Animated performance: 'Better' music means larger movements

Caitlyn Trevor & David Huron

ABSTRACT:

A study was carried out to determine the effect of the amount of performer movement on judgments of performance quality. The movements of eight live solo performances were captured (two each for flute, clarinet, violin, and cello). From each original recording, three stick figure animations were created: one with augmented performance motion, one with the original motion, and one with diminished motion. The three animations were combined into single dynamic videos that allowed participants to continuously adjust the range of motion in the animation via a slider — from diminished through original to augmented motion. Participants were instructed to adjust the overall amount of performance motion to create the best musical performance. Consistent with the hypothesis, participants elected to significantly augment the motions of the performers.

I. INTRODUCTION

Among humans, the visual modality tends to dominate the auditory modality in many stimulus contexts. Although music is nominally a purely sonic art, extant research suggests that visual information plays an important role in nominally auditory perceptions. Of course, musicians have long known that performer attractiveness, dress, and deportment can influence how a performance is received. Wapnick *et al.* (1997, 1998), for example, found that the attractiveness of a performer can influence judgments of performance quality. Goldin and Rouse (1997) assembled data pertaining to the auditions of orchestral musicians — comparing those auditions carried out with and without the use of an occluding screen. Their analyses demonstrated a marked gender bias against female performers in non-blind auditions.

Apart from deportment, gender, and attractiveness, however, other visual cues appear to be in play, such as performer motion. Music theorists, music educators, and musicologists have long discussed gesture and motion in music (Berry 2009; Fisher and Lochhead 2002; Gritten and King 2011; Hatten 2004; Larson 2012; McCreless 2006; Montague 2012). In a series of experimental studies, Tsay (2013) asked participants to predict the winners from the three finalists from each of ten concerto competitions. Participants were exposed to three conditions: sound-only, video-only, and combined audio-visual. Although participants insisted that sound alone is the most important criterion for judging musicality, participants were surprisingly most accurate in predicting competition winners in the video-only condition. A study by Morrison *et. al.* (2009) focused more specifically on the influence of conductor movement on performance ratings. They paired a single audio recording with videos of different conductors conducting the same work. Participants judged the sounded performance as more expressive when paired with more expressive conductors. Similarly, Juchniewicz (2008) compared ratings of a piano performances with amounts of expressive motion and showed that performance elements such as dynamics, phrasing, and overall musicality are rated higher when performer motion increases. These studies suggest that the amount of movement a performer employs may positively influence judgments of the performance quality.

Apart from these musical studies, other research suggests that visual information can generally take precedence over auditory information in a multimodal context. A well-known example is found in the McGurk effect where observing the facial movements of a speaker can dramatically override the perceived spoken sounds (McGurk and MacDonald 1976). Like the McGurk effect, some visual gestures appear to directly influence auditory perception in music. For example, in a study of multimodal perception, Schutz and Lipscomb (2007) recorded video and audio clips of contrasting articulations on marimba: a short sharp stroke, and a nominally longer legato stroke. When the videos of the strokes were switched with each other's audio recording, participants reported hearing longer or shorter tones based solely on the accompanying video clip. In short, the visual action of the mallet stroke altered what listeners heard.

Despite the widely held view by musicians that regards visual elements as secondary considerations, these studies nevertheless suggest a preeminent role of visual components in nominally musical judgments. This raises the question of what visual factors are responsible for these effects. The work of Morrison *et al.* (2009) and Juchniewicz (2008) suggest that the amount of gross motion might be influential. In the current study, we follow up this notion by examining the potential influence of the magnitude of performer motion. Specifically, we test two hypotheses:

> H1. When given the opportunity to amplify or attenuate a visual rendering of a performance in order to enhance the musicality, participants will tend to augment the performer's movements.

In considering performer motion, one might suppose that different styles of music would affect the amount of performer motion both employed by the musicians and preferred by the participants. Because of the difficulty of a fast, technical passage, it is possible that a musician would tend to employ less expressive motion for fear of interfering with their technique. On the other hand, for a slower more lyrical passage, a performer might feel freer to add expressive motions. As outlined in the above scenario, we predicted that performers would move less during a faster, more technically challenging passage than while playing a slower, more lyrical passage. Consequently, we also propose to test the following second hypothesis:

H2. When given the opportunity to adjust a performer's movements, participants will prefer greater amounts of motion for lyrical passages compared with faster passages.

To anticipate the results, we will see that, on average, participants preferred to augment the motions of the performers.

II. METHODOLOGY

In brief, we recorded motion capture data from performances of four instrumental musicians. The performances were visualized as stick figure animations in order to remove visual factors other than gross movement. Stimuli were presented using a custom interface that allowed for a continuous adjustment of the magnitude of movement — from an extremely movement-attenuated version to an extremely movement-amplified version.

1. Creation of Stimuli

A. Musicians

Four musicians from Ohio State University, two doctoral performance candidates (clarinet and violin) and two faculty members (cello and flute), volunteered to be recorded for the materials used in the main study. In order to investigate a wide range of motion capabilities afforded by different instruments, we initially planned to capture one performance from each of the four main instrument families: a string instrument, a woodwind instrument, a brass instrument, and a percussion instrument. However, due to coordination difficulties, our pool of volunteer performers ultimately comprised two woodwind players and two string players. An important consideration is whether a performer is seated or stands. In our study, the cellist and clarinetists played in a seated position, whereas the violinist and flutist played standing. In order to test hypothesis 2, musicians were instructed to bring two, already-learned musical passages that were each about one minute in length. Specifically, they were asked to bring a slower, more lyrical passage, and a faster, more technically challenging passage.

B. Recording Procedure

Each of the four musicians wore a motion capture suit to which small, spherical markers were applied in conventional anatomical locations (as shown in Figure 1). Some experimentation was conducted in order to optimize placement of finger markers.

Figure 1. Experimenter in the motion capture suit and markers.

Due to their close proximity, fingers are particularly challenging parts of the body to track using a motion capture system (in this case VICON Blade). Unfortunately, for these smaller movements, the system often confuses the markers, mislabeling fingers and joints. To avoid using the gloves and in an attempt to yield a more refined capture of finger motions, we attempted to use make-up tape in order to attach miniature markers directly onto their hands and fingers (as shown in Figure 2). However, for most of the musicians, these failed to stay attached while performing. Therefore, after testing to be sure that the gloves minimally interfered with performing, the gloves were ultimately used, plus two additional markers taped to the index fingers and the minimus (fifth) fingers. Markers were also attached to various points on each instrument in order to allow for the possibility of adding instruments to the animations.

C. Method of Animation

The data for each performance capture were filtered and imported into MotionBuilder (Catalogue n.d.) for animation (as shown in Figure 3). From the originals, a diminished motion version and an augmented motion version were created. This was achieved by hand-editing the performer's motions.

Figure 2. Examples of motion-capture markers affixed to the musicians' hands.

Two forms of performer motion might be conceptually distinguished. Primary motions are those forms of motion that are essential for sound production such as the movements of fingers, arms, lips, etc. Secondary motions are those forms of motion that appear to be not strictly necessary for sound production. For example, a cellist may sway her/his upper body from side to side or backwards and forwards. Only secondary motions were targeted by the authors for possible augmentation or diminution in the edited animations. For the augmented condition, these secondary performance motions were amplified. For the diminished condition, these motions were rendered with a smaller amplitude, often resulting in a more stationary spine. The most commonly altered motions included head motions, knee bends, and full-body swaying.

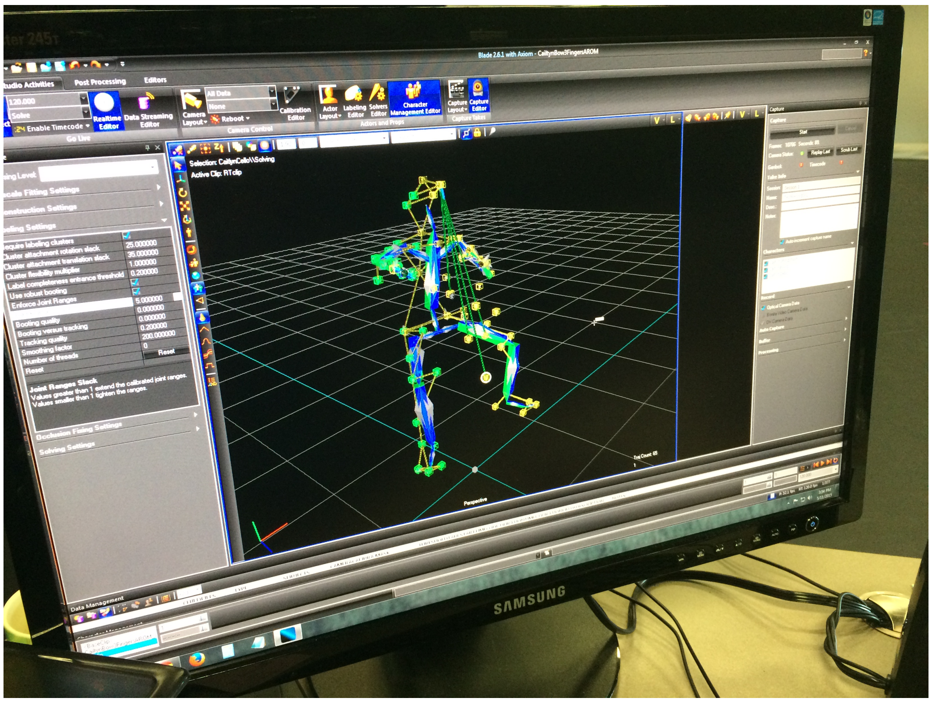

Figure 3. Sample motion capture data realized in MotionBuilder.

After creating these animated versions, each animated figure was fitted with white cylinders in a skeletal arrangement and rendered in a virtual environment with a wooden floor and a dark, neutral background (as shown in Figure 4). A skeleton model was chosen after attempts to fit more realistic models were found to reduce the apparent motions overall since some body parts were visually obscured. Additionally, a major motivation for using motion capture in this study was that it eliminated all other physical influences, such as gender and attractiveness. A skeletal stick figure animation is arguably devoid of all these other visual aspects. Additionally, there were no chairs or instruments animated into the scenes. This decision was made based on the sheer difficulty of attempting to align the instruments and chair with the live capture data without having them collide with a limb or the main body. In post-experiment interviews, participants reported minimal aversion to the stick figures nor were there expressions of concern regarding the absence of instruments or chairs.

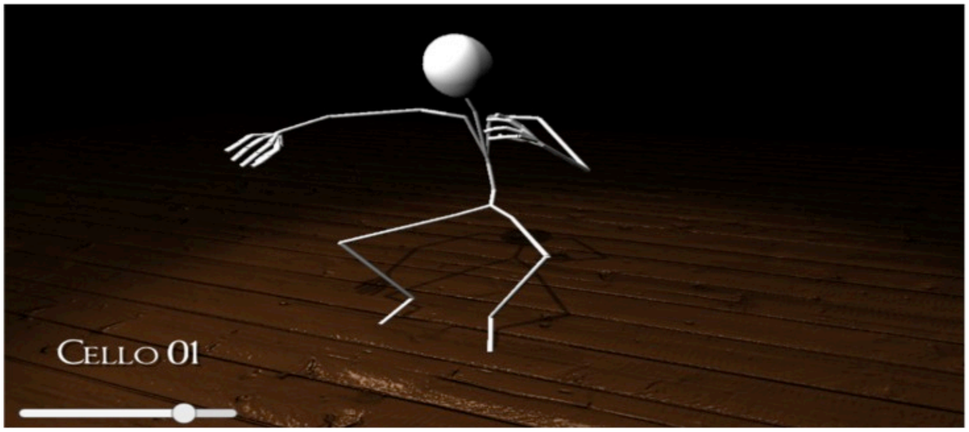

Figure 4. Sample (static) image from the animated performance stimuli. All performer movements were simplified to animated stick figures devoid of environmental, instrument and chair elements. Stick figure animation is likely to reduce possible confounding influences of gender and attractiveness features.

D. Interface

The interface was created using Unity (Goldstone 2009). The three versions of each performance (augmented, original, and diminished) were combined into one dynamic video that was adjustable by a horizontal slider. The effect of the slider was to interpolate the marked body positions between the three underlying animations so the resulting image changes are entirely smooth. This structure resulted in the illusion of being able to manipulate the overall amount of performance motion in real time. Slider orientation was randomized from trial to trial: in some trials the most diminished version was associated with the left-most position, whereas for other trials the behavior was reversed. In addition, in order to avoid possible anchoring effects from trial to trial, the initial position of the slider was randomized. Finally, the order of the musical performances was randomized for each participant in the experiment. Videos clips were presented in real time coordinated with the (un-manipulated) audio of the recorded performance. Notice that manipulating the slider had no effect whatsoever on the audio portion of the stimulus.

2. Experiment

A. Participants

There were 39 participants from the School of Music at Ohio State University. Participants ranged in age between 18-65.

B. Procedure

The procedure involved the method of adjustment where participants were asked to tune the animation slider. The task was described to participants by presenting instructions conveying the following hypothetical scenario.

> "You are going to see eight different performances. For each performance, we want you to consider the following scenario.
>
> Suppose there is a musical competition in which contestants are asked to submit a video performance. The rules of the competition forbid showing any information regarding the age or sex of the performer. Therefore, stick figure animations are encouraged so that the judges are able to see how the musician moves along with the sound.
>
> In preparing their video to submit for the competition, some performers realize that there is perhaps room to cheat a bit by manipulating the movement.
>
> With the slider at the bottom of each video clip you can manipulate the overall amount of movement in the performance. For each animated clip we want you to tune the slider to create what you think is the best performance.
>
> When you have adjusted all of the sliders, please press the submit button at the bottom of the page.
>
> Thank you for participating!"

In order to later test whether musical experience might have an effect on the results, the experiment began by having participants complete the Ollen Musical Sophistication Index (Ollen 2006). No *a priori* hypothesis was made regarding possible effects of musical sophistication.

III. RESULTS

Recall that our first hypothesis predicted that participants would adjust the overall amount of performer motion towards the augmented version. The data gathered from the slider (invisible to the participant) ranged numerically from 0.0 (diminished) to 0.5 (original) to 1.0 (augmented). In order to reject the null hypothesis, the mean must be significantly higher than 0.5. As the data are not normally distributed, a Wilcoxon signed-rank test was used rather than a simple *t*-test. The results proved to be statistically significant in the predicted direction ($M = 0.601$, $SD = 0.18$, $p = .002$).

Our second hypothesis predicted that participants would prefer more motion for lyrical passages than for technical passages. The results proved to be contrary to our hypothesis. For technical passages, the amount of motion was again greater than the original movement amplitude ($M = 0.662$, $SD = 0.19$, $p = 0.001$). However, for lyrical passages, the amount of motion did not differ significantly from the original movement ($M = 0.541$, $SD = 0.20$, $p = 0.20$). Note that in contrasting the technical performances with the lyrical performances, a paired Wilcoxon Signed-rank test of the difference proved significant ($p < .001$). That is, our participants preferred augmented movements for the fast technical passages while preferring only slightly more movement than normal for the slow lyrical passages.

Recall that we collected musicality data using the Ollen Musical Sophistication Index. Having no *a priori* hypothesis, we nevertheless examined whether the degree of musical sophistication of the participants influenced the results. Using Spearman's R, we found a significant negative correlation between the musical sophistication scores and the amount of preferred motion ($r = -0.40$, $p = 0.013$). This relationship is shown in Figure 5. Notice that the regression slope suggests that the most musically sophisticated participants preferred the magnitude of motion to be quite similar to the original performer motion. Said another way, the effect of augmented performer motion is most marked for those participants with the least sophisticated musical background — as measured using the Ollen instrument.

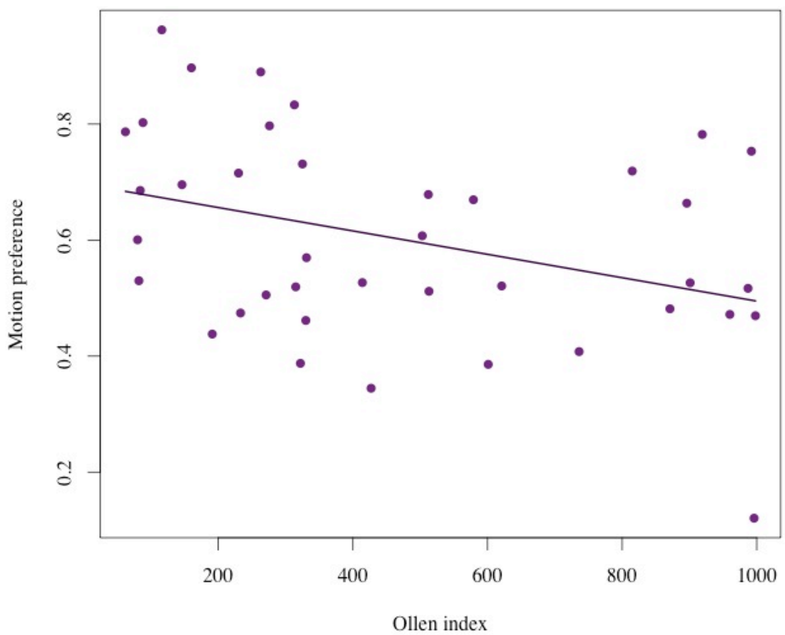

Figure 5. Relationship between musical sophistication and preferred performer motion. Less musically sophisticated participants (as measured using the OMSI index) preferred augmented or exaggerated motion. More musically sophisticated participants preferred the magnitude of motion to be quite similar to the original performer motion.

Finally, recall that we collected data for four distinct instruments. A series of *post hoc* analyses found that the amount of preferred movement was significantly higher for the cello ($p = .02$) and flute ($p = .04$) compared with the clarinet. However, due to the small number of participants involved in the creation of the stimuli, strong between-instrument conclusions are not warranted.

IV. DISCUSSION

The results of our study are consistent with our main hypothesis that participants judge the best performances to exhibit greater (exaggerated or augmented) secondary performance motions. These results corroborate findings by Morrison *et al.* (2009) and Juchniewicz (2008). On the other hand, our data are not consistent with our second hypothesis. That is, participants did not favor greater secondary performer motion for lyrical passages compared with technical passages.

Prima facie, these results suggest that our theory may be incorrect. At the same time, we might entertain a related alternative explanation for why the data may be skewed in a direction opposite to what was anticipated. When playing a technically challenging faster passage, it is reasonable to assume that musicians will tend to move less because they are focused on managing the notes as best they can. On the other hand, for a slower more lyrical passage where there is less technical demand, a musician might add more expressive motions because they can afford to. Perhaps the data are skewed towards higher adjustments for technical pieces because the performers moved less for those passages than for the lyrical passages. In other words, perhaps when adjusting motions for lyrical passages, participants felt less inclined to increase the motion since it was already fairly high.

Another consideration was whether the level of musical sophistication of the participants affected the responses. It is interesting to note that participants with lower musical sophistication scores adjusted the animations to produce more average motion than the more musically sophisticated participants. Perhaps this difference in responses is related to knowledge that a more musically sophisticated participant might have of technical restraints when adding expressive motion, thereby leading them to augment the motion less. This finding appears to be consistent with informal observations of audience-dependent performer behaviors. Popular musicians engage in far more stage movement than is found among classical musicians. However, even among classical musicians, it is often observable that musicians who appear to a more "popular" classical audience (such as Lang Lang and "2Cellos") move considerably more than classical musicians whose target audiences are less "popular" oriented. Of course, the cause and effect here may be reversed. That is, performers who move more may attract more popular audiences rather than performers who aim to attract popular audiences endeavoring to move more.

V. FUTURE RESEARCH

It would be of interest to further explore the physical visual elements of performance. Motion capture technology in particular provides a useful way to study performer motion with reduced confounding factors such as attractiveness, gender, race, etc. Additionally, animating the motion data allows for manipulation of the visual elements without altering the original audio recording of the performance.

Upon closer inspection of how the sliders manipulated the performer motions, there were sometimes marked changes in posture and head placement. For example, the cellist, when augmented, also adopted a more expansive, open posture and a raised head. Research on posture and eye contact indicate that open postures and good eye contact communicate friendliness, trustworthiness, and greater emotionality (Burgoon, Birk, and Pfau 1990; Sundaram and Webster 2000). Further future research might focus on the role of specific movement components in conveying performance information to viewers. Such studies may better account for the specific effects of head motions, eye contact, body posture, and other elements as forming a sort of affective vocabulary of performance motion.

ACKNOWLEDGMENT

Our thanks to the Center for Cognitive and Brain Sciences at Ohio State University for funding this research through a summer research grant. Additional thanks to our performers (Evan

Lynch, Mark Rudoff, Snow Shen, and Ann Stimson), our motion capture operators (Vita Berezina-Blackburn and Lakshika Udakandage), our unity app developer (John Luna), and the head of the Advance Computing Center for the Arts and Design, Maria Palazzi. Finally, we would like to thank members of Cognitive and Systematic Musicology Laboratory at Ohio State University for critical feedback and support.

REFERENCES

Berry, Michael. 2009. "The Importance of Bodily Gesture in Sofia Gubaidulina's Music for Low Strings." *Music Theory Online* 15: np.

Burgoon, Judee K., Thomas Birk, and Michael Pfau. 1990. "Nonverbal Behaviors, Persuasion, and Credibility." *Human communication research* 17 (1): 140–69.

———————. n. d. "Motion Builder: Making Motion Control Programming Easy." *Servo Catalogue*: 86–87.

Fisher, George, and Judy Lochhead. 2002. "Analyzing from the Body." *Theory and Practice* 27: 37–67.

Goldin, Claudia, and Cecilia Rouse. 1997. "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians" *National bureau of economic research* 5903.

Goldstone, Will. 2009. "Unity Game Development Essentials." *Packt Publishing Ltd.*

Gritten, Anthony, and Elaine King, eds. 2011. *New Perspectives on Music and Gesture.* Aldershot: Ashgate Publishing, Ltd.

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

Juchniewicz, Jay. 2008. "The Influence of Physical Movement on the Perception of Musical Performance." *Psychology of Music* 36 (4): 417–27.

Larson, Steve. 2012. *Musical Forces: Motion, Metaphor, and Meaning in Music.* Indiana University Press.

McCreless, Patrick. 2006. "Anatomy of a Gesture: From Davidovsky to Chopin and Back." *Approaches to Meaning in Music*: 11–40.

McGurk, Harry, and John MacDonald. 1976. "Hearing Lips and Seeing Voices." *Nature* 264: 746–8.

Montague, Eugene. 2012. "Instrumental Gesture in Chopin's Etude in A-flat Major, Op. 25, No. 1." *Music Theory Online* 18.

Morrison, Steven J., Harry E. Price, Carla G. Geiger, and Rachel A. Cornacchio. 2009. "The Effect of Conductor Expressivity on Ensemble Performance Evaluation." *Journal of Research in Music Education* 57 (1): 37–49.

Ollen, Joy E. 2006. "A Criterion-Related Validity Test of Selected Indicators of Musical Sophistication Using Expert Ratings." PhD diss., The Ohio State University.

Schutz, Michael, and Scott Lipscomb. 2007. "Hearing Gestures, Seeing Music: Vision Influences Perceived Tone Duration." *Perception*36 (6): 888–97.

Sundaram, D. S., and Cynthia Webster. 2000. "The Role of Nonverbal Communication in Service Encounters." *Journal of Services Marketing* 14 (5): 378–91.

Tsay, Chia-Jung. 2013. "Sight Over Sound in the Judgment of Music Performance." *Proceedings of the National Academy of Sciences*110 (36): 14580–5.

Wapnick, Joel, Alice Ann Darrow, Jolan Kovacs, and Lucinda Dalrymple. 1997. "Effects of Physical Attractiveness on Evaluation of Vocal Performance." *Journal of Research in Music Education* 45 (3): 470–9.

Wapnick, Joel, Jolan Kovacs Mazza, and Alice-Ann Darrow. 1998. "Effects of Performer Attractiveness, Stage Behavior, and Dress on Violin Performance Evaluation." *Journal of Research in Music Education* 46 (4): 510–21.