Journal of New Music Research

ISSN: 0929-8215 (Print) 1744-5027 (Online) Journal homepage: https://www.tandfonline.com/loi/nnmr20

Sensor Choice for Parameter Modulations in Digital Musical Instruments: Empirical Evidence from Pitch Modulation

Mark T. Marshall , Max Hartshorn , Marcelo M. Wanderley & Daniel J. Levitin

To cite this article: Mark T. Marshall , Max Hartshorn , Marcelo M. Wanderley & Daniel J. Levitin (2009) Sensor Choice for Parameter Modulations in Digital Musical Instruments: Empirical Evidence from Pitch Modulation, Journal of New Music Research, 38:3, 241-253, DOI: 10.1080/09298210903085865

To link to this article: https://doi.org/10.1080/09298210903085865

Published online: 17 Dec 2009.

Journal of New Music Research
2009, Vol. 38, No. 3, pp. 241–253

Routledge
Taylor & Francis Group

Sensor Choice for Parameter Modulations in Digital Musical Instruments: Empirical Evidence from Pitch Modulation

Mark T. Marshall, Max Hartshorn, Marcelo M. Wanderley, and Daniel J. Levitin

McGill University, Canada

Abstract

This paper describes ongoing research into the design of new digital musical instruments (DMIs). While many new DMIs have been created using a variety of sensors, there has been relatively little empirical research into determining the optimal choice of sensor for control of specific musical functions. In this paper we attempt to identify an optimal choice of sensor for the control of parameter modulations in a DMI.

Two experiments were conducted. In the first, pianists and violinists were tested on three strategies for producing pitch modulations. Both subjective user ratings and objective performance scores were analysed. The results suggest that modulated applied pressure is the optimal control for pitch modulation. Preference and performance did not appear to be directly mediated by previous musical experience. In the second experiment, the accuracy, stability and depth of modulation were measured for a number of musicians performing modulations with each of three control strategies. Results indicate that some options offer improved stability or accuracy over others and that performance with all strategies is significantly dependent on the speed of modulation. Overall results show that the optimal choice of sensor should be based on a combination of subjective user preference ratings and objective performance measurements.

1. Introduction

In recent years a large body of research has grown around the development of new digital musical instruments. A digital musical instrument (DMI) is essentially any musical instrument that makes use of a computer for sound generation and in which the control interface is separable from the sound generator (Miranda & Wanderley, 2006). A sensor is typically used to convert a user's physical movement into an electronic signal. That signal is then mapped onto one or more musical aspects of a sound generated by the computer, such as pitch, volume or tempo. Currently available sensors can measure performance parameters including physical force and motion to heart rate and even brain waves (Bongers, 2000). The same sensor can be used to control different musical functions, and, just as importantly, different sensors can be used to control the same function. This separation between physical input and musical output gives designers the freedom and control to create new mappings that were previously extremely difficult or even impossible to create in acoustic instruments. The question is, which sensors are useful for controlling which musical functions? Are the musical gestures elicited by a specific sensor useful for controlling a specific musical attribute (Vertegaal et al., 1996)?

Here the question of usability is brought to the fore. There is a need for a conceptual and experimental framework for identifying precisely what makes some sensors more suitable for certain tasks than others (Levitin et al., 2002). The idea that interfaces should be intuitive, that devices should provide visible clues as to their function, has already proved an essential principle in the design of everyday non-musical objects (Norman, 2002). New instruments should also be such that they allow for a degree of explorability; that is they

Correspondence: Marcelo M. Wanderley, Input Devices and Music Interaction Laboratory, Centre for Interdisciplinary Research in Music Media and Technology, McGill University, 555, Sherbrooke Street West, H3A 1E3 Montreal, QC, Canada.
E-mail: marcelo.wanderley@mcgill.ca
Daniel J. Levitin, Department of Psychology, Centre for Interdisciplinary Research in Music Media and Technology, McGill University, 555, Sherbrooke Street West, H3A 1E3 Montreal, QC, Canada. E-mail: daniel.levitin@mcgill.ca

DOI: 10.1080/09298210903085865

should be appropriately complex as to allow for the development of musical skill and expertise (Hunt, 2000). At a glance these may seem like two completely opposing demands (Wessel & Wright, 2002). In fact, as Levitin et al. (2002) note, the challenge of instrument design is to strike a balance between simplicity and complexity, to allow for just the right amount of challenge, frustration and boredom to maintain a user's interest without alienating her.

Vertegaal et al. (1996) argue that digital musical instrument usability can be greatly enhanced by taking into account the visual, tactile and kinesthetic feedback offered by different sensing devices. They propose a categorization of sensors based on the type, direction and resolution of the parameter sensed, as well as the primary feedback offered by the sensors. Along with this sensor classification they identify three classes of basic musical function: *absolute dynamical*, *relative dynamical* or *static*. From these classifications they produce a system which maps classes of sensor to classes of musical function.

However, there exists relatively little empirical research to identify optimal sensors to control specific functions in digital musical instruments. As the field of Human–Computer Interaction (HCI) has long examined the relationship between users and computer interfaces, it has been suggested that it could provide useful paradigms for the study of digital musical interfaces (Wanderley & Orio, 2002; Isaacs, 2003). In particular, the use of representative tasks to evaluate an interface is a fundamental part of HCI that could be applied to the evaluation of digital musical instruments. In order to do this, we require a list of representative musical tasks, the counterparts of interaction tasks used in human–computer interaction such as those described by Buxton (1987). Just as HCI research measures subject performance on discretely defined interaction tasks, DMI research can measure subject performance on discretely defined *musical* tasks.

This paper deals with ongoing research into the use of sensors in digital musical instruments. It begins with a survey of sensors in existing DMIs and examines certain sensors that are commonly used in new instruments. This is followed by the report of a series of experiments that attempt to evaluate the usefulness of some of the most common sensors for specific musical tasks.

2. Sensor use in existing digital musical instruments

In order to examine the current state of sensor use in digital musical instruments, we performed a detailed literature review of all of the papers and posters from all eight years of the conference on *New Interfaces for Musical Expression* (*NIME*). This review comprised 577 papers and posters, containing descriptions of 266 different instruments. Some papers described multiple instruments and some instruments were described in multiple papers. Those instruments described in multiple different papers (generally in different years) usually involved descriptions of new applications or design improvements over the original. While the review covered a number of different aspects of the design on these instruments (Marshall, 2009), for this paper we concentrate on their use of sensors.

Table 1 shows the most popular sensors in digital musical instruments presented at the NIME conferences, along with the number of instruments in which one or more of each particular sensor was found. Note that this is not a count of the number of sensors used (as an instrument may include multiple copies of a particular sensor), but instead offers a measure of the relative popularity of particular sensors. The total sum of sensor types used was 595 sensors, implying that on average each instrument used 2.25 sensors.[1]

From Table 1, we can see that certain sensors are used in a large number of instruments, particularly FSRs (26% of all instruments), accelerometers (21%), video cameras (20%) and buttons or switches (19%). Interestingly, when we examine the use of these sensors in different instruments we find that the same sensor can be used to control completely different types of parameter in different instruments. Also, in some cases the sensors themselves are used to detect different movements, such as using an accelerometer to detect either tilt or acceleration, or both.

How then do we decide which sensors to use to design a new digital musical instrument? Which sensors (or sensor/gesture combinations) are most suited to controlling a specific task?

Wanderley et al. (2000) carried out an exploratory experiment to examine this issue. In their study subjects performed on a digitizing tablet with a stylus and two

Table 1. Most popular sensors from NIME instruments.

Sensor	Occurrences	Property Sensed
FSR	68	Force
Accelerometer	56	Acceleration
Video Camera	54	
Button/Switch	51	Position (On/Off)
Rotary Potentiometer	31	Rotary Position
Microphone	29	Sound Pressure
Linear Potentiometer	28	Linear Position
Infrared Distance Sensor	27	Linear Position
Linear Position Sensor	23	Linear Position
Bend Sensor	21	Rotary Position

[1]It should be noted that because many of the instruments based around video cameras (40 out of 54) used only one sensor, the average number of sensor types per instrument for non-camera-based instruments is correspondingly higher than this, at 2.55 sensors per instrument.

extra sensors: a Force Sensitive Resistor (FSR) and a Linear Position Sensor placed on the tablet. The FSR is sensitive to pressure, and subjects were able to modulate pitch up by pressing on it with their finger. Repeated increases and decreases of pressure allowed them to produce a low-frequency oscillation of the pitch. The linear position sensor used operates similarly to the FSR, but rather than measure only finger pressure, it measures both the pressure and the position of one's finger across a tape-like strip (and might therefore more accurately be called a Linear Position and Force Sensitive Resistor). Here pitch could be modulated by moving one's finger back and forth across the sensor.

Wanderley and colleagues examined which sensors participants preferred for the control of pitch modulation. Subjects would start with the stylus at a specific point representing one pitch. They would then move the stylus to a target representing a different pitch and modulate that pitch in one of three ways: by pressing the FSR in a repeated motion, by sliding their finger back and forth across the linear position sensor, or by tilting the stylus along the tablet's x-axis.[2] Subjects performed the task with each method and then provided subjective ratings for those methods. Results showed an approximately linear downward trend: pressing was the preferred method for modulating pitch, followed by sliding, then tilting.

A later study had subjects perform simple tasks, such as selecting pitches to play a melody, or modulating pitch with their finger using touch sensitive devices (Marshall & Wanderley, 2006). For the modulation task, subjects cycled through a four-note melody by pressing a button with one hand and, on the fourth note, modulated that pitch with the other hand. While a variety of sensing devices were evaluated, we will comment on the two that are of special interest to us, namely the force sensitive resistor (FSR) and the linear position sensor.

Subjects performed the task using each method and then subjectively rated the methods as in the previous experiment. Here subjects rated the linear position sensor above the FSR. However closer examination of the data revealed that subjects modulated pitch on the linear position sensor one of two ways: they either slid their finger back and forth across the sensor, or rolled it back and forth (in a manner more similar to vibrato production on a fretless stringed instrument). Subjects who rolled their finger tended to prefer the linear position sensor to the FSR while subjects who slid their finger preferred the FSR to the linear position sensor (as would be suggested by the findings from the digitizing tablet study). However, the rolling method was rated highest despite its similarity to the x-axis tilt method of the tablet study, which was rated lowest (possibly due to its being performed with the dominant hand).

Given these results however, several issues still remain to be dealt with. In particular these experiments did not examine issues such as:

- the effect of previous musical experience on user preference;
- objective measurements of performance as well as subjective user preference measurements;
- the effects of control parameters such as the speed of control on the preference and performance measures;
- the use of different gestures with the same sensor.

The remainder of this paper deals with two experiments carried out to further examine these issues. As in the previous two studies, the present study examines the production of pitch modulation on digital musical instruments. We chose to focus on pitch modulation because of the previous work on the topic and because it represents a widely used musical skill that can be learned, measured and manipulated. What follows is a discussion of these experiments, which aimed to examine the ability of subjects to modulate a note of fixed pitch, using three different methods of modulation.

3. Experiment 1: User preference and previous musical experience

Experiment 1 investigated whether previous musical experience plays a role in determining method preference and performance. The previously cited studies did not take into account or control for the previous musical experience of subjects and it is possible that the influence of previously learned performance gestures may be one of the reasons for the discrepancy between the above two studies. In this experiment we tested pianists and violinists. We hypothesize that violinists will perform better than pianists overall as they presumably have much more experience with the performance of pitch modulation (through vibrato). We also hypothesize that violinists will prefer and perform better with the rolling method since it is most similar to the way they produce such modulations on a violin. For this study we also introduce an objective measure of performance: the stability of the speed of the modulation. This allows us to perform comparisons between a user's subjective rating and objective performance, which may also prove useful.

3.1 Participants

Twenty-seven right-handed musicians with at least eight years of musical experience on their instrument were recruited from McGill University and paid CAD$10. Handedness was assessed using the Edinburgh

[2]Note that this third strategy was performed with the right hand, i.e. the hand holding the stylus, while the first two were produced with the left hand.

Handedness Inventory (Oldfield, 1971). Nine subjects were dropped from the final analysis because an equipment error rendered their data unusable. Of the remaining 18 subjects there were 9 pianists and 9 violinists.

3.2 Design and materials

The experiment followed a mixed design. The between subjects factor was instrument played (two levels, piano or violin) and the within subjects factor was modulation method (three levels, pressing, sliding and rolling).

For the *pressing method*, subjects used an FSR. Increasing finger pressure on the sensor raised the pitch while a decrease in finger pressure lowered the pitch back down. Modulation was produced by applying pressure to the sensor in a pulsating motion.[3] The FSR was mounted flat on a small block of wood without any padding.

For the sliding and rolling method, subjects used a linear position sensor. A light finger pressure is required to activate the sensor. Moving the position of one's finger to the right causes pitch to increase, while movement to the left causes pitch to decrease. For the *sliding method* subjects produced modulations by sliding their finger back and forth across a limited portion of the sensor (about 3 cm wide). Too large a spacing would have made it difficult for the subject to maintain vibrato at the speeds we were looking for, while too small a spacing would have made the vibrato harder to control. Pilot testing in our laboratory prior to the experiment determined that the optimal spacing for this gesture was ~ 3 cm. Scaling on the sensor was changed for the *rolling method* from 3 cm to <1 cm. Subjects here could produce modulations by simply pivoting their finger back and forth at a demarcated point on the sensor (similar, though not identical to the way such modulations are produced on a violin). For all three methods, subjects alternated using the index finger of either their right or left hand, depending on the conditions currently being tested. The sensor was placed at elbow level to the subject whose forearm rested on the table at approximately a 90° angle to the upper arm.

The experiment was run on a 17 inch Macintosh Powerbook G4. The sensor output was converted to a computer-usable format using an AVR-HID analogue-to-digital converter,[4] at a rate of 100 Hz and with a 10 bit resolution. The visual programming environment Max/MSP was used to map the signal onto a musical output. The musical tone that subjects were able to modulate was created with the sound synthesis software Tassman 4.[5] We used preset tone #44: Simple Sine Lead, a straightforward sine-wave tone, modified to have no built in pitch variation, set to G3 (196 Hz). Subjects produced the tone by holding down on the spacebar of the laptop with one hand, using the opposite hand to modulate the tone. They were able to hear the tone through the laptop speakers which were set to a comfortable volume.

To calculate the frequency of the modulation over time (as it changes over the course of the recording), we used an additive analysis performed from the short-term Fourier transform of the measured control signal. Only one peak was selected in the 1–7 Hz range, using 192-sample (1920 ms) Blackman–Harris 92 dB windows, with a hop size of 16 samples (160 ms). The 'instantaneous' frequency was then obtained at a sample rate of 7 Hz. This frequency adequately represents the periodicity of the control signal, as the only remaining components of the additive analysis were at multiples of the fundamental frequency and had much lower amplitudes than the fundamental. These additional components contribute mainly to the control curve shape, which varies between sinusoidal and triangular, but do not change the centre frequency or depth of the curve. This algorithm was adapted from the second-order sinusoidal model (Marchand & Raspaud, 2004). From this we calculated a participant's mean speed for a specific modulation, along with their standard deviation from that mean.

While all subjects tested were right-handed, right-handed violinists traditionally produce vibrato with their left hand. Pianists, however, finger and play notes with both hands and presumably have little to no experience with pitch modulation, so we should expect right-handed pianists to both perform better with and prefer using their right (dominant) hand relative to their left hand, but to perform worse than violinists using their (well practiced) left hand. We tested all subjects with both hands to see if there was really any difference.

Subjects were instructed to perform modulations at both slow and fast speeds. While no exact speeds were specified to the participants, a recorded sample of a slow and a fast pitch modulation (produced using the linear position sensor with the sliding method) was played to subjects a few times at the beginning of the experiment to demonstrate what we meant by a slower or faster modulation. We were not interested in having subjects match the speeds precisely; we just wanted to obtain generally similar slow and fast modulations throughout the experiment. The order of left–right slow–fast modulation tasks was also randomized throughout the experiment.

[3]This means that the modulation could only be produced higher than the base pitch for this method. This is similar to pitch modulation on fretted stringed instruments, such as the guitar. The other methods could both raise and lower the pitch allowing for a modulation more like that of fretless string instruments, wind instruments or the human voice.

[4]See http://www.idmil.org/projects/avr-hid for a detailed description of this device.

[5]Produced by Applied Acoustics Systems. See http://www.applied-acoustics.com/tassman.htm for more information.

3.3 Procedure

Subjects arrived at the lab and were given an Information/Consent form to read and sign. They were shown the experimental interface and told that they would be producing pitch modulations on this interface using three different methods.[6] They were played the slow and fast modulation samples and told to try and maintain a uniform rate for each modulation they produced. In the first trial subjects were introduced to the sliding, rolling and pressing methods. They were verbally and visually instructed on how to perform each method and then given up to a minute to practice before we began recording their output.

After going through all three methods once, subjects were then given the Queens Musical Background Questionnaire (Cuddy et al., 2005) and the Edinburgh Handedness Inventory (Oldfield, 1971) to verify their musical background and handedness. No subjects were discarded. The second trial was exactly the same as the first except the order of the tasks was counterbalanced. After each method subjects were given a questionnaire that asked them to rate the ease of use of the method, whether they preferred it for slow or fast modulation speeds, whether they preferred it with their right or left hand, and an overall preference rating for that method.[7] They were also given an opportunity to add any comments they might have with regard to the method in question. After the second trial was completed subjects were paid and debriefed as to the nature of the experiment. Subjects each completed a total of two trials.

3.4 Data analysis

Results were analysed using the Statistical Package for the Social Sciences software (SPSS). The subjective questionnaire data and objective measurements were analysed separately at first. Correlation analyses were then performed between the two sets of data. Both sets were analysed using repeated measure ANOVAs with t-tests for specific relevant comparisons. *Post hoc* tests were performed using Tukey's Honestly Significant Difference (HSD).

[6]While the term 'vibrato' might seem more intuitive, vibrato includes apsects of pitch, amplitude and timbral modulation. As we are only interested in modulation of a single parameter and to remove any ambiguity, we consistantly used the term 'pitch modulation' when instructing the participants.

[7]Ease of use refers to a rating by the participant of how easy it is to control the parameter using a particular sensor and gesture. Overall user preference refers to a rating by the participant of how they liked performing the task using a particular sensor and gesture.

3.5 Results

3.5.1 Subjective data

Participant ease of use and preference ratings were in most cases very similar, suggesting that participants did not really distinguish between the two measures. The overall differences in preference ratings between the violinist and pianist groups were marginally significant [$F(1, 16) = 3.39$, $p = 0.08$], while there was no significant difference between instrument groups for the ease of use rating [$F(1, 16) = 0.49$, $p = 0.49$]. The violinist group rated the pressing method significantly lower than the pianists [$t(16) = 3.06$, $p < 0.01$], the sliding method marginally lower [$t(16) = 1.95$, $p = 0.07$], and the rolling method insignificantly higher [$t(16) = -0.73$, $p = 0.48$]. This does not fully confirm our hypothesis, which was that violinists would prefer the rolling method to pianists; nevertheless the effect is in the right direction.

We found a significant effect of method used across all subjects [preference: $F(2, 32) = 7.34$, $p < 0.01$, ease of use: $F(2, 32) = 5.38$, $p < 0.01$]. *Post hoc* tests indicated that subjects rated the pressing method significantly higher than the other methods in terms of both ease of use [rolling: $p < 0.05$, sliding: $p < 0.05$] and preference [rolling: $p < 0.05$, sliding: $p < 0.01$]. No significant differences were found between the sliding and rolling methods for either ease of use or preference ratings.

Looking at the instrument groups individually, the pianist group preferred the pressing method, followed by the sliding method and then the rolling method (see Figure 1(a)) [preference: $F(2, 16) = 4.26$, $p < 0.05$, ease of use: $F(2, 16) = 4.30$, $p < 0.05$]. *Post hoc* tests show that the pressing method was rated significantly higher than the sliding and rolling methods in terms of both ease of use [sliding: $p < 0.05$, rolling: $p < 0.05$] and preference ratings [sliding: $p < 0.05$, rolling: $p < 0.01$]. Once again, no significant differences were found between the sliding and rolling methods for either rating.

The violinist group preferred the pressing method, followed by the rolling method and then the sliding method (see Figure 1(b)). In this case the differences were marginally significant [preference: $F(2, 16) = 3.79$, $p < 0.05$, ease of use: $F(2, 16) = 3.50$, $p = 0.05$]. *Post hoc* tests indicated only that the sliding method was rated significantly lower than the other methods in terms of both preference [pressing: $p < 0.05$, rolling: $p < 0.05$] and ease of use [pressing: $p < 0.05$, sliding: $p < 0.05$].

Pianists show significantly greater ratings for their right hand than violinists [$F(1, 16) = 10.29$, $p < 0.01$]. Speed preference hovers generally around the middle with a slight preference in both groups on all the methods for slower speed vibrato [average rating: 4.17]. There were no significant differences in speed preference between violinists and pianists [$F(1, 16) = 0.38$, $p = 0.55$] or between the different methods [$F(2, 32) = 1.31$, $p = 0.28$]. However both groups rated slower speed

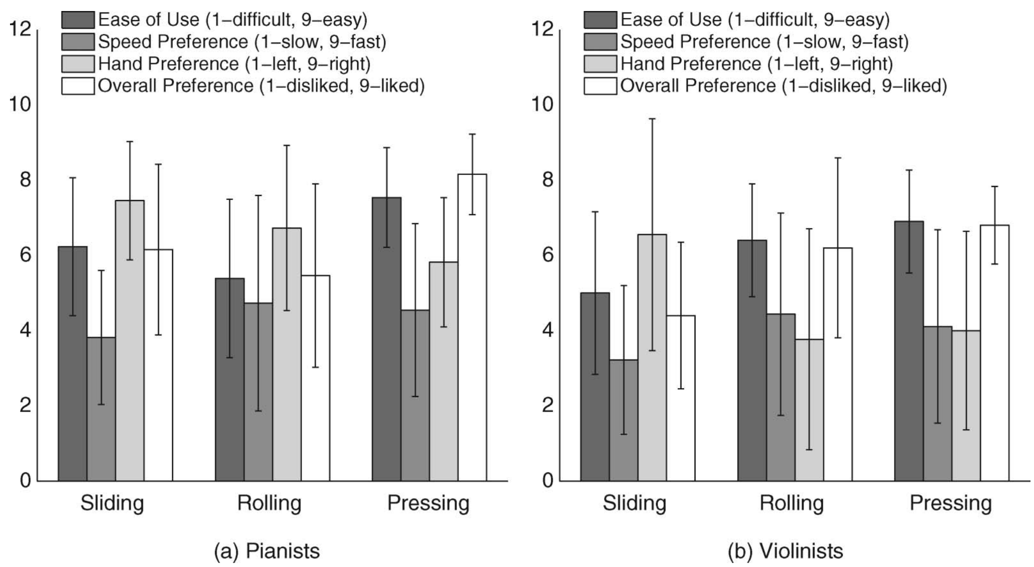

Fig. 1. Mean questionnaire responses.

modulations slightly higher with the sliding method. This makes sense as the finger moves over a somewhat larger distance with this method.

3.5.2 Objective data

The standard deviation dependent variable measures the extent to which modulation samples deviate from a constant rate. Repeated measure ANOVAs were performed with instrument as the between groups factor. No significant difference between groups was found for the fast modulation [left hand: $F(1, 16) = 0.71$, $p = 0.42$, right hand: $F(1, 16) = 0.01$, $p = 0.93$] or slow modulation [left hand: $F(1, 16) = 0.51$, $p = 0.49$, right hand: $F(1, 16) = 0.34$, $p = 0.57$].

There was also no significant overall difference in stability between the left hand and the right hand in both the pianist [fast: $F(1, 8) = 0.06$, $p = 0.82$, slow: $F(1, 8) = 0.31$, $p = 0.59$] and violinist [fast: $F(1, 8) = 2.28$, $p = 0.16$, slow: $F(1, 8) = 1.71$, $p = 0.21$] groups.

Fast modulations were significantly less stable than slow modulations for both pianists [left hand: $F(1, 8) = 15.70$, $p < 0.01$, right hand: $F(1, 8) = 4.99$, $p < 0.05$] and violinists [left hand: $F(1, 8) = 13.80$, $p < 0.01$, right hand: $F(1, 8) = 24.86$, $p < 0.001$]. There were also significant overall differences in mean frequency for each method [$F(2, 32) = 34.623$, $p < 0.001$]. The rolling method was performed the fastest at an average frequency of 3.99 Hz, while the sliding and pressing methods were played significantly slower at a frequency of 2.65 and 2.82 Hz respectively [Tukey's HSD, $p < 0.01$ for both sliding and pressing]. Pearson's correlation coefficient reveals significant correlations between the mean frequency and frequency deviation on each method [sliding: $r(16) = 0.77$, $p < 0.001$, rolling: $r(16) = 0.73$, $p < 0.001$, pressing: $r(16) = 0.83$, $p < 0.001$]. This means that there is a significant correlation between the mean frequency of modulation and the stability of the modulation, which is in keeping with the negative relationship found between speed and stability.

We found a significant difference in stability based on the method used [$F(2, 32) = 55.84$, $p < 0.001$]. The rolling method was significantly less stable than the sliding method [Tukey's HSD, $p < 0.01$] and the pressing method [Tukey's HSD, $p < 0.01$], with no significant difference between the stability of the pressing and sliding methods. We found no effect of instrument on stability [$F(1, 10) = 0.06$, $p = 0.81$], nor of an interaction between instrument and method [$F(2, 32) = 0.70$, $p = 0.51$].

While there are no significant correlations between preference ratings and stability for any method, some slight trends are apparent when we look at the data. In the piano group, the rolling method receives lower preference ratings and is less stable than the other two methods (see Figure 2(a)). This suggests a relationship between preference and performance stability; perhaps participants give lower ratings to this method because their performance is less stable. However when we look at the data from the violin group this relationship disappears (see Figure 2(b)). Even though performance with the rolling method is less stable than the sliding method, it receives a higher preference rating. It is likely here that another factor besides performance stability has influenced the participant's preference ratings. One possible explanation is that the rolling method is most similar to how a violinist performs vibrato on their own instrument, resulting in higher preference ratings.

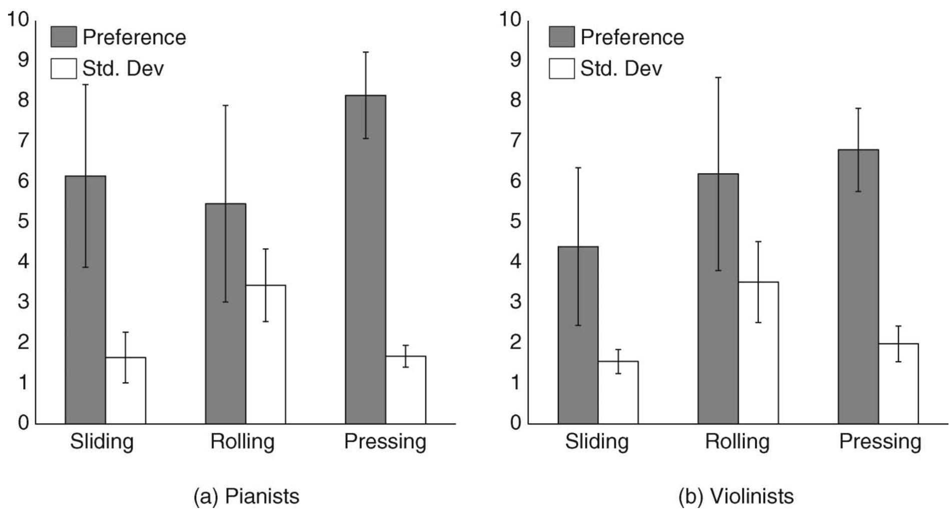

Fig. 2. Subjective preference compared to standard deviation (scaled for comparison).

3.6 Discussion

Overall, we found that participants prefer the pressing method and that this is true of both violinists and pianists. While violinists did exhibit slightly higher ratings for the rolling method than the pianists, the effect was statistically insignificant. Moreover we found no evidence that violinists produce modulations that are more stable than those of pianists. Therefore we cannot directly conclude that preference for (or performance on) novel mappings is influenced by musical experience.

Conversely, we also cannot conclude the opposite, that musical experience has no effect on preference or acquisition of new musical mappings. It may be that the mapping used was in reality not all that similar to the traditional method of violin vibrato production. Violinists typically hold their instrument up to their chin, with their arm curved upwards and their hand perpendicular to their body. In contrast, subjects here were asked to perform modulations with their hand flat on an elbow high table. A number of the violinist subjects even asked if they could hold the sensors in the manner they would normally hold a violin, but were not allowed to do so. Moreover, the tactile feedback from the sensor is undoubtedly different from that to which an experienced violinist is accustomed. The failure of 'skill transfer' from creating vibrato on a violin neck with a rolling motion to creating one on our sensors may be simply due to the different feel of the two mediums.

The fact that pianists show a slight correlation between stability and preference suggests they relied to a certain extent on a subjective assessment of their own performance in making their ratings. This correlation was absent in violinists, leading us to question what else may have been informing their preference. Perhaps the violinists' previous vibrato experience has, in some indirect way, influenced their method preference. However, it is difficult to discern from these data the nature of the effect.

As can be seen in Figure 1(b), violinists gave a lower rating to the sliding method than the rolling method, despite the fact that their sliding method modulations were more stable. When we look more closely at the data we see that both the pianists' and the violinists' performances are significantly more stable on the sliding method with their right hand versus their left. It is the only method where this is the case, and it only occurs for fast modulations. Right-handed pianists (or any right-handed control) might expect to perform better with their dominant hand. However this may not be the case with right-handed violinists. These musicians have extensive experience creating pitch modulations (in the form of vibrato) with their left hands and therefore might expect to perform better with their left hand. During the experiment, a number of violinists expressed hesitation when asked to perform the tasks using their right hand. Some were doubtful they could perform accurately. Perhaps this resulted in them giving lower subjective ratings to this method.

The factor most clearly responsible for mediating performance (and possibly in turn preference) was the speed of the modulation. Fast modulations were found to be significantly less stable than slow modulations. Moreover there were significant differences in mean speed between the different methods, and these differences were significantly correlated with differences in accuracy. Modulation speed therefore represents a serious potential confounding variable. Differences in preference and performance among the different methods may primarily be the result of how fast modulations were performed on each method.

In this experiment the three methods were played at different speeds despite the fact that subjects were told to try and maintain consistent speeds throughout all experimental trials. While some random variability is to be expected, the fact that subjects averaged faster speeds for some methods and slower speeds for others suggests they felt more comfortable performing at these speeds. Just as different instruments may naturally elicit their own specific set of movements and gestures, these three methods naturally elicit modulations at specific frequencies. From a holistic perspective, modulation frequency on a given method can even be seen as an integral component of the mapping itself, a function of the interface's design.

Nevertheless speed is something that should be controlled for, it just needs to be done carefully. Pilot testing would be necessary to determine an experimental speed that works well for each method being evaluated. There is no sense in testing each method with a frequency of 7 Hz if two of the methods only work well at about 5 Hz. In addition the results of any such experiment would need to be interpreted cautiously. For example, even if a method tests very well at 3 Hz it may not be suitable if the intent is to mimic violin vibrato, as this type of vibrato is traditionally produced at a speed of 5–7 Hz (Papich & Rainbow, 1974).

From this we can see the possibility of examining subject's performance at a variety of different fixed speeds. This would allow us to determine whether speed is a factor in the production of modulations with each of these different gesture mappings. This question forms the basis of the second experiment, described in the next section, which is centred on gaining objective measures of the performance of pitch modulations using each gesture mapping and at a variety of fixed speeds.

4. Experiment 2: Objective performance measurement and the effect of target speed

Our second experiment examined the ability of participants to perform modulations at different speeds using each of the three previous methods. In particular we sought to improve upon the methodology of the previous studies by including multiple objective measures of performance. In addition to measuring stability as we did in the previous study, here we added measures of accuracy and modulation depth, which we will detail shortly, to provide a more thorough objective assessment.

This experiment also offered an opportunity to validate some of the results from our first experiment, while controlling more carefully for the effects of speed. As was noted in the previous section, the speed of modulation influenced stability far more than any other variable. Simply playing subjects a sample modulation speed and asking them to mimic the speed was not enough to control for this. In Experiment 2 subjects were asked to attempt to match the speed of a pitch modulation which was being played to them at the same time. We were interested in discovering whether or not the higher stability of the sliding and pressing methods relative to the rolling method held with more stringent control on speed. We were also interested to see if the effects were robust enough to hold up under a range of speeds.

4.1 Participants

There were 10 participants in this study, each of whom was compensated CAD$10 for taking part. All were musicians with at least eight years of experience on their instrument and all were right-handed.

4.2 Design and materials

This experiment follows a within-subjects design. The factors examined are the method of modulation (three levels, sliding, rolling, pressing) and speed of modulation (six levels, 1, 2, 3, 4, 5, and 6 Hz).

The sensor setup used by the participants to perform each method was the same as those described in the previous study. The experiment was run using a 15-inch MacBook Pro computer. As a result of the equipment problems for some subjects during the first study, the sensor output was converted to a computer-usable format using an Electrotap Teabox sensor interface (Allison & Place, 2005), which offers a higher resolution and faster sampling rate than the interface used in the first study. This was connected to the computer using an S/PDIF optical cable.

Max/MSP was once again used to process the incoming sensor data. It was also used to produce the sound output, using a simple sine wave generator. However, the patch created in Max/MSP was also able to produce sample signals at each of the speeds which were to be tested. These signals were also output as a sine tone, but at a fundamental frequency which was a perfect fifth higher than that of the sound output from the participants' actions. The sample signals provided a sound containing a pitch modulation that the participant was to attempt to match in frequency of modulation. Again, subjects produced the tone by holding down on the spacebar of the laptop with one hand while using the opposite hand to modulate the tone.

At the beginning of the experiment subjects were informed that they would be producing modulation using three different methods at six different speeds. They were told that for each modulation they were to produce, there would be a sample playing, the speed of which they would attempt to match.

The sensor input was sampled in Max/MSP at a rate of 8000 Hz and recorded to an audio file for later processing. Processing was performed in Matlab, using the same algorithm as was used in the first experiment. However, due to the higher sampling rate of the recorded signals, some parameters were changed and additional filtering was added. Firstly, the data were low-pass filtered using a fourth-order Butterworth low-pass filter with a cut-off frequency of 100 Hz to remove any high-frequency content before processing. It was then resampled to a sampling rate of 2000 Hz. For the analysis we used a window length of 4196 samples (524.5 ms) and a hop size of 100 samples (12.5 ms), resulting in the output being determined at a rate of 20 Hz. The algorithm then determined and recorded the mean modulation frequency produced, the standard deviation from this mean (as a measure of stability), the deviation of the signal from the target frequency (as a measure of accuracy) and the RMS amplitude of the signal (which gives a measure of depth).

It is important here to note the difference between stability and accuracy. The accuracy of the performance is a measure of how close to the target speed the performance was. This is measured as the deviation from the target frequency. The stability, on the other hand, is a measure of how much variation occurred over the course of the production. This is determined as the deviation of the signal from the mean speed the participant produced (i.e. the standard deviation). It should also be noted that as we are calculating deviations the result is actually the inverse of the stability and accuracy. This means that a more accurate or stable performance will result in a lower deviation (from the target and mean frequencies respectively).

It should also be noted that although we are measuring the depth of the modulations produced by the participants, the sample sounds that were presented to the participants all used the same depth of modulation. This depth of modulation was also stable over the length of the sample sounds. Also, the participants were not asked to reproduce the depth of the sample sound, but only its speed of modulation. Our aim in examining the depth of the modulation then was to determine the effect of changes in the speed of modulation on other parameters of the modulation.

As with the previous study, the order of the presentation of the combination of methods and speeds was randomized throughout the experiment.

4.3 Procedure

Subjects arrived at the lab and were given an Information/Consent form to read over and sign. Subjects were shown the experimental interface and told that they would be producing modulations on this interface using three different methods. They were also told that in each case the system would produce a signal and that they would have to try to match as close as possibly the speed of that modulation. A sample was played for them at this time so that they would know what to expect.

Subjects were introduced to the three methods and were then verbally and visually instructed on how to perform each method. They were then allowed up to a minute to practice before we began recording their output. After completing two of the three trials, the subjects were given the Queens Musical Background Questionnaire (Cuddy et al., 2005). Upon finishing the experiment, subjects were given monetary compensation and debriefed as to the nature of the experiment.

4.4 Data analysis

Results were analysed using SPSS. All data was analysed using a 3×6 (Method $\times$ Speed) factorial ANOVA, with *post hoc* tests performed using Tukey's HSD.

4.5 Results

4.5.1 Accuracy

The largest significant factor found for the accuracy of performance was the speed [$F(5, 35) = 39.83$, $p < 0.001$]. This factor accounted for most of the variance in accuracy ratings [$\text{eta}^2 = 0.66$]. *Post hoc* tests showed that each speed is significantly more accurate than any higher speed [Tukey's HSD, $p < 0.05$ in all cases].

Significant effects were also found for the method used [$F(2, 14) = 38.07$, $p < 0.001$] and the method $\times$ speed interaction [$F(10, 70) = 3.95$, $p < 0.001$]. Specifically, the pressing and rolling methods were both significantly more accurate than the sliding method [Tukey's HSD, $p < 0.01$ for both methods], but were not significantly different from each other. *Post hoc* tests on the interaction showed significant differences between low speed (≤ 3 Hz) vibrato using any method and high speed (> 3 Hz) vibrato using any method [Tukey's HSD, $p < 0.05$].

Figure 3 shows the mean achieved frequency versus the target frequency for each method. Note the increased distance from the target line for each method as the target frequency increases.

Examining the effect of method at each speed separately showed that at low speeds there was no significant difference between the accuracy of each method. At high speeds however (> 3 Hz), both the pressing and rolling methods prove to be significantly more accurate than the sliding method [Tukey's HSD, $p < 0.05$ for each comparison]. This means that using these methods the participants' achieved speeds were closer to the target speed than when using the sliding method. There were no significant differences found between the pressing and rolling methods at any speed.

4.5.2 Stability

The largest significant factor found to effect stability is again that of speed [$F(5, 35) = 55.69$, $p < 0.001$, $\text{eta}^2 = 0.75$]. Again, *post hoc* tests showed that performance at any speed was significantly more stable than performance at higher speeds [Tukey's HSD, $p < 0.05$ for all comparisons]. We also found significant effects of method of vibrato production [$F(2, 14) = 9.24$, $p < 0.01$]. Interestingly, in this case the sliding method proved to be significantly more stable than either the pressing or rolling methods [Tukey's HSD, pressing: $p < 0.01$, rolling: $p < 0.01$]. Once again the scores for the pressing and rolling methods were not significantly different. There was also a significant effect of the method × speed interaction [$F(10, 70) = 3.12$, $p < 0.01$]. As with the accuracy measurement, the method × speed interaction indicates that low speed using any method is more stable than high speed using any method [Tukey's HSD, $p < 0.05$ for all comparisons].

We also found no significant differences between any of the methods when controlling low speed (≤3 Hz) modulations. Above 3 Hz we found significant differences between the stability of the sliding method and that of the other two methods. In this case however, the sliding method is significantly more stable than the other methods [Tukey's HSD, $p < 0.05$].

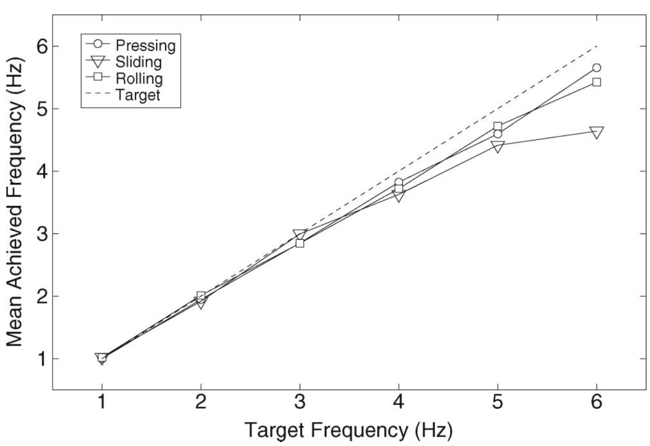

Fig. 3. Mean achieved frequency versus target frequency for each method.

Figure 4 shows the deviation from the mean achieved frequency for each method at each of the target frequencies. A lower deviation indicates a higher level of stability. It can clearly be seen that at higher frequencies the sliding method is more stable than the other methods [$p < 0.01$].

4.5.3 Modulation depth

Examining the depth, we again found a significant effect of speed [$F(5, 35) = 10.39$, $p < 0.001$] and of the method × speed interaction [$F(10, 70) = 2.09$, $p < 0.05$]. There was also a marginally significant effect of method [$F(2, 14) = 4.20$, $p = 0.08$].

Post hoc tests showed that higher speed modulations generally had a lower depth than low speed modulations. Modulations at frequencies of ≤3 Hz had a significantly higher depth than those of >3 Hz [Tukey's HSD, $p < 0.05$ for all comparisons].

Examining the modulation depth for each method depending on speed, we found no significant effect of speed on depth for the pressing method [$F(5, 35) = 1.47$, $p = 0.25$] or for the rolling method [$F(5, 35) = 0.96$, $p = 0.45$]. While there is a decrease in depth as speed rises, it is not a significant decrease. The sliding method on the other hand showed a significant difference in depth due to speed [$F(5, 35) = 16.50$, $p < 0.001$]. *Post hoc* tests showed significant differences in depth between speeds below 4 Hz and speeds from 4 Hz upwards [Tukey's HSD, $p < 0.05$ for all comparisons]. Figure 5 shows the modulation depth as a function of frequency for each of the mappings.

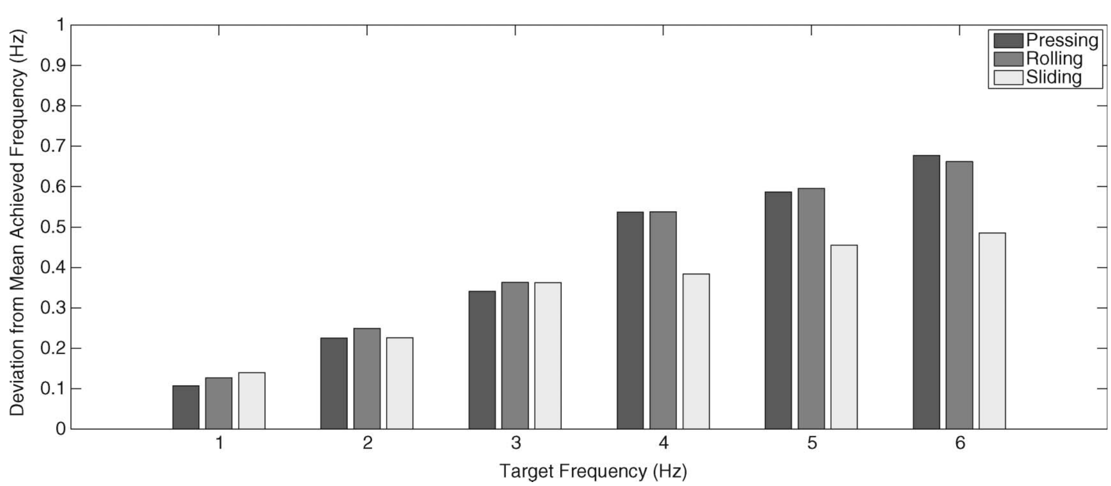

Fig. 4. Deviation from the mean achieved frequency for each method at each of the target frequencies. A lower score indicates a higher level of stability.

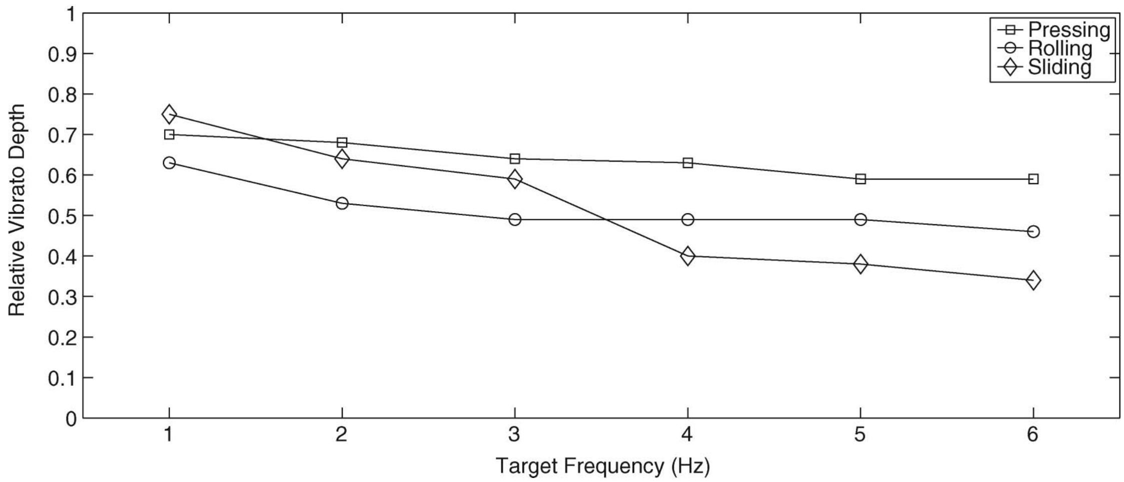

Fig. 5. Modulation depth for each mapping as a function of frequency, normalized over the range ± 1 semitone.

4.6 Discussion

As noted in the first experiment, the most significant effect on the stability of performance is that of the speed. Interestingly, this also holds for both accuracy and depth (although the power of this effect is much less for depth). This would indicate that when deciding on a control for a modulation task we must be aware that there is a decrease in performance at higher speeds. Also interesting to note is that there appears to be a cut-off point between 3 and 4 Hz, as accuracy, stability and depth all vary significantly above and below this point.

There also appears to be a link between the method used and the stability of the modulation being performed, as noted in the first experiment. In this experiment the sliding method was significantly more stable, although only for modulations above 3 Hz. At the same time, we can see a decrease in both depth and accuracy for the sliding method at these speeds.

Another possible reason for this could be the availability of intrinsic visual feedback for the sliding method when compared to the other methods. When sliding his finger back and forward over a distance it is possible for the participant to see where his finger is and to use this feedback to ensure that he is consistent in how he plays. As noted by Marshall and Wanderley (2006), this visual feedback is only available with the linear position sensor and only when moving over a distance along the sensor (as is the case with the sliding method but not with the rolling method).

The decrease in modulation depth at higher speeds for the sliding method may also be the result of the mechanics of the hand/arm movement. In order to increase the rate of the modulation while still maintaining some control over it, the performers must decrease the magnitude of the sliding movement. It is also possible that as the instructions for the experiment emphasized the speed of modulation as the primary interest, the participants may have purposely reduced the depth to allow them to concentrate on the speed. Further experimentation where the depth of the modulation is fixed could allow us to see if the accuracy of the sliding method would decrease even further and whether the stability of performance would also suffer.

Taken together, these results could indicate that the sliding method is suited to slower modulations. For instance, as already mentioned, most violin vibrato is in the 5–7 Hz range, a range for which the sliding method would not be suited due to its reduced accuracy and depth at this range.

Overall there were no significant differences between the pressing and rolling methods for any of the examined factors (stability, accuracy or depth). If we were to choose between those two methods for a modulation of pitch in an interface then performer's subjective preference ratings (such as those in the first experiment or the previous work already discussed) would seem to be a good indicator of which is the most suited.

Finally, it should be noted that the reduced stability of the pressing and rolling methods at higher frequencies could be compensated for by the performer and might disappear with practice. Again, this provides a possible area for further research.

5. General discussion

Overall, we have seen that both subjective ratings (such as performer preference ratings) and objective measures (such as stability and accuracy) can offer an indication of the suitability of a gesture mapping for the modulation of a parameter such as frequency. However, it would seem that neither the objective nor the subjective measures are sufficient in and of themselves to be used as the sole guideline in choosing a sensor and gesture for the control of a parameter modulation. For example, there were no significant differences in user preference ratings between the sliding and rolling methods, yet there were significant differences in stability and accuracy between these methods. Conversely, there were no significant differences in stability and accuracy between the pressing and rolling methods, yet there were significant differences in user preference ratings for these sensors. To differentiate between these sensors then requires that we take into account both the subjective user ratings and the objective measurements.

In addition to this, we found that there are a number of other factors which affect both the subjective user ratings and the objective measurements, which must be taken into account when attempting to determine the optimal sensor for a specific task. Firstly, it would seem that the speed of the modulation has a direct effect on the results of our objective measures. At lower speeds all three methods of control performed very similarly across all three measures (stability, accuracy and depth). At higher speeds we find clear trade-offs. The sliding method is significantly more stable, but significantly less accurate and with a significant decrease in depth of the modulation. Both the rolling and pressing methods show themselves to be accurate and without a reduction in depth, but each of these are less stable than the sliding method.

Looking at the participant preference ratings from the first experiment, we can see that the participants significantly preferred the pressing method to both other methods of control. This is consistent with previous studies (Wanderley et al., 2000) and so would seem to be an important factor in the choice of which method to use for control of the modulation. Combining the preference ratings with the objective measurements would seem to indicate that the pressing method is the best method for control of modulation, given that it is more accurate, preferred by the participants and has no loss of depth at higher speeds. As already stated, the lower stability (compared to the sliding method) could perhaps be compensated for by the performer with practice. However, it is also possible that practice with any of the methods could affect the accuracy, stability and depth measurements.

The issue of training is also one which merits some discussion. As already mentioned, both the user ratings and the objective measurements could change with practice. However, for this study the goal was to evaluate the sensors specifically for novice users. The aim here is to determine which sensors would offer a 'low entry fee' (Wessel & Wright, 2002). An instrument designed using the methods discussed in this paper should be easy to use, not requiring much initial effort to learn.[8] However, there still remains the issue of ensuring that such an instrument offers much musical potential. An instrument which is too simple may seem toy-like and may only have a limited musical potential. Therefore, effort must be made to ensure that these instruments offer 'No ceiling on virtuosity' (Wessel & Wright, 2002). One method of doing this could be the use of complex mappings such as those discussed by Hunt (2000). By having some simple musical functions (such as pitch selection and modulation) controlled by gestures determined using the methods described here and more complex musical functions (such as timbral modifications) controlled by the interaction of various gestural parameters it may be possible to create an instrument which is both easy to use and also offers much potential for musical creation and exploration.

[8]The issue of initial ease of use is also particularly important for interfaces which are designed for use by casual users. One example of such an instrument is the reacTable* (Jorda et al., 2005).

6. Conclusion

The work described in this paper has attempted to empirically determine the suitability of sensors for specific tasks in a musical interface. The results of our work indicate that a combination of user preference ratings and objective measurements can determine which sensors should be used to control a specific task.

Interestingly, we found no link between previously learned performance gestures (such as the production of vibrato on a violin) and preference ratings or performance of pitch modulation with novel mappings. However, we did find that pianists gave significantly higher preference ratings than violinists for producing modulations with their right hand, which is most likely due to the fact that right handed violinists produce vibrato on the violin with their left hand and so were uncomfortable performing modulations with their right hands. This is further illustrated by those violinists who asked to perform the rolling method with the sensor on their shoulder, in the position in which they normally play their instrument.

For the control of pitch modulation in particular we have found that at low speeds objective measurements of accuracy, stability and depth are insufficient as no significant differences were found across methods using these measurements. This means that if an interface requires only low-frequency modulations, we can be free to choose any of the methods discussed here.

At higher speeds we see a trade-off between stability on one hand and accuracy and depth on the other hand. With the sliding method participants are more stable, but are significantly less accurate and also produce significantly shallower vibrato. This could indicate that the sliding method is less suited for use at higher speeds, unless stability of control is the highest priority. For the other methods, accuracy and depth are maintained at higher speeds, but with a loss of stability.

Coupled with the participant preference ratings, which indicated a preference for the pressing method, these results indicate that the pressing method is the best control of pitch modulation, particularly over a wider range of speeds. The fact that this is in line with previous findings would also seem to be a positive indicator.

Together, these results show that for computer music interfaces it is possible to derive a measure of sensor suitability for a specific task, so long as we take into account both the subjective measurements of user preference and the more objective measurements of accuracy and stability in performance.

Acknowledgements

The authors would like to thank Bertrand Scherrer for invaluable input on the analysis of the recorded data and Vincent Verfaille for detailed discussions regarding vibrato analysis and the use of his Matlab code for parameter measurement.

This work was partially funded by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC)/Canada Council for the Arts (CCA) for the first and third authors, as well as by NSERC discovery grants to the last two authors.

A version of some sections of this paper was submitted in partial fulfilment for the Bachelor's Degree with Honours at McGill University for the second author.

References

Allison, J.T., & Place, T.A. (2005). Teabox: A sensor data interface system. In *Proceedings of the 2005 conference on New Interfaces for Musical Expression (NIME05)*, Vancouver, Canada, pp. 56–59.

Bongers, B. (2000). Physical interfaces in the electronic arts: Interaction theory and interfacing techniques for real-time performance. In M.M. Wanderley & M. Battier (Eds.), *Trends in Gestural Control of Music* (pp. 124–164). Paris, France: IRCAM – Centre Georges Pompidou.

Buxton, W.A.S. (1987). The haptic channel. In R.M. Baecker & W.A.S. Buxton (Eds.), *Readings in Human-computer Interaction: A Multidisciplinary Approach* (pp. 357–365). San Mateo, CA: Morgan Kaufmann.

Cuddy, L.L., Balkwille, L.-L., Peretz, I., & Holden, R.R. (2005). Musical difficulties are rare: a study of 'tone deafness' among university students. *Annals of the New York Academy of Sciences*, *1060*, 311–324.

Hunt, A. (2000). *Radical user interfaces for real-time musical control* (Doctoral dissertation). University of York, UK.

Isaacs, D. (2003). *Evaluating input devices for musical expression* (B. Inf. Tech. (Hons.) Thesis). University of Queensland, Australia.

Jorda, S., Kaltenbrunner, M., Geiger, G., & Bencina, R. (2005). The reacTable*. In *Proceedings of the 2005 International Computer Music Conference (ICMC05)*, Barcelona, Spain.

Levitin, D., McAdams, S., & Adams, R.L. (2002). Control parameters for musical instruments: A foundation for new mappings of gesture to sound. *Organised Sound*, *7*(2), 171–189.

Marchand, S., & Raspaud, M. (2004). Enhanced time-stretching using order-2 sinusoidal modeling. In *Proceedings of the Digital Audio Effects (DAFX04) Conference*, Naples, Italy, pp. 76–82.

Marshall, M.T. (2009). *Physical interface design for digital musical instruments* (PhD thesis). McGill University, Canada.

Marshall, M.T., & Wanderley, M.M. (2006). Evaluation of sensors as input devices for computer music interfaces. In R. Kronland-Martinet, T. Voinier & S. Ystad (Eds.), *CMMR 2005 – Proceedings of Computer Music Modelling and Retrieval 2005 conference: LNCS 3902* (pp. 130–139). Berlin, Heidelberg: Springer-Verlag.

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

Norman, D. (2002). *The Design of Everyday Things*. New York: Basic Books (Perseus).

Oldfield, R. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Papich, G., & Rainbow, E. (1974). A pilot study of performance practices of twentieth-century musicians. *Journal of Research in Music Education*, *22*(1), 24–34.

Vertegaal, R., Ungvary, T., & Kieslinger, M. (1996). Towards a musician's cockpit: Transducers, feedback and musical function. In *Proceedings of the 1996 International Computer Music Conference (ICMC96)*, San Francisco, CA, pp. 308–310.

Wanderley, M.M., & Orio, N. (2002). Evaluation of input devices for musical expression: Borrowing tools from HCI. *Computer Music Journal*, *26*(3), 62–67.

Wanderley, M.M., Viollet, J.-P., Isart, F., & Rodet, X. (2000). On the choice of transducer technologies for specific musical functions. In *Proceedings of the 2000 International Computer Music Conference (ICMC2000)*, Berlin, Germany, pp. 244–247.

Wessel, D., & Wright, M. (2002). Problems and prospects for intimate musical control of computers. *Computer Music Journal*, *26*(3), 11–22.