of a gesture. Fig. 4 (c) and (d) show a comparison between the template data and a participant's data after smoothing and truncating for gesture 1 (c) and 2 (d).

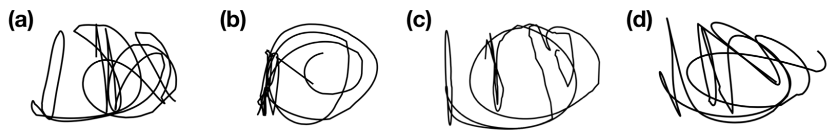

Fig. 5. (a) Trajectory of template gesture 1; (b)–(d) trajectory from 3 participants on gesture 1 illustrating the gesture appropriation by each participant.

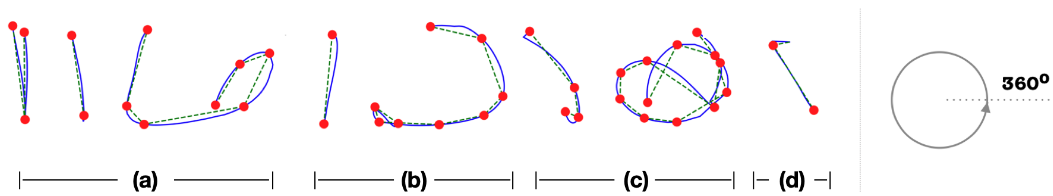

Fig. 6. Describing template gesture 1 using the Ramer-Douglas-Peucker algorithm with 32 states regardless of position. (a)–(d) correspond to the sequence in Fig. 1.

5.2 Dependent Measure

We focus our analyses on one of the three dimensions that we gave participants in the instructions – shape, which is a representative feature that captures motor variability in the learning process in our case. It describes the trajectory in space and demonstrates whether participants follow the gesture sequences, in other words, how well they imitate the template phrase from videos. We are aware that other metrics could be used such as the change of speed, and we plan to test them in the followup studies. We release collected rich dataset online [8].

To compute variability in trajectory, we need to find an appropriate approach that takes into account the unique characteristics of the complex gestures we used in the experiment, which challenges traditional methods. First, we observe that participants appropriate gestures based on their perception of the instruction video and their movement style, thus making the overall shape look very different from the template (Fig. 5). Second, after considering different approaches based on manual annotation or recognition of position data, we eventually chose to transform the gesture trajectory into directions, measured in degrees. This gives us the advantage of checking whether participants follow the structure of the sequence, so that, for example, the position of the percussive movement or the size of the circle do not matter (Fig. 1). We transform the trajectory using the Ramer-Douglas-Peucker algorithm [21, 60] and describe each trial using directions. Fig. 6 shows the segmentation of template gesture 1, described using 32 states regardless of position: [276°, 90°] hands down hands up, [274°] hands down, [75°, 37°, 259°, 213°, 178°, 85°] circle inward, [267°] hands down, [129°, 277°, 359°, 42°, 97°, 126°, 178°] circle outward, etc.

This approach has two advantages: first, we need not identify the position nor segment the gesture sequence; second, we can tell whether a participant follows the template by comparing two vectors. We compared different methods for calculating the difference between two vectors including dynamic time warping [4], Euclidean distance [15], and Hamming distance [35], Levenshtein distance [44] and Jaro-Winkler distance [83] after converting the directions from degrees to strings. We finally chose the Euclidean distance as a measure quantifying typical changes we observe in our dataset: for a given participant, gesture elements are either added or missing

[8] https://nubo.ircam.fr/index.php/s/p3sr7rMfDwRr9mb

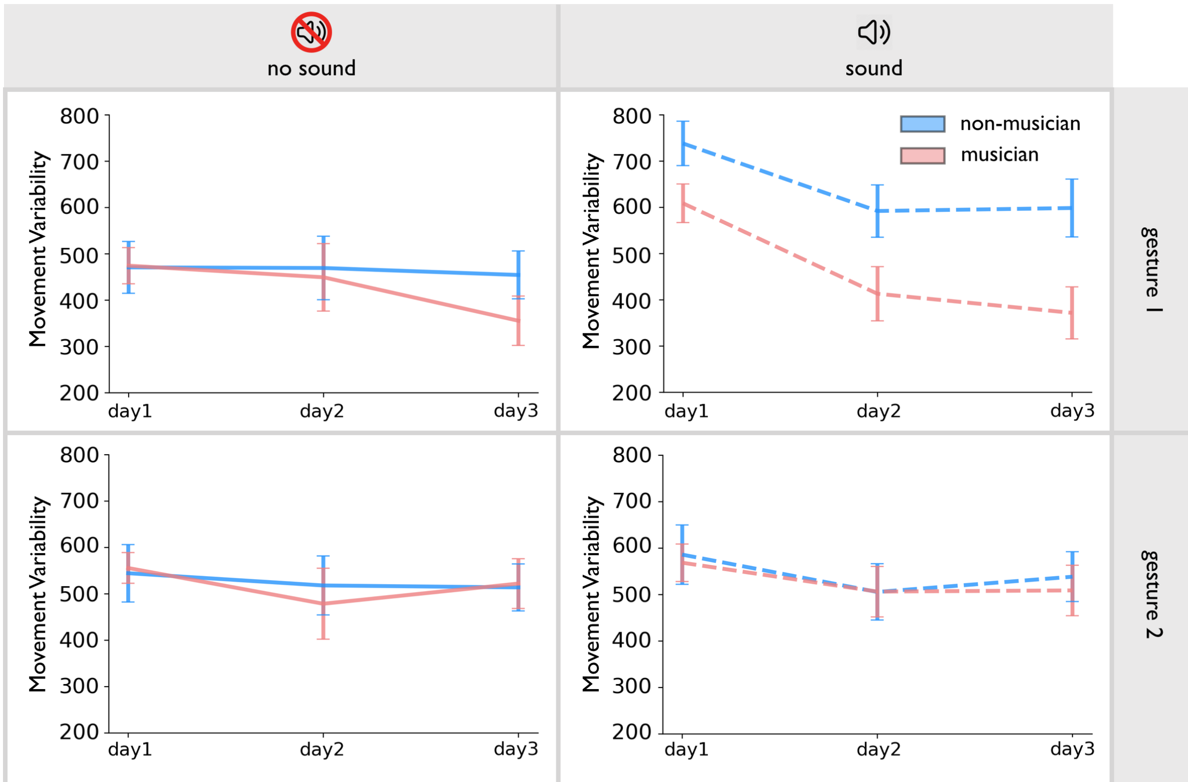

Fig. 7. An overview of movement variability for musicians and non-musicians with or without auditory feedback over 3 days for gesture 1&2 (with 95% confidence intervals).

between successive trials. In all figures, the reported variability corresponds to cumulative differences in direction (in degrees) over the whole shape, between two consecutive trials. This variability typically diminishes until the shape is performed in a consistent manner, independent of whether it is close to the template or not (See Appendix for a more detailed description).

5.3 Statistical Results

We first performed a Shapiro-Wilk normality test and found that data is not normally distributed ($W = 0.99; p < 0.0001$), therefore we rule out parametric ANOVA tests. Instead, we performed a generalized linear mixed model with repeated measures (Type III Wald χ^2 tests) to examine the effects of FEEDBACK, USER GROUP and DAY on movement variability. Fig. 7 shows an overview of variability for musicians and non-musicians with or without auditory feedback over 3 days for gesture 1 and 2. We note strong interaction effects for the independent variables, as well as different effects on gesture 1 and 2, therefore we separate the analysis between gesture 1 and gesture 2.

5.3.1 Intra-individual Variability for Gesture 1. Table 2 shows analysis of Deviance Table (Type III Wald χ^2 tests) for gesture 1. We found significant effect of FEEDBACK, as well as two interaction effects: FEEDBACK × DAY and USER GROUP × DAY. Since the two interaction effects are present and strong, we further analyze these two interaction effects, rather the main effect FEEDBACK.

Factors	*Chisq*	*Df*	*Pr(>Chisq)*	
FEEDBACK	27.7358	1	< 0.0001	***
USER GROUP	0.0052	1	$= 0.9426$	
DAY	0.8955	2	$= 0.6391$	
FEEDBACK × USER GROUP	3.4318	1	$= 0.0639$	
FEEDBACK × DAY	32.6792	2	< 0.0001	***
USER GROUP × DAY	15.4246	2	$= 0.0004$	***
FEEDBACK × USER GROUP × DAY	0.7418	2	$= 0.6901$	

Table 2. Analysis of Deviance Table (Type III Wald χ^2 tests) for gesture 1. Main effect FEEDBACK and interaction effects: FEEDBACK × DAY and USER GROUP × DAY are significant. ***: 0.0001; **: 0.001; *: 0.01. The same significance encoding is used throughout the paper.

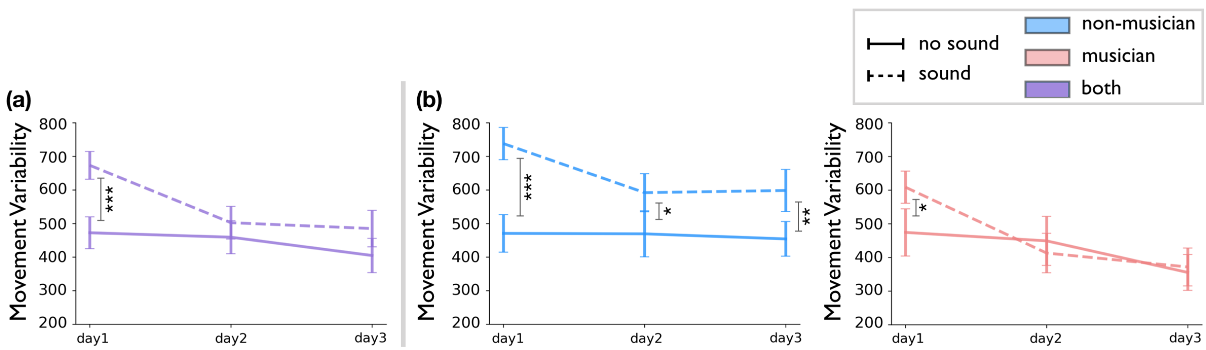

Fig. 8. Interaction effects FEEDBACK × DAY of gesture 1 for (a) both groups and (b) non-musicians (left) and musicians (right) with 95% conference intervals.

Least-Squares Means (LSM) analysis shows that auditory feedback induces significantly more variability for participants on day 1 ($Df = 24.5, t.ratio = -5.59, p < 0.0001$), but not on day 2 or day 3 (Fig. 8a). Further investigating the differences between musicians and non-musicians (Fig. 8b), LSM analysis demonstrates that for non-musicians, auditory feedback induces significantly more movement variability for all 3 days: day 1 ($Df = 24.5, t.ratio = -5.27, p < 0.0001$), day 2 ($Df = 24.5, t.ratio = -2.42, p = 0.0235$) and day 3 ($Df = 24.5, t.ratio = -2.84, p = 0.0089$). However, it only has a significant effect on musicians on day 1 ($Df = 24.5, t.ratio = -2.65, p = 0.0140$), but not on day 2 or day 3.

Fig. 9 shows the interaction effects of USER GROUP × DAY. LSM analysis shows that musicians and non-musicians exhibit similar movement variability on day 1 but musicians reduce this variability significantly more than non-musicians on day 2 ($Df = 24.5, t.ratio = 2.77, p = 0.0105$) and day 3 ($Df = 24.5, t.ratio = 4.53, p = 0.0001$) (Fig. 9a). Further separating the FEEDBACK conditions (Fig. 9b), we find that musicians and non-musicians have similar variability across 3 days when auditory feedback is not present. However, when there is auditory feedback, musicians demonstrate significantly lower variability compared to non-musicians for all 3 days: day 1 ($Df = 24.5, t.ratio = 2.55, p = 0.0175$), day 2 ($Df = 24.5, t.ratio = 3.53, p = 0.0017$) and day 3 ($Df = 24.5, t.ratio = 4.46, p = 0.0002$).

In summary, musicians and non-musicians demonstrate similar motor variability across 3 days when auditory feedback is not present. When auditory feedback is present, it induces significantly higher motor variability

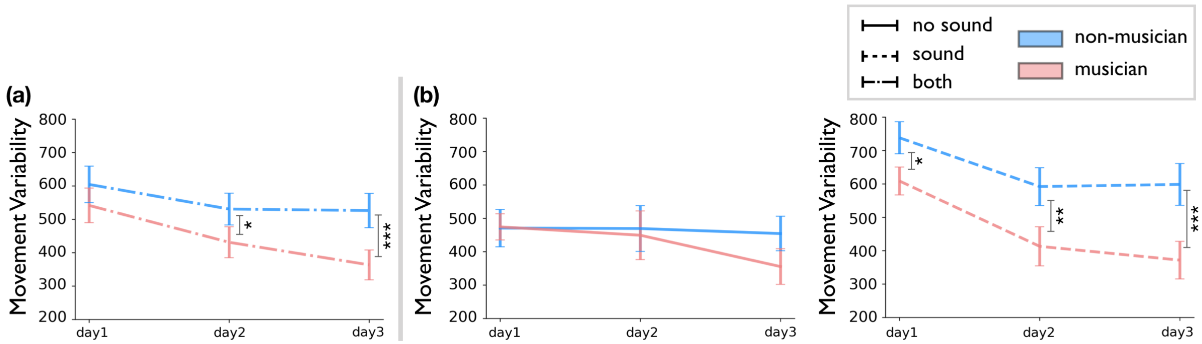

Fig. 9. Interaction effects USER GROUP × DAY of gesture 1 for (a) both feedback conditions and (b) no sound (left) and sound (right) with 95% conference intervals.

Factors	*Chisq*	*Df*	*Pr(>Chisq)*
FEEDBACK	0.7905	1	= 0.3739
USER GROUP	0.0578	1	= 0.8101
DAY	3.2278	2	= 0.1991
FEEDBACK × USER GROUP	0.1881	1	= 0.6645
FEEDBACK × DAY	4.4704	2	= 0.1069
USER GROUP × DAY	4.7356	2	= 0.0936
FEEDBACK × USER GROUP × DAY	5.3263	2	= 0.0697

Table 3. Analysis of Deviance Table (Type III Wald χ^2 tests) for gesture 2. No significant effects are found.

for both groups on day 1. While musicians reduce this variability to a similar level than no auditory feedback condition on day 2 and day 3, non-musicians remain to have significantly higher variability. Therefore, we reject **H2** that movement sonification reduces motor variability. Across 3 days, musicians also have significantly lower variability than non-musicians, validating **H3**. **H1** (motor variability decreases over three days) is only partially validated between day 1 and day 2 for both musicians and non-musicians when auditory feedback is present. The 3-session experiment enables us to examine how these effects unfold over time.

5.3.2 Intra-individual Variability for Gesture 2. Unlike gesture 1, independent variables FEEDBACK, USER GROUP and DAY seem to have no effects on variability for gesture 2. Table 3 shows analysis of Deviance Table (Type III Wald χ^2 tests). We find no significant effects. Auditory feedback does not change motor variability compared to no auditory condition. Musicians and non-musicians perform equally well regardless of the presence of feedback over 3 days (Fig. 10). Therefore, we reject all three hypotheses.

5.4 Subjective Feedback

Regarding subjective feedback, Table 1 shows the results of question 1 comparing participants' gestures to that of the video. Additional results can be found in the accompanying material. Overall, participants are more confident when there is no auditory feedback than when there is (average 3.7 vs. 3.2, $p < 0.0001$). Participants also give higher scores for gesture 2 than gesture 1 (average 3.7 vs. 3.1, $p < 0.0001$). The two groups (non-musician and musician) give similar scores (average 3.5 vs. 3.4) (Fig. 11).

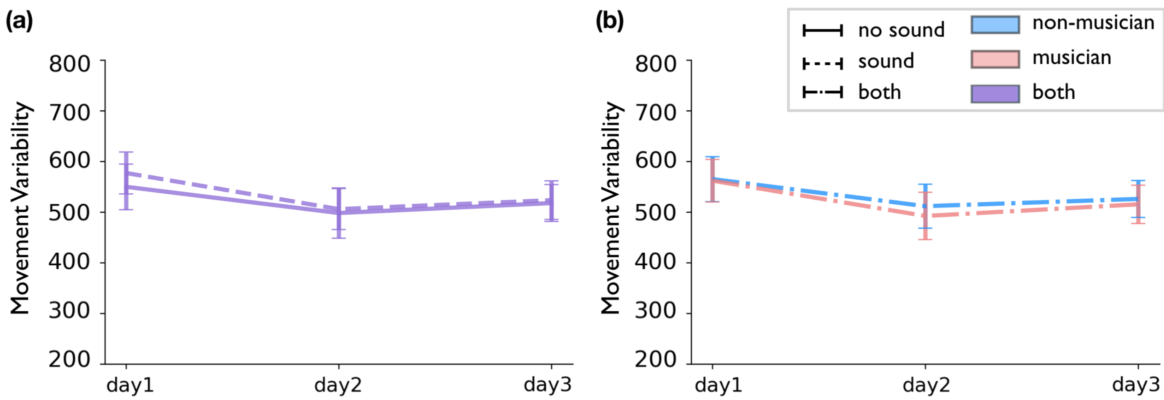

Fig. 10. Interaction effects of gesture 2: (a) FEEDBACK × DAY and (b) USER GROUP × DAY. There are no significant effects.

Auditory feedback: Overall, participants find it more difficult to learn the gestures with auditory feedback, particularly on the first day as "*there is a lot of information to process*", "*I'm not sure what should I focus on*", "*it is difficult to understand why my sound is different from the one in the video*" and "*very frustrating*". This is consistent with the observed significant effects of FEEDBACK on day 1 for gesture 1. However, it is important to note that they also appreciate auditory feedback once they have learned the gestures: "*it helps tremendously to guide and remind the movement*", "*I know my movement is not correct if I don't hear the right sound*" and "*I think I'll do less well if there is no sound*". While the statistical results support only the effects of the sonification on gesture 1, most participants feel sonification as motivating or helpful even for gesture 2.

Gestures: Regarding the gestures themselves, all participants apart from two musicians agree that gesture 2 is much easier as "*it's repetitive and slow so you have time to see what's happening*" and "*the elements are very different from each other*". They also liked the sound as "*it's soft and musical*". By contrast, gesture 1 is considered more difficult as "*the elements are similar but not the same*", "*the orientations of the circles are very confusing*" and "*it's very fast*". Interestingly, while the sound feedback for gesture 1 is more direct than for gesture 2, such as the percussive part, participants appreciated it less than the sound of gesture 2: "*I found it (the percussion sound) aggressive*", "*the sound is not so intuitive and provides less feedback*". This is surprising since gesture 1 is designed toward creating a higher sound interaction agency. It also triggered emotional responses such as "*it's very intense and stressful*". In general, perception depends on the participants background. For examples, two musicians mentioned that "*gesture 2 is definitely more difficult as in music, everything that's slow and requires you to keep something constant is difficult*".

Musical background: Even though musicians and non-musicians give similar self-evaluation scores, musicians demonstrate significantly less variability than non-musicians for gesture 1 when auditory feedback is present. This might be because non-musicians tend to focus on the movement in the video and try to find the correct sound. On the other hand, musicians tend to focus on the sound and are motivated to perfect it by perfecting their movements. This is related to the notion of *attentional focus* [86]: research on focus of attention has consistently demonstrated that an external focus (i.e., on the movement effect, here the auditory feedback) enhances motor performance and learning relative to an internal focus (i.e., on the gestures themselves). Interestingly enough, this effect seems to be also interlaced with gestures themselves as well as the sound design: musicians and non-musicians perform equally well for gesture 2 regardless of auditory feedback.

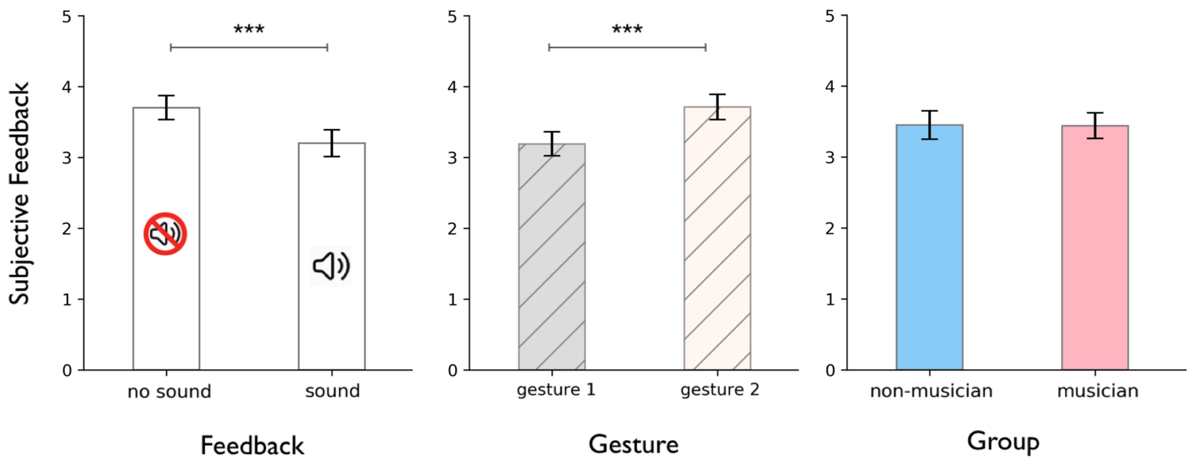

Fig. 11. Subjective feedback with 95% confidence intervals (higher is better) for question 1: My gesture is close to that of the video.

5.5 Summary

In summary, how movement sonification and musical background affect motor variability in complex gesture learning seems to be much more complex than one would expect. For gesture 1, a fast and dynamic gesture similar to conducting gestures with a direct "sound-action" sonification approach, auditory feedback induces significantly higher motor variability for both musicians and non-musicians on day 1. While musicians reduce the variability on day 2 and 3, non-musicians remain significantly more variable compared to their counterparts in the condition without auditory feedback. Across 3 days, musicians also have significantly lower variability than non-musicians. For gesture 2, a slow and smooth gesture where sonification follows an "action-music" metaphor, there are virtually no effects of movement sonification and musical background on variability. Therefore, motor variability does not necessarily decrease over 3 days; it depends on the gesture complexity, the feedback mechanism as well as user backgrounds. These observations can be explained by the qualitative feedback that participants generally find it more difficult to learn the gestures with auditory feedback; gesture 2 is considered significantly easier than gesture 1 in terms of the movement, the sound and their relationship; and musicians focus on producing the correct sound by perfecting their movements whereas non-musicians try to imitate the movements and check whether they are correct by verifying the sound outcome. Furthermore, the appreciation of sonifcation can vary significantly among users, from negative to positive aspects, which might or might not be directly linked to quantitative changes.

6 DISCUSSIONS

Our main contribution is the empirical study that takes into account the important factors that have not been systematically studied in the literature: the effects of movement sonification and users' musical background on motor variability in complex gesture learning. Using two gestures that have been designed to exhibit different characteristics, we illustrate the interlaced interaction effects of these factors and how they unfold in the 3-day learning process. In this section, we summarize relevant outcomes as a recommendation list for future work. We also outline the limitation of our current study and directions for future research.

6.1 Lessons Learned

6.1.1 Participants' musical background should be taken into account in studies that involve movement sonification. While prior study has investigated the differences between musicians and non-musicians in general gesture imitation tasks, for instance, Spilka et al. [75] showed that musicians were able to imitate American Sign Language more accurately than non-musicians, to the best of our knowledge, our study offers some first results on the differences between these groups in learning complex gestures with or without auditory feedback. We show that, in particular for gesture 1, musicians and non-musicians demonstrate significantly different levels of variability across 3 days when auditory feedback is present. Qualitative results also highlight that musicians, with years of musical training that have strengthened tehir perceptual and motor action representations, adopt a very different approach in learning gestures with sound – they focus on producing the correct sound by perfecting their movements. Non-musicians tend to imitate the movements and verify whether the sound output is correct. Future studies should consider the fundamental differences between these two groups when recruiting participants in movement sonification studies.

6.1.2 Longitudinal study is important to examine the effects of sonification on complex gestures. Different from previous studies where participants complete a series of trials in a single day, we purposefully designed an experiment that lasts for 3 days, taking into consideration that learning takes place over time, especially learning complex gestures. While Hummel et al. speculate that "novices could not benefit from auditory feedback as they have no idea of the correct movement sonification" [37], our findings suggest that movement sonification is difficult at the beginning as participants need to focus on different information sources as well as to understand their relationships. Musicians seem to better cope with this, reducing motor variability to a similar level as their counterparts in the condition without auditory feedback. Non-musicians remain variable across 3 days. The 3-day experiment makes it possible to see how the effects of sonification unfold over time. For future studies that involve complex gesture learning, we therefore recommend longitudinal studies to better investigate the effects of movement sonification.

6.1.3 The effects of sonification and musical background vary given a specific design of gesture and sound. It is striking that participants react differently to the two gestures, both quantitatively and qualitatively. Gesture 1 is considered more difficult because of its speed and similar yet different elements, whereas gesture 2 is slow with distinct gesture segments. Participants also appreciate the music of gesture 2. By contrast, sonification of gesture 1 is considered more aggressive by some participants and less informative, while users had more control on the sound in this case than in case of gesture 2. This highlights that the perceived 'agency', as well as the adherence to the sound/musical metaphor, might be more important than the actual effects of the gesture to the sounds. Therefore, the effects of sonification and musical background vary given a specific design of gesture and sound. Indeed, we did not explicitly study the choice of gesture-sound combination and it is possible that it affects the perceived difficulty. For instance, Lemaitre et al. [41] and Tajadura et al. [78] point out the effects of auditory feedback on emotional responses; Tajadura et al. [79] also show that our body perception can be altered by sonification; and Dyer et al. [25] demonstrate the advantages of melodic over rhythmic movement sonification in bimanual coordination performance. Future work should further explore how different sounds and different movement–sound mappings affect perception and motor variability in complex gesture learning.

6.2 Limitations

6.2.1 Motor variability metrics. Our study relies on a specific chosen method that computes motor variability based on shape. This constitutes one of the limitations of the currently reported results, reflecting only specific features of what movement variability could encompass. As mentioned earlier, we chose the Euclidean distance because of its sensitive accounting for added or missing gesture elements, but it could not adequately quantify small variations that occur during learning. Moreover, we do not consider the motor variability involved in expressivity, typically found among experienced musicians. We plan to examine such aspects of movement variability in followup studies, and explore other variability metrics.

6.2.2 Number of participants and number of gestures. We recruited 12 musicians and 12 non-musicians to learn two complex gestures that are designed to exhibit very different characteristics over 3 days. In total, we collected 2160 gesture profiles, which we release online as open source data, as well as 24 interview transcriptions. While there are only 6 participants in each FEEDBACK × USER GROUP condition, we argue that the current study offers a useful first step in exploring the effects of movement sonification and users' musical background on motor variability in complex gesture learning. Furthermore, the two gestures we designed and used should not be taken as a generalization of all complex gestures – gesture 1 is fast and rhythmic while gesture 2 is slow and smooth. A larger-scale study is needed to explore other gestures as well as their respective sound designs.

7 CONCLUSION

We conducted what we believe is the first study that systematically evaluates the effects of movement sonification and users' musical background on motor variability in complex gesture learning over three sessions. We contribute an experimental protocol and a user study comparing musicians and non-musicians under two conditions, with and without auditory feedback.

The results indicate that these effects are much more complex than generally acknowledged, interlacing users' capability in perceptual-motor decoding, gesture and sound design, as well as which stages they are at in the learning process. The 3-day experiment made evident the differences between musicians and non-musicians: For gesture 1, auditory feedback induces higher motor variability for both groups. While musicians reduce the variability on day 2 and 3, non-musicians remain to have high variability. For gesture 2, there are virtually no effects of these factors with the metrics we proposed. This can be explained by the perceived difficulty of auditory feedback, perceived difficulty of gestures, as well as the two user groups' different learning strategy.

Future work includes examining the effects of movement sonification using other variability metrics; conducting a larger scale study with more participants of different backgrounds, such as dancers; and explore more gesture&sound designs and their effects on motor variability. To this end, we release the dataset, including gesture and audio recordings, and invite interdisciplinary researchers to further advance the use of pleasant and motivational sound in movement-based interactive systems.

ACKNOWLEDGMENTS

We thank the reviewers and our colleagues at Ircam and Exsitu for their helpful comments, and the participants in our studies for their time. This research was partially funded by ELEMENT – Enabling Learnability in Embodied Movement Interaction (ANR-18-CE33-0002), by European Research Council (ERC) grants n° 695464 "ONE: Unified Principles of Interaction" and n° 321135 "CREATIV: Creating Co-Adaptive Human-Computer Partnerships", and by Insead-Sorbonne University Behavioral Lab.

REFERENCES

[1] Jack A Adams. 1971. A closed-loop theory of motor learning. *Journal of motor behavior* 3, 2 (1971), 111–150.

[2] Sarah Fdili Alaoui, Baptiste Caramiaux, Marcos Serrano, and Frédéric Bevilacqua. 2012. Movement qualities as interaction modality. In *Proceedings of the Designing Interactive Systems Conference*. ACM, New York, NY, USA, 761–769.

[3] Jessalyn Alvina, Joseph Malloch, and Wendy Mackay. 2016. Expressive Keyboards: Enriching Gesture-Typing on Mobile Devices. In *Proceedings of the 29th ACM Symposium on User Interface Software and Technology (UIST 2016)*. ACM, ACM, Tokyo, Japan, 583 – 593. https://doi.org/10.1145/2984511.2984560

[4] Frédéric Bettens and Todor Todoroff. 2009. Real-time dtw-based gesture recognition external object for max/msp and puredata. In *Proceedings of the SMC 2009 Sound and Music Computing Conference*. SMC network, 30–35.

[5] Frédéric Bevilacqua, Eric O Boyer, Jules Françoise, Olivier Houix, Patrick Susini, Agnès Roby-Brami, and Sylvain Hanneton. 2016. Sensori-motor learning with movement sonification: perspectives from recent interdisciplinary studies. *Frontiers in neuroscience* 10 (2016), 385.

[6] Frédéric Bevilacqua, Baptiste Caramiaux, and Jules Françoise. 2016. Perspectives on Real-time Computation of Movement Coarticulation. In *Proceedings of the 3rd International Symposium on Movement and Computing (MOCO '16)*. ACM, New York, NY, USA, 35:1–35:5.

[7] Frederic Bevilacqua, Fabrice Guédy, Norbert Schnell, Emmanuel Fléty, and Nicolas Leroy. 2007. Wireless Sensor Interface and Gesture-follower for Music Pedagogy. In *Proceedings of the 7th International Conference on New Interfaces for Musical Expression (NIME '07)*. ACM, New York, NY, USA, 124–129. https://doi.org/10.1145/1279740.1279762

[8] Frédéric Bevilacqua, Norbert Schnell, Jules Françoise, Éric O. Boyer, Diemo Schwarz, and Baptiste Caramiaux. 2017. Designing action-sound metaphors using motion sensing and descriptor-based synthesis of recorded sound materials. In *The Routledge Companion to Embodied Music Interaction*, Micheline Lesaffre, Pieter-Jan Maes, and Marc Leman (Eds.). Taylor & Francis, New York, NY, USA, 391–401.

[9] Eric O Boyer, Frederic Bevilacqua, Emmanuel Guigon, Sylvain Hanneton, and Agnes Roby-Brami. 2020. Modulation of ellipses drawing by sonification. *Experimental Brain Research* 238 (2020), 1011–1024.

[10] Daniel A Braun, Ad Aertsen, Daniel M Wolpert, and Carsten Mehring. 2009. Motor task variation induces structural learning. *Curr Biol* 19, 4 (Feb 2009), 352–7.

[11] Roberto Bresin, Maurizio Mancini, Ludvig Elblaus, and Emma Frid. 2020. Sonification of the self vs. sonification of the other: Differences in the sonification of performed vs. observed simple hand movements. *International Journal of Human-Computer Studies* 144 (2020), 102500.

[12] Antonio Camurri, Corrado Canepa, Nicola Ferrari, Maurizio Mancini, Radoslaw Niewiadomski, Stefano Piana, Gualtiero Volpe, Jean-Marc Matos, Pablo Palacio, and Muriel Romero. 2016. A system to support the learning of movement qualities in dance: a case study on dynamic symmetry. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, New York, NY, USA, 973–976.

[13] Baptiste Caramiaux, Alessandro Altavilla, Scott G. Pobiner, and Atau Tanaka. 2015. Form Follows Sound: Designing Interactions from Sonic Memories. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3943–3952.

[14] 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), 18.

[15] Per-Erik Danielsson. 1980. Euclidean distance mapping. *Computer Graphics and image processing* 14, 3 (1980), 227–248.

[16] Jérémy Danna, Maureen Fontaine, Vietminh Vietminh.Paz@univ-Amu.Fr Paz-Villagrán, Charles Gondre, Etienne Thoret, Mitsuko Aramaki, Richard Kronland-Martinet, Sølvi Ystad, and Jean-Luc Velay. 2015. The effect of real-time auditory feedback on learning new characters. *Human Movement Science* 43 (2015), 216–228.

[17] Jérémy Danna and Jean-Luc Velay. 2017. Handwriting Movement Sonification: Why and How? *IEEE Transactions on Human-Machine Systems* 47, 2 (2017), 299–303.

[18] Jérémy Danna, Jean-Luc Velay, Vietminh Vietminh.Paz@univ-Amu.Fr Paz-Villagrán, Annabelle Capel, Céline Petroz, Charles Gondre, Etienne Thoret, Mitsuko Aramaki, Solvi Ystad, and Richard Kronland-Martinet. 2013. Handwriting Movement Sonification for the Rehabilitation of Dysgraphia. In *10th International Symposium on Computer Music Multidisciplinary Research (CMMR) - Sound, Music & Motion - 15 - 18 oct. 2013 - Marseille, France*. Marseille, France, 200–208.

[19] Keith Davids, Paul Glazier, Duarte Araujo, and Roger Bartlett. 2003. Movement systems as dynamical systems. *Sports medicine* 33, 4 (2003), 245–260.

[20] Ashesh K. Dhawale, Maurice A. Smith, and Bence P. Ölveczky. 2017. The Role of Variability in Motor Learning. *Annual Review of Neuroscience* 40, 1 (2017), 479–498.

[21] David H. Douglas and Thomas K. Peucker. 2011. *Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature*. John Wiley & Sons, Ltd, New York, NY, USA, Chapter 2, 15–28. https://doi.org/10.1002/9780470669488.ch2 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470669488.ch2

[22] Paul Dourish. 2004. *Where the action is: the foundations of embodied interaction*. MIT press, Cambridge, MA, USA.

[23] Gaël Dubus and Roberto Bresin. 2013. A systematic review of mapping strategies for the sonification of physical quantities. *PloS one* 8, 12 (2013), e82491.

[24] Gaël Dubus and Roberto Bresin. 2015. Exploration and evaluation of a system for interactive sonification of elite rowing. *Sports Engineering* 18, 1 (2015), 29–41.

[25] JF Dyer, P Stapleton, and MWM Rodger. 2017. Advantages of melodic over rhythmic movement sonification in bimanual motor skill learning. *Experimental brain research* 235, 10 (2017), 3129–3140.

[26] Alfred O Effenberg. 2005. Movement sonification: Effects on perception and action. *IEEE multimedia* 12, 2 (2005), 53–59.

[27] Alfred O Effenberg, Ursula Fehse, Gerd Schmitz, Bjoern Krueger, and Heinz Mechling. 2016. Movement sonification: effects on motor learning beyond rhythmic adjustments. *Frontiers in neuroscience* 10 (2016), 219.

[28] Paul M Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 47, 6 (1954), 381.

[29] Paul M Fitts and Michael I Posner. 1967. *Human performance.* Brooks/Cole, Belmont, CA, USA.

[30] Jules Françoise, Olivier Chapuis, Sylvain Hanneton, and Frédéric Bevilacqua. 2016. SoundGuides: Adapting Continuous Auditory Feedback to Users. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16).* ACM, New York, NY, USA, 2829–2836. https://doi.org/10.1145/2851581.2892420

[31] Karmen Franinović and Stefania Serafin. 2013. *Sonic Interaction Design.* MIT Press, Cambridge, MA, USA.

[32] R. I. Godøy and M. Leman (Eds.). 2009. *Musical Gestures: Sound, Movement and Meaning.* Routledge, New York, NY, USA.

[33] Rolf Inge Godøy and Marc Leman. 2010. *Musical gestures: Sound, movement, and meaning.* Routledge, New York, NY, USA.

[34] J. Guerra, L. Smith, D. Vicinanza, B. Stubbs, N. Veronese, and G. Williams. 2020. The use of sonification for physiotherapy in human movement tasks: A scoping review. *Science & Sports* 35, 3 (2020), 119 – 129.

[35] Richard W Hamming. 1950. Error detecting and error correcting codes. *The Bell system technical journal* 29, 2 (1950), 147–160.

[36] Luke S Hopper, Andries L Weidemann, and Janet Karin. 2018. The inherent movement variability underlying classical ballet technique and the expertise of a dancer. *Research in Dance Education* 19, 3 (2018), 229–239.

[37] Jessica Hummel, Thomas Hermann, Christopher Frauenberger, and Tony Stockman. 2010. Interactive sonification of german wheel sports. In *Proceedings of ISon 2010-Interactive Sonification Workshop: Human Interaction with Auditory Displays.* interactive-sonification.org, 17–22.

[38] A. R. Jensenius, M.M. Wanderley, R. I. Godøy, and M. Leman. 2009. Musical gestures: concepts and methods in research. In *Musical gestures: Sound, Movement, and Meaning.* Routledge, New York, NY, USA.

[39] Richard Kirby. 2009. Development of a real-time performance measurement and feedback system for alpine skiers. *Sports Technology* 2, 1-2 (2009), 43–52.

[40] John W Krakauer and Pietro Mazzoni. 2011. Human sensorimotor learning: adaptation, skill, and beyond. *Current opinion in neurobiology* 21, 4 (2011), 636–644.

[41] Guillaume Lemaitre, Olivier Houix, Patrick Susini, Yon Visell, and Karmen Franinović. 2012. Feelings elicited by auditory feedback from a computationally augmented artifact: The flops. *IEEE Transactions on Affective Computing* 3, 3 (2012), 335–348.

[42] Marc Leman. 2007. *Embodied Music Cognition and Mediation Technology.* MIT Press, Cambridge, Massachusetts.

[43] Marc Leman. 2016. *The expressive moment: How interaction (with music) shapes human empowerment.* MIT Press, Cambridge, MA, USA.

[44] Vladimir I Levenshtein. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics doklady* 10, 8 (1966), 707–710.

[45] Pieter-Jan Maes, Valerio Lorenzoni, and Joren Six. 2019. The SoundBike: musical sonification strategies to enhance cyclists' spontaneous synchronization to external music. *Journal on Multimodal User Interfaces* 13, 3 (2019), 155–166.

[46] Richard A Magill and David I Anderson. 2007. *Motor learning and control: Concepts and applications.* Vol. 11. McGraw-Hill, New York, NY, USA.

[47] Claire F Michaels and Claudia Carello. 1981. *Direct perception.* Prentice-Hall, Englewood Cliffs, NJ, USA.

[48] Benjamin Matuszewski Michelle Agnes Magalhaes and Frédéric Bevilacqua. 2020. *Constella(c)tions.* Technical Report. STARTS Residency Public Report.

[49] Florian Mueller and Katherine Isbister. 2014. Movement-based game guidelines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA, 2191–2200.

[50] Mauro Murgia, Federica Corona, Roberta Pili, Fabrizio Sors, Tiziano Agostini, Carlo Casula, Massimiliano Pau, and Marco Guicciardi. 2015. Rhythmic auditory stimulation (RAS) and motor rehabilitation in Parkinson's disease: new frontiers in assessment and intervention protocols. *The Open Psychology Journal* 8, (Suppl 3: M7) (2015), 220–229.

[51] Elena Nava and Ana Tajadura-Jiménez. 2020. Auditory-induced body distortions in children and adults. *Scientific reports* 10, 1 (2020), 1–14.

[52] John G Neuhoff, Rebecca Knight, and Joseph Wayand. 2002. Pitch change, sonification, and musical expertise: Which way is up?. In *Proceedings of the 2002 International Conference on Auditory Display.* Georgia Institute of Technology, Atlanta, GA, USA, Article 2, 6 pages.

[53] Joseph W. Newbold, Nadia Bianchi-Berthouze, Nicolas E. Gold, Ana Tajadura-Jiménez, and Amanda CdC Williams. 2016. Musically Informed Sonification for Chronic Pain Rehabilitation: Facilitating Progress & Avoiding Over-Doing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16).* Association for Computing Machinery, New York, NY, USA, 5698–5703.

[54] Karl M. Newell and Andrew B. Slifkin. 1998. The nature of movement variability. In *Motor behavior and human skill: A multidisciplinary approach.* Human Kinetics, Champaign, IL, USA, 143–160.

[55] Radoslaw Niewiadomski, Maurizio Mancini, Andrea Cera, Stefano Piana, Corrado Canepa, and Antonio Camurri. 2019. Does embodied training improve the recognition of mid-level expressive movement qualities sonification? *Journal on Multimodal User Interfaces* 13, 3 (2019), 191–203.

[56] Nikou Nikmaram, Daniel S. Scholz, Michael Großbach, Simone B. Schmidt, Jakob Spogis, Paolo Belardinelli, Florian Müller-Dahlhaus, Jörg Remy, Ulf Ziemann, Jens D. Rollnik, and Eckart Altenmüller. 2019. Musical Sonification of Arm Movements in Stroke Rehabilitation Yields Limited Benefits. *Frontiers in Neuroscience* 13 (2019), 1378.

[57] Nikou Nikmaram, Daniel S. Scholz, Michael Großbach, Simone B. Schmidt, Jakob Spogis, Paolo Belardinelli, Florian Müller-Dahlhaus, Jörg Remy, Ulf Ziemann, Jens D. Rollnik, and Eckart Altenmüller. 2019. Musical Sonification of Arm Movements in Stroke Rehabilitation Yields Limited Benefits. *Frontiers in Neuroscience* 13 (2019), 1378.

[58] Benjamin O'Brien, Brett Juhas, Marta Bieńkiewicz, Frank Buloup, Lionel Bringoux, and Christophe Bourdin. 2020. Online sonification for golf putting gesture: reduced variability of motor behaviour and perceptual judgement. *Experimental Brain Research* 238, 4 (2020), 883–895.

[59] Katerina El Raheb, Marina Stergiou, Akrivi Katifori, and Yannis Ioannidis. 2019. Dance interactive learning systems: A study on interaction workflow and teaching approaches. *ACM Computing Surveys (CSUR)* 52, 3 (2019), 1–37.

[60] Urs Ramer. 1972. An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing* 1, 3 (1972), 244 – 256. https://doi.org/10.1016/S0146-664X(72)80017-0

[61] P Ripollés, N Rojo, J Grau-Sánchez, JL Amengual, E Càmara, J Marco-Pallarés, M Juncadella, L Vaquero, F Rubio, E Duarte, et al. 2016. Music supported therapy promotes motor plasticity in individuals with chronic stroke. *Brain imaging and behavior* 10, 4 (2016), 1289–1307.

[62] Matthew WM Rodger and Cathy M Craig. 2016. Beyond the metronome: auditory events and music may afford more than just interval durations as gait cues in Parkinson's disease. *Frontiers in neuroscience* 10 (2016), 272.

[63] Giulio Rosati, Antonio Rodà, Federico Avanzini, and Stefano Masiero. 2013. On the role of auditory feedback in robot-assisted movement training after stroke: review of the literature. *Computational intelligence and neuroscience* 2013 (2013), 11.

[64] Emanuele Ruffaldi, Alessandro Filippeschi, Antonio Frisoli, Oscar Sandoval, Carlo Alberto Avizzano, and Massimo Bergamasco. 2009. Vibrotactile perception assessment for a rowing training system. In *World Haptics 2009-Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, Piscataway, NJ, USA, 350–355.

[65] Abraham Savitzky and Marcel JE Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry* 36, 8 (1964), 1627–1639.

[66] Nina Schaffert, Thenille Braun Janzen, Klaus Mattes, and Michael H Thaut. 2019. A review on the relationship between sound and movement in sports and rehabilitation. *Frontiers in psychology* 10 (2019), 244.

[67] Andrea Schiavio and Eckart Altenmüller. 2015. Exploring music-based rehabilitation for Parkinsonism through embodied cognitive science. *Frontiers in neurology* 6 (2015), 217.

[68] Richard A Schmidt. 1975. A schema theory of discrete motor skill learning. *Psychological review* 82, 4 (1975), 225.

[69] Richard A Schmidt. 1991. Frequent augmented feedback can degrade learning: Evidence and interpretations. In *Tutorials in motor neuroscience*. Springer, Heidelberg, Germany, 59–75.

[70] Richard A Schmidt, Timothy D Lee, Carolee Winstein, Gabriele Wulf, and Howard N Zelaznik. 2018. *Motor control and learning: A behavioral emphasis*. Human kinetics, Champaign, IL, USA.

[71] Gerd Schmitz, Bahram Mohammadi, Anke Hammer, Marcus Heldmann, Amir Samii, Thomas F Münte, and Alfred O Effenberg. 2013. Observation of sonified movements engages a basal ganglia frontocortical network. *BMC neuroscience* 14, 1 (2013), 32.

[72] Ludovic Seifert, Chris Button, and Keith Davids. 2013. Key properties of expert movement systems in sport. *Sports Medicine* 43, 3 (2013), 167–178.

[73] Ludovic Seifert, Hugues Leblanc, Romain Herault, John Komar, Chris Button, and Didier Chollet. 2011. Inter-individual variability in the upper–lower limb breaststroke coordination. *Human movement science* 30, 3 (2011), 550–565.

[74] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic bulletin & review* 20, 1 (2013), 21–53.

[75] Michael J Spilka, Christopher J Steele, and Virginia B Penhune. 2010. Gesture imitation in musicians and non-musicians. *Experimental Brain Research* 204, 4 (2010), 549–558.

[76] Nicholas Stergiou and Leslie M Decker. 2011. Human movement variability, nonlinear dynamics, and pathology: is there a connection? *Human movement science* 30, 5 (2011), 869–888.

[77] Sandra Sülzenbrück and Herbert Heuer. 2011. Type of visual feedback during practice influences the precision of the acquired internal model of a complex visuo-motor transformation. *Ergonomics* 54, 1 (2011), 34–46.

[78] Ana Tajadura-Jimenez, Nadia Bianchi-Berthouze, Enrico Furfaro, and Frederic Bevilacqua. 2015. Sonification of Surface Tapping Changes Behavior, Surface Perception, and Emotion. *MultiMedia, IEEE* 22, 1 (2015), 48–57. https://doi.org/10.1109/MMUL.2015.14

[79] Ana Tajadura-Jiménez, Maria Vakali, Merle T Fairhurst, Alisa Mandrigin, Nadia Bianchi-Berthouze, and Ophelia Deroy. 2018. Author Correction: Contingent sounds change the mental representation of one's finger length. *Scientific reports* 8, 1 (2018), 4875.

[80] Atau Tanaka. 2019. Embodied Musical Interaction. In *New Directions in Music and Human-Computer Interaction*. Springer, Heidelberg, Germany, 135–154.

[81] Michael Thaut and Volker Hoemberg. 2014. *Handbook of neurologic music therapy*. Oxford University Press, Oxford, UK.

[82] Carlos Trenado, Anaí González-Ramírez, Victoria Lizárraga-Cortés, Nicole Pedroarena Leal, Elias Manjarrez, and Diane Ruge. 2019. The Potential of Trial-by-Trial Variabilities of Ongoing-EEG, Evoked Potentials, Event Related Potentials and fMRI as Diagnostic Markers for Neuropsychiatric Disorders. *Frontiers in neuroscience* 12 (2019), 850.

[83] William E Winkler. 1990. String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage.. In *Proceedings of the Survey Research Methods Section*. American Statistical Association, Alexandria, VA, USA, 354–359.

[84] Daniel M Wolpert, Jörn Diedrichsen, and J Randall Flanagan. 2011. Principles of sensorimotor learning. *Nature Reviews Neuroscience* 12, 12 (2011), 739–751.

[85] Howard G Wu, Yohsuke R Miyamoto, Luis Nicolas Gonzalez Castro, Bence P Ölveczky, and Maurice A Smith. 2014. Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience* 17, 2 (2014), 312–321.

[86] Gabriele Wulf. 2013. Attentional focus and motor learning: a review of 15 years. *International Review of sport and Exercise psychology* 6, 1 (2013), 77–104.

[87] Gabriele Wulf and Charles H Shea. 2002. Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic bulletin & review* 9, 2 (2002), 185–211.

[88] Goh Yamamoto, Kensuke Shiraki, Masami Takahata, Yutaka Sakane, and Yoichi Takebayashi. 2004. Multimodal knowledge for designing new sound environments. In *Proceedings of the workshop "Mobile HCI and Sound Workshop" of the 6th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI)*. 6. http://mcs.open.ac.uk/mobsound/

A DATA PROCESSING

Algorithm 1 describes the steps to locate the true start and end of a gesture. Fig. 12 shows four trajectory instances of the same participant in the same day: (a) and (b) similar gestures; (c) a gesture with added elements and (d) a gesture with missing elements. We use the Euclidean distance to compute the difference:

$$V_{\text{diff}} = V_d(S_i, S_j) \tag{1}$$

where S_i is the direction vector of a particular trial and S_j is that of trial $j = i + 1$.

Using Algorithm 2, the distance between (a) and (b) is $V_{a,b} = 365.71$, between (b) and (c) is $V_{b,c} = 928.18$ and between (c) and (d) is $V_{c,d} = 1011.23$. The code can be found online [9].

Algorithm 1: Locate the start and the end of gesture data.

Data: Data-List, Min, Max
Result: Return New-Data-List, start, end

```
// Find all the peaks within range of Min and Max
Find peaks (Data-List, prominence = (Min, Max))
start = peaks[0]
end = peaks(Length(peaks)−1)
while Data-List[start] >= Data-List[start −1] do
    start = start −1
while Data-List[end] >= Data-List[end +1] do
    end = end +1
for i in range(start, end) do
    New-Data-List.append(Data-List[i])
return New-Data-List, start, end
```

[9] https://nubo.ircam.fr/index.php/s/p3sr7rMfDwRr9mb

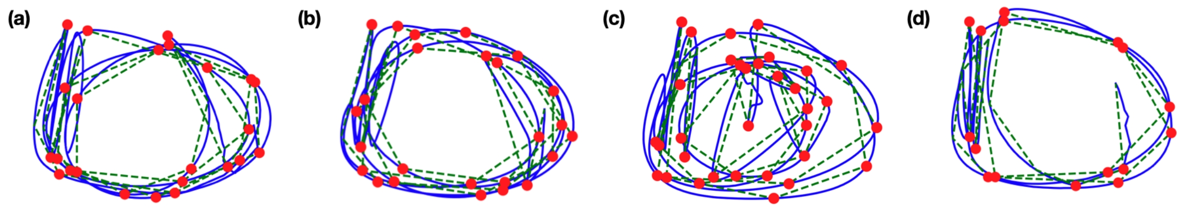

Fig. 12. Examples of four trajectories of p5: (a) and (b) similar gestures (28 directions) (c) a gesture with added elements (36 directions) (d) a gesture with missing elements (21 directions).

Algorithm 2: Compute the Euclidean distance between two direction vectors.

Data: Direction vector 1 (Dv1), Direction vector 2 (Dv2)
Result: Return Euclidean distance

```
if Length(Dv1) < Length(Dv2) then
    for i in range(Length(Dv2)-Length(Dv1)) do
        Dv1.append(0)
else
    for j in range(Length(Dv1)-Length(Dv2)) do
        Dv2.append(0)
Ed = distance.euclidean(Dv1, Dv2)
return Ed
```