ABSTRACT
The Microsoft Kinect depth sensor could offer a convenient, markerless solution for quantifying the head and torso movements of pianists to examine the impact of somatic training on playing postures and movement. To assess the suitability of the Kinect for this application, we tracked four professional piano teachers performing scales immediately before and after a week-long workshop involving daily Feldenkrais Awareness through Movement (ATM) lessons. We compared Kinect skeletal tracking data with 2D reference data obtained simultaneously using Dartfish.

KEYWORDS
Kinect
Dartfish
posture
motion tracking
Feldenkrais
piano pedagogy

Assessing the suitability of Kinect for measuring the impact of a week-long Feldenkrais method workshop on pianists’ posture and movement

JILLIAN F. BEACON
School of Human Kinetics, University of Ottawa

GILLES COMEAU
School of Music, University of Ottawa

PIERRE PAYEUR
School of Electrical Engineering and Computer Science, University of Ottawa

DONALD RUSSELL
Department of Mechanical and Aerospace Engineering, Carleton University
video analysis software. Analysis revealed frequent tracking errors in the Kinect data compared to reference data from Dartfish. Differences in pre- and post-test measurements of forward head position, head height, C7 vertebra height and shoulder displacement did not correspond between Dartfish and Kinect. Our results suggest that one Kinect sensor does not provide enough accuracy to track torso movements of pianists for the purposes of ergonomic assessment in response to somatic training.

INTRODUCTION

The quality of pianists’ postural alignment may play a role in the development of playing-related musculoskeletal disorders (PRMDs) (Brandfonbrener 1997; Allsop and Ackland 2010), which have been estimated to affect between 39 per cent and 47 per cent of professional adult musicians (Zaza 1998). Ergonomic playing postures are also thought to facilitate fluency and control during performance, optimizing a performer’s ability to play expressively (Mark et al. 2003; Osada 2009; Wheatley-Brown et al. 2014). As a result, many musicians have turned to somatic training methods, such as the Alexander Technique (Alexander 1932), the Feldenkrais method (Feldenkrais 1981) or Body Mapping (Conable 2009), to improve their posture and increase body awareness during performance. These methods are designed to help individuals replace potentially harmful movement patterns with more healthful alternatives by improving kinaesthetic awareness and equipping them with more detailed knowledge of the body’s structure and function (Spire 1989; Conable 1995; Ginsburg 1999). Somatic methods have come to play a prominent role in music education as approaches to aid in injury prevention and rehabilitation, with many pedagogues gaining notoriety for their application of somatic principles when retraining injured musicians, including pianists (Taubman 1995; Mark et al. 2003; Fraser 2011; Daniel 2012; Stewart 2015). Somatic training methods are also included in the training programmes offered at many prestigious music education institutions and festivals. For instance, students at the Juilliard School can study the Alexander Technique with practitioner Lori Schiff, who is on faculty as a somatic trainer and trumpet instructor. She also teaches the Alexander Technique at the renowned Aspen Music Festival (Schiff 2014). Similarly, Stewart (2015) is regularly invited to teach Feldenkrais seminars and workshops at various music schools, notably the Mannes School of Music and the Juilliard School. She has also worked as an annual resident at the Marlboro Music Festival, which is artistically directed by international piano superstar, Mitsuko Uchida (Marlboro Music Festival 2015), and the Yellowbarn Music School and Festival, which is associated with the Manhattan School of Music. Despite the growing awareness and popularity of somatic training in professional music education, much of the evidence purporting improvements to performance quality or improvement of PRMDs is subjective and consists of practitioner-reported results (Rosenthal 1987; Mayers and Babits 1989; Nelson 1989) or student testimonials (Goldansky 2008; Stewart 2010; Boyd 2015; Fraser 2015; Johnson 2015). It is possible that motion tracking technology could provide researchers with means of objectively assessing pianists’ movements during performance to gain a more sophisticated understanding of how somatic training might help performers learn new movement patterns and habits of postural alignment that may contribute to their overall health and performance ability.
Human motion tracking with Kinect

Researchers have successfully used advanced optical motion tracking systems, such as Vicon or Qualisys, for accurate human motion tracking in rehabilitation research (Zhou and Hu 2008). However, these systems require anatomical markers to be positioned on subjects’ bodies, which could distract performers and potentially cause them to move atypically during testing. These optical-based systems are expensive to acquire and also require careful set-up and calibration, and extensive data extraction procedures (Balan et al. 2005). Many researchers with a background primarily in music may not have access to this equipment and may not possess the technological expertise to operate them and extract the data.

The Microsoft Kinect for Xbox 360 could offer music researchers a promising solution to the problem of accessible, reliable and unobtrusive motion tracking for assessing the impact of somatic training on posture in the context of piano playing. It was initially developed for gestural control in video games, circumventing the need for hand-held controllers and allowing for a more immersive gaming experience. The Kinect apparatus contains a regular RGB video camera and a depth-sensing camera that projects a dense array of structured infrared light points into the room to create a 3D image of objects in front of the sensor. It is equipped with software that identifies body parts by shape and tracks their location in three dimensions. The Kinect software predicts the likeliest position of the skeletal points it is searching for based on shapes detected by the infrared sensor instead of measuring the precise location of active or passive markers placed on participants’ bodies as would occur with optical systems. Detailed descriptions of the Kinect tracking process and the software operation can be found in the overviews of Duffy (2010) and Hadjakos (2012).

Many studies have compared Kinect tracking data with data simultaneously captured using 3D optical systems, such as Optitrack (Webster and Celik 2014), Optotrack (Tao et al. 2013), MediaLab (Fernández-Baena et al. 2012), Codemotion (Alnowami et al. 2012) and Vicon (Clark et al. 2012) to assess the accuracy of Kinect tracking. Although it is difficult to make a precise estimate of the Kinect accuracy due to the diversity of Kinect applications across studies, estimates suggest the Kinect is able to localize joint positions within 1–4cm at a distance of 1–4m (Obdržálek et al. 2012). Estimates for the accuracy of joint angle measurement are reported to be within five and thirteen degrees after major tracking errors are filtered out (Fernández-Baena et al. 2012). Although the measurement errors reported for Kinect are much greater than those of optical tracking systems, some researchers have found Kinect tracking to be adequate for assessing movement quality and posture in specific applications. For instance, Scano et al. (2014) compared joint position data of reaching movements tracked in the sagittal plane from the Kinect and a passive motion capture system. They found that measurement error was within an acceptable range for the assessment of upper-limb movement quality, and their results were precise enough to distinguish between subjects with Parkinson’s disease and healthy subjects. Similarly, Webster and Celik (2014) successfully used the Kinect to measure arm movements of stroke victims performing rehabilitative exercises. Kusaka et al. (2014) used the Kinect to measure arm joint angles of elderly people with hemiplegia to an accuracy of ten degrees to determine whether a therapeutic intervention improved arm mobility. Thus far, research demonstrates that the suitability of the Kinect as a quantitative measuring tool for human posture depends on the movement context and data collection procedures.
Pilot testing the Kinect with pianists

Hadjakos (2011, 2012) successfully used Kinect sensors to track the position of a pianist’s head, shoulders and arms from a perspective above the keyboard during virtuosic performance. We were therefore interested to know whether the Kinect would also be suitable for tracking pianists in the sagittal plane to help us assess the impact of somatic training interventions on pianists’ head and torso alignment. We conducted a pilot test to investigate the suitability of the Kinect in this application using software developed by engineers at the University of Ottawa and l’Université du Québec en Outaouais (Gauthier and Cretu 2014; Payeur et al. 2014). Software engineers modified the existing frontal-plane skeletal tracking model to track the pianists sitting in the sagittal plane for measuring posture variables pertaining to the vertical alignment of the head, shoulders, spine and hips. Results of our pilot test demonstrated that the average $x$, $y$ and $z$ coordinates of the head, shoulder centre, right shoulder and lower spine position tracked by the Kinect reflected expected differences in position when comparing the tracking data of exaggerated slouched or sway-backed postures with neutral postures during piano performances of scales and short musical phrases (Payeur et al. 2014). However, since this study did not compare the Kinect tracking results to a reliable reference measurement, it is unclear how closely the Kinect tracking reflected the actual movements of the pianist. Therefore, it is not possible to conclude whether or not the reported average differences reflect the detection of postural change, or artefacts of the tracking process. Furthermore, it is not clear from our pilot study whether the resolution of the Kinect allows for the detection of posture change in the head and shoulders in realistic performance situations, since the exaggerated postures modeled by our participant in the pilot study were more pronounced than the subtler changes to posture expected to result from somatic training interventions.

RESEARCH QUESTIONS

The present study aims to assess the suitability of the Kinect as a quantitative measurement tool for assessing the impact of somatic training on pianists’ posture by comparing the $x$- and $y$-axis components of 3D-Kinect tracking results with 2D reference coordinates measured from videos using Dartfish motion tracking software (Beacon 2015b; Dartfish 2015). Dartfish ProSuite software includes an object-tracking feature that follows pixels of a selected colour as a video plays. Researchers have used this feature to quantitatively measure posture for a variety of different purposes, including the assessment of sitting posture of subjects with postural backache (Womersley and May 2006), the assessment of a sit-and-reach test for hamstring flexibility (Mier 2011), the influence of neck pain on neck flexion during a reaching task (Constand and MacDermid 2013), the thoracic posture of rugby players (Bolton et al. 2013), standing posture in asthmatics after diaphragmatic and aerobic breathing training (Shaw and Shaw 2011) and comparing the impact of strength and stretch interventions in range of motion in dancers (Wyon et al. 2013). In an earlier study, we found that Dartfish video–based motion analysis is capable of measuring anatomical markers on pianists’ bodies within 0.5cm (Beacon 2015a). Our research seeks to answer two main questions:

1. How well do time plots of $x$- and $y$-axis coordinate data tracked by the Kinect match reference plots obtained using Dartfish when tracking live pianists?
2. Do Kinect tracking results reveal differences in post-test sagittal-plane posture variables of pianists that agree with reference results obtained using Dartfish?

In response to these questions, we present the following hypotheses:

1. We hypothesize that the time plots of coordinate data from the Kinect will contain frequent tracking errors in comparison with Dartfish, since the time plots in our initial pilot study with the Kinect displayed evidence of multiple tracking errors (Payeur et al. 2014).
2. We hypothesize that differences in posture variables will be measurable for some individuals, since a week-long exposure to somatic training is expected to impact posture and movement habits (Beacon 2015b).

We compared only the x and y coordinates between the two systems when assessing tracking quality for anatomical positions tracked by both systems to answer question 1, since Dartfish is a 2D tracking tool and the Kinect tracks in 3D. For question 2, we addressed this issue by asking participants to perform a playing task that did not require torso movement in a third plane, towards and away from the camera. This allowed us to measure sagittal plane posture variables using Kinect skeletal coordinates visible on the right side of the pianists’ bodies for comparison with measurements from Dartfish’s 2D coordinate data.

**METHODOLOGY**

**Design**

This study uses a repeated-measures design to track four pianists’ head, shoulder, right arm and spine movements during the performance of scales before and after participating in a week-long workshop with pianist and Feldenkrais practitioner Alan Fraser that applied the Feldenkrais method to piano movement and posture. We compared sagittal plane variables taken from the 3D Kinect tracking data to benchmark 2D coordinate data obtained using Dartfish video–based motion tracking software.

**Participants**

We recruited four professional piano teachers (three females, one male; ages = 24, 29, 50 and 51) from among the workshop participants. All participants had achieved a minimum of a bachelor degree in music studying piano. Two had attended prior institutes of Professor Fraser, and two had no previous experience with the Feldenkrais method.

**Playing requirements**

We asked participants to perform four repetitions of a C major contrary-motion scale in sixteenth notes at approximately 80 beats per minute, starting on C4 and extending to the lowest and highest octave of the piano. We chose this test because it requires symmetrical movements on both sides of the body and does not require torso rotation. Furthermore, during this exercise, pianists primarily sway their torso toward and away from the piano bench in the x and y axes while their bodies remain almost stationary with respect to distance from the camera (the z-axis), permitting meaningful comparison of the 2D
data from Dartfish with the 3D data from the Kinect for sagittal plane posture variables of the head and torso.

**Experimental set-up**

**Anatomical markers**

We placed red Kinotape markers on participants’ right ear-tragus, right acromion process (top of shoulder), right olecranon process (elbow) and right ulnar styloid process (wrist) prior to each recording session to permit accurate tracking of skeletal positions with Dartfish. We marked the C7 vertebral process with a spherical white marker fastened securely with medical tape. We provided participants with a tight-fitting sports top to prevent loose clothing from occluding any markers. A medical student placed all markers to ensure accurate and consistent placement between pre- and post-test recording sessions.

**Apparatus**

We recorded video with a Sony HD HandyCam (HDR-XR260V, 8.9 megapixels) set to record at a frame rate of 60i (capturing 30 frames per second). We mounted it on a Manfrotto tripod positioned at an appropriate height and distance for each participant. We used a Kinect for XBox 360, equipped with an infrared depth-camera (640 x 480 pixels, 30 images per second) and an RGB camera (1280 x 1024 pixels, ten images per second) to track the pianists’ movement. The Kinetics’ original motion-capture software platform was modified to track pianists from the sagittal viewpoint (Gauthier and Cretu 2014; Payeur et al. 2014).

We positioned the video camera perpendicular to the right shoulder of the pianist and levelled it. We adjusted the distance and height of the camera for each pianist and maintained the participant-specific camera positions for the post-test. We also maintained the participant-specific height of the piano bench and the distance of the bench from the piano. We placed the Kinect at approximately a 45-degree angle to the front and right of participants. Although this approximate Kinect with respect to the subject had to be adjusted by a distance of up to 30cm to help facilitate tracking initiation, which did not occur immediately in some of the tests.

**Procedure**

We collected all data at the Piano Pedagogy Research Laboratory at the University of Ottawa, and all participants provided consent prior to data collection. We recorded the pre-test scale performance before activities began on the first morning of the workshop. A research assistant began recording with the Kinect the moment the pianist began playing the exercise and stopped Kinect recording as they finished playing the last note. Another assistant operated the video camera. Subsequently, each participant attended a one-hour-long piano lesson and a one-hour-long Feldenkrais Awareness Through Movement (ATM) group lesson on each of the six days of the workshop. Participants were also invited to attend lectures on piano technique and to observe other students’ piano lessons. The piano lessons and lectures explored how the Feldenkrais method could be applied to piano technique and posture. We reapplied the anatomical markers and re-recorded participants playing the
post-test scale performance after they completed the final ATM lesson on the last day of the workshop.

**Measurement**

**Kinect tracking procedure**

The Kinect automatically generated \( x \), \( y \) and \( z \) coordinates of the existing Kinect skeletal model (including positions for the head, shoulder-centre, right shoulder, right elbow and right wrist) and exported them into Excel files for analysis. The direction of the axes of the coordinate system and an example of how the Kinect skeleton points would correspond to a participant’s body is depicted in Figure 1.

**Dartfish tracking procedure**

The videos were cropped to begin and end with the first and last note of the scale performances. Two experienced users tracked all anatomical markers in the videos using Dartfish TeamPro software, version 7.0, according to the procedure described in our initial Dartfish study (Beacon 2015a). They set the reference distance to the diameter of the lowest ball marker on the spine (3.7cm), and they marked the origin of the coordinate system at a stationary point visibly marked on the piano bench behind the participants.

**Analysis**

**Comparing Dartfish and Kinect time plots**

We compared the Kinect and Dartfish time-plots of the \( x \) and \( y \) coordinates of the body positions described in Table 1 to visually assess how well the Kinect tracked the various anatomical points in the \( x \) and \( y \) planes compared

![Figure 1: Left: Kinect skeletal tracking points overlaid on a performing participant. Right: Directions of axes for coordinate system. The origin of the coordinate system is positioned at the Kinect sensor camera, with the z-axis increasing away from the sensor, the x-axis increasing to the left of the sensor (when standing behind the sensor facing the same direction as the camera) and the y-axis increasing upward (Microsoft 2015).](image-url)
to Dartfish. We rated the tracking performance of the Kinect for each anatomical position based on how closely it matched the movement pattern depicted in the Dartfish reference plots using the rating-scale in Table 2. After carefully examining all 80 time plots, we identified four categories of tracking quality based on the types of tracking errors observed.

Comparing average measurement of posture variables

We calculated the pre- and post-test average values for the single-plane postural variables listed in Table 3 from the Kinect and Dartfish $x$ and $y$ coordinate data. We subtracted the pre-test averages from the post-test averages to compare the between-session posture differences measured by the Kinect against the reference measurements attained using Dartfish. Since the Kinect predicts skeletal positions using algorithms containing anthropomorphic information about the size and orientation of skeletal elements based on the depth-map image it produces, the Kinect body positions cannot be considered to be in precisely the same anatomical positions as the markers positioned for Dartfish. For instance, the shoulder centre projected by the Kinect is a few centimeters lower than the C7 vertebra used in standard calculations of forward head position as depicted in Figure 2, and therefore the average forward head angles are expected to be larger from the Kinect data. However, assuming that the Kinect head coordinate reflects the true centre of the participant’s head, any significant difference in average forward head posture apparent in the angles calculated from Dartfish data should also be reflected in the angles calculated from the Kinect coordinates.

RESULTS

Research question 1: Comparing Dartfish and Kinect time-plots

Figure 3 illustrates the distribution of $x$ and $y$ plots according to tracking quality as rated using the scale presented in Table 3. The Kinect appears to track at an ‘excellent’ or ‘good’ level more frequently in the $x$-axis compared to the $y$-axis. Out of a total of 40 tracking sessions, 25 $y$-axis plots qualified as ‘poor’
Assessing the suitability of Kinect for measuring the impact...

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
<th>Visual example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Excellent</td>
<td>Kinect plot is nearly identical to the Dartfish tracking plot</td>
<td><img src="image1" alt="Excellent" /></td>
</tr>
<tr>
<td>2. Good</td>
<td>The Kinect plot clearly depicts a pattern similar in shape and magnitude to the Dartfish plot, but contains up to two large tracking errors</td>
<td><img src="image2" alt="Good" /></td>
</tr>
<tr>
<td>3. Average</td>
<td>Dartfish pattern is partially visible in the Kinect plot, but contains 3–5 significant diversions or displays a magnified range of motion</td>
<td><img src="image3" alt="Average" /></td>
</tr>
<tr>
<td>4. Poor</td>
<td>Dartfish pattern is invisible in the Kinect plot and/or Kinect tracking was incomplete and/or the object was frequently lost during Kinect tracking</td>
<td><img src="image4" alt="Poor" /></td>
</tr>
</tbody>
</table>

The time-plots used in the examples above were taken from the following tests in descending order: (1) Participant 4 x-axis coordinates of head during session 1, (2) Participant 3 y-axis coordinates of elbow in session 2, (3) Participant 2 x-axis coordinates of elbow in session 2 and (4) Participant 3 y-axis coordinate of elbow in session 1.

Table 2: Qualitative rating scale for assessing the Kinect tracking performance over Dartfish solution.

tracking examples. No y-axis plots were rated as excellent, and only four were rated as ‘good’. Only x-axis plots received ‘excellent’ rankings, but 60 per cent of the x-axis tests were categorized as ‘average’ or ‘poor’, with the remaining 40 per cent split evenly between ‘excellent’ and ‘good’. Table 4 illustrates that some skeletal points were reliably tracked more often than others. For example, the
head centre/ear tragus and C7/shoulder centre positions were rated ‘good’ or ‘excellent’ more frequently than the elbow, wrist and right shoulder. Tracking of the right shoulder was the least reliable, receiving only ‘average’ or ‘poor’ ratings. These results suggest that the quality of Kinect tracking varies according to the skeletal position and the plane of motion being observed.

Although the majority of Kinect tracking sessions qualified as poor in comparison with Dartfish, a few of the $x$-axis time plots of the head and shoulder centre closely match Dartfish performance (Figure 4). This indicates that Kinect is sometimes capable of producing reliable tracking data in our testing context, but it is unclear from our study as to why the Kinect appeared to track reliably for some sessions while tracking very poorly for most others.

**Research question 2: Comparing pre- and post-test sagittal plane posture variable measurements**

**Forward head angle**

Table 5 presents the average angle of forward head position for each participant in each recording session, along with the difference in angle size between the two sessions. Results indicate that the Kinect and Dartfish tracking data did not yield similar results for the difference in average pre- and post-test forward head angle measurements. According to our previous reliability testing (Beacon 2015a), differences in angle measurements need to be greater than 2.5 degrees to be considered outside of measurement error for Dartfish measurements. Only participant 4 exhibited a change well outside
the measurement error for angles in Dartfish, with a decrease in forward head angle of 5.0 degrees from session 1 to session 2. This strongly suggests that participant 4 held his head farther forward in the post-test. The Kinect reported a 19.1 degree decrease in forward head position from pre-test to post-test for the same participant, which would suggest the opposite result — that he held his head much more erectly in the post-test. As can be seen in Table 6, even when the horizontal head displacement on the x-axis is considered separately from the y-axis, (considering that the time-plot comparisons from research question one revealed tracking to be less consistent in the y-axis), the changes from session 1 to session 2 are still not similar in either magnitude or direction between the two tracking methods. Although the results from Dartfish indicate that measurable differences in posture variables may be
observable for some participants after a week of somatic training, the Kinect data are not in agreement as to the direction and magnitude of the changes.

**Head and C7 height**

Table 7 presents the differences in pre- and post-test average head and C7 height above the hips/bench as measured by Kinect and Dartfish. Results

<table>
<thead>
<tr>
<th>Position</th>
<th>Excellent</th>
<th>Good</th>
<th>Average</th>
<th>Poor</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>C7 (shoulder centre)</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>Elbow</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Wrist</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>Right shoulder</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>9</td>
<td>16</td>
</tr>
</tbody>
</table>

*Table 4: Distribution of Kinect tracking performance ratings for different body positions.*

*Figure 4: Best-case tracking scenarios for Kinect motion tracking of the head during scale performance.*

observable for some participants after a week of somatic training, the Kinect data are not in agreement as to the direction and magnitude of the changes.
indicate that the differences in pre- and post-test averages do not correspond between the two sets of data. As shown in Table 7, the Dartfish values for participants 2 and 4 appear to report similar changes in average height from pre-test to post-test in both C7 and the ear-tragus, and that these differences are of a magnitude outside of the range of measurement error of 0.5cm. Therefore, it is likely that each of these participants displayed a true decrease in head height and C7 height from session 1 to session 2. The Kinect values do not seem to correspond with the Dartfish data. In fact, the Kinect data appears to indicate that participant 2’s torso extended upward by about 5cm instead of lowering by about 5cm as reported by Dartfish.
DISCUSSION

**Research question 1: Comparing Dartfish and Kinect time-plots**

As hypothesized, 2D tracking of the head, shoulder centre, right shoulder, elbow and wrist using the Kinect was less reliable and of inferior quality compared to the Dartfish video–based tracking system. However, comparing Dartfish and Kinect time-plots of anatomical positions provided specific information on how Kinect performance varied according to the skeletal position and axis of motion of the body part being tracked. The following list outlines the most common tracking errors that we observed when comparing Dartfish and Kinect time-plots:

**Frequent momentary loss of tracking**

The Kinect frequently lost track of skeletal positions momentarily and often reported extremely high or low values that appeared as ‘noise’ on the time-plots, as pictured in the upper-right plot in Figure 5. Frequent occurrences of these uncharacteristically high or low values corrupt data, preventing researchers from accurately measuring posture or examining trends in posture change over time.

**Amplification of movement magnitude**

The coordinates reported by the Kinect often reflected the appropriate direction of the movement, but at a greatly amplified magnitude. For example, the range of motion for the horizontal movements of the elbow for participant 2 in Figure 5 ranges from 5 to 6cm in the Dartfish plots, whereas the same motion appears to have a range of 18–23cm in the simultaneous Kinect tracking plot. This type of amplification can make it difficult to comprehend the true range of the motion using Kinect tracking data.

![Figure 5: (a–d) Examples of common Kinect tracking errors.](image-url)
Unrealistic pose-estimation

Sometimes the Kinect reported skeletal coordinates that represent a movement that would be anatomically impossible to perform in reality. Furthermore, reported coordinates often suggest unrealistic body-segment lengths. For example, in Figure 5, it can be seen that participant 1’s head suddenly increases in its average height by 22cm. In this example, it is possible that the Kinect mistook a lower part of the torso, such as the shoulders, for the head at first, but located the real head at the appropriate height further into the scale performance. This illustrates the possibility that values reported from the Kinect as belonging to a specific body position might actually have been tracked from an entirely different position.

Tracking errors appear as movement

The Kinect continuously updates its prediction of the location and size of a body segment during tracking, even if the object is stationary. This results in plots similar to the example on the bottom left of Figure 5. Here, the Dartfish plot illustrates that participant 2’s shoulder moved minimally in the y-axis, but the corresponding Kinect plot makes it appear that the shoulder moved up and down through an expansive range of about 15cm. Examples like these make it difficult to trust that the data from the Kinect reflects the true movements of participants.

Future researchers using the Kinect to quantitatively assess posture would need to address these types of tracking errors by using more advanced technology or by further customization of Kinect software. For instance, software could be programmed to filter out unrealistically high or low values generated from momentary loss of tracking to smooth out the plots and improve the validity of calculations of average values, as has been done in other studies using the Kinect to measure posture (Clark et al. 2012; Fernández-Baena et al. 2012; Tao et al. 2013). A potential solution to the problem of movement magnification would be to apply a scaling factor to the movements of specific body positions that tend to get amplified by the Kinect (Obdržálek et al. 2012; Clark et al. 2012; Tao et al. 2013).

Research question 2: Kinect suitability for assessing posture changes

Although we observed occasional examples of accurate head and shoulder tracking using the Kinect, ultimately we were unable to use coordinate data derived from the Kinect to measure changes in sagittal plane posture measurements. The Dartfish reference measurements reporting significant decreases in head height by as much as 5cm (well outside the range of measurement error for the tool), for three participants suggests that the height of the head and upper back were measurably different from session 1 to session 2 for some individuals. Kinect coordinates did not report posture changes in the same direction or magnitude. It can therefore be concluded that one Kinect for Xbox 360 is not an adequate tracking tool for measuring changes in posture variables over time in the context of piano playing.

Most of the posture variables of interest when assessing piano posture require accurate measurement of changes to the vertical position of body markers, since researchers are interested in changes to vertical alignment of the spine and head as pianists move through body positions that vary in
degrees between more slouched or more erect postures. The lack of precision in the y-axis coordinates reported by the Kinect for Xbox 360 further discount it as a tool for measuring posture variables quantitatively for the purpose of assessing somatic training. Similar to Obdržálek et al. (2012), we noticed that the position of the hip was often projected at unrealistically high or low positions in our Kinect data. Therefore, using the Kinect tracking position of the hip centre as the point for reference for collecting height data from Kinect data while using a point marked on the piano bench as a reference for the Dartfish data likely contributed to the discrepancies in C7 and head height data reported by the two systems.

**Accounting for Kinect error**

The limitations of Kinect skeletal tracking in the context of piano performance can be attributed to a variety of factors. One source of error is the Kinect’s unreliable pose estimation system, which has been shown to frequently identify poses incorrectly, especially in the sagittal plane, or in positions where one joint position might occlude another (Huber et al. 2014; Obdržálek et al. 2012). The problem of body occlusion is especially significant for the shoulder position when recording pianists in the sagittal plane, since the Kinect will have trouble deciding whether it is tracking a right or left arm at any given moment. Research also suggests that more reliable markerless systems for human pose estimation, such as Impulse, by PhaseSpace Inc., tend to have customizable features that allow the researcher to record the bone lengths of individual participants. In the Kinect algorithm, the lengths of skeletal segments do not remain consistent and often vary frame to frame even if a participant is not moving. The repeatability of Kinect results is higher for objects centred in the frame, and the standard deviation of results increases predictably in the periphery of the image and as the distance of the object from the Kinect increases (Alnowami et al. 2012; Dutta 2012; Pedro and Caurin 2012). Multiple studies have provided evidence that a proportional bias exists in the size of some Kinect tracking measurements, especially in the sternum region (Clark et al. 2012; Obdržálek et al. 2012; Tao et al. 2013). The Kinect also appears to track more reliably in the x and y-axes compared to the z-axis, and error tends to vary between different joint positions or angles (Pedro and Caurin 2012; Webster and Celik 2014). Changes to software and improvements to sensor technology could help to remedy some of these issues. However, researchers should also consider that the Kinect for Xbox 360 is sensitive to environmental factors, such as ambient lighting condition. In the present study, we were unable to determine which experimental factors accounted for drastic differences in tracking quality from trial to trial. It is possible that unpredictable tracking performance might persist despite changes to software, potentially making data collection with the Kinect unpredictable.

**CONCLUSION**

Our results suggest that skeletal tracking using one Kinect for Xbox 360 is not accurate enough to measure quantitative differences in pianists’ posture variables for the purposes of ergonomic assessment in somatic training interventions. Time-plots revealed frequent loss of tracking in the Kinect, and a general amplification of the true magnitude of some movements in comparison with Dartfish. The difference in pre- and post-test measurements for average head height, C7 height and forward head angle did not correspond between the two tracking
Assessing the suitability of Kinect for measuring the impact...

methods. Kinect tracking plots from x-axis coordinates were rated as ‘excellent’ or ‘good’ compared to Dartfish more than twice as often as y-axis plots. Kinect tracking quality of the head and C7 vertebral positions were rated as ‘excellent’ or ‘good’ more often than the shoulder, elbow and wrist positions. Kinect tracking was particularly poor for the right shoulder, with all sixteen tracking trials rated as ‘average’ or ‘poor’. This suggests that the suitability of the Kinect as a tracking tool depends on which part of the body the researcher is interested in studying.

Limitations of this study
We used a 1000-watt spotlight to ensure that Kinotape markers were visible in the videos for easy Dartfish tracking. This lamp may have interfered with the infrared wavelengths necessary for the Kinect camera to create an accurate depth map, and future studies requiring comparison to video-data should try to use cool LED lights instead. Furthermore, since the Kinect tracking often did not initiate automatically, we often had to relocate the unit at the beginning of a new test so it could locate the participant. Future studies comparing video recordings with coordinates from the Kinect’s infrared camera should try to record video data simultaneously from the Kinect itself using the on-board RGB camera instead of using an external video camera so that video data and the depth-tracking data are done from identical vantage points.

RECOMMENDATIONS FOR FUTURE RESEARCH
Although our trials demonstrated that the use of a single Kinect does not provide a means of accurately measuring the impact of somatic training interventions on pianists’ posture, markerless motion tracking remains a promising tool for investigating musician movement non-invasively in performance situations. Researchers should continue to explore how more sophisticated depth sensor technology equipped with a more anatomically correct skeletal model could be used in studies on musicians’ movement. Music researchers could also try using more recent versions of the Kinect sensor to see if improvements to pose estimation have been addressed in more current technology. Some researchers have greatly improved Kinect tracking accuracy by using three or more Kinects simultaneously to triangulate skeletal positions in human gait studies (Yang et al. 2016). Future research with pianists should employ a multiple-Kinect set-up to prevent occlusion problems that may arise from interference of body parts or equipment, including the piano itself.

REFERENCES


Conable, B. (Director) and Andover Productions (Producer) (2009), ‘Move well, avoid injury. What everyone needs to know about the body’, DVD, Chicago, IL: GIA Publications.


SUGGESTED CITATION


CONTRIBUTOR DETAILS

Jillian Beacon earned a bachelors in music in integrated studies from the University of Calgary, where she studied piano performance and composition. She holds a masters diploma in piano pedagogy research and a masters of arts in music from the University of Ottawa. She is currently a Ph.D. candidate in human kinetics at the University of Ottawa and conducts research on injury prevention for musicians at the Piano Pedagogy Research Laboratory. Her research explores the application of motion tracking technology to investigate piano ergonomics and the impact of somatic training methods on musicians’ movement and practice behaviour.

Contact: Piano Pedagogy Research Laboratory, University of Ottawa, Pérez Hall, 50 University Private, Room 204, Ottawa, ON, Canada K1N 6N5.
E-mail: jbeac048@uottawa.ca

Gilles Comeau received his Ph.D. in the foundations of music education in 1995 and completed post-doctoral studies in piano pedagogy from 1995 to 1997 with Marc Durand and Gilles Manny. Dr Comeau has been the beneficiary of many research grants, including a large grant from the Canadian Foundation for Innovation to set up a 1.2 million dollars research laboratory in piano pedagogy. As head of this infrastructure, he has established partnership with many other research laboratories and research institutes and set up different multidisciplinary research groups that study various aspects of piano learning and piano teaching: music reading, motivation, physiological aspect of piano performance, piano-playing health injuries, video-mediated learning. He has written various scholarly research papers and his research findings have received coverage in popular media outlets (television, radio and newspapers). He has authored many books, including Piano Pedagogy: A Research and Information Guide; Comparing Dalcroze, Orff and Kodály; the five volume...
series *Histoire illustrée de la musique pour les jeunes musiciens* and over twenty education kits to be used by music and arts teachers.

Contact: Piano Pedagogy Research Laboratory, University of Ottawa, Pérez Hall, 50 University Private, Room 204, Ottawa, ON, Canada K1N 6N5. E-mail: gcomeau@uottawa.ca

Pierre Payeur received his Ph.D. degree in electrical engineering from Université Laval, Quebec City, Canada, in 1999. Since then, he has been a Professor at the School of Electrical Engineering and Computer Science, at the University of Ottawa, Ontario, Canada. He is the founding director of the Sensing and Machine Vision for Automation and Robotic Intelligence (SMART) Research Group and a founding member of the Video and Image Processing, Computer Vision, and Autonomous Systems (VIVA) Research Laboratory. Dr. Payeur also serves as Associate Dean for the Faculty of Graduate and Postdoctoral Studies of the University of Ottawa.

Contact: School of Electrical Engineering and Computer Science, University of Ottawa, 800 King Edward, Room 5066, Ottawa, ON, Canada K1N 6N5. E-mail: ppayeur@eecs.uottawa.ca

Donald Russell (Ph.D. from Massachusetts Institute of Technology) is a professor in the Department of Mechanical and Aerospace Engineering at Carleton University (Ottawa, Ontario, Canada). He also holds an associate diploma and a choir master’s diploma from the Royal Canadian College of Organists and a diploma in church music studies from the Royal School of Church Music. He is an adjunct professor of music at the University of Ottawa and an active organist, chorister and composer. His primary research interests are the biomechanics of musical performance and performance-related injuries.

Contact: Department of Mechanical and Aerospace Engineering, Carleton University, 3091 Mackenzie, 1125 Colonel by Drive Ottawa, ON, Canada K1S 5B6. E-mail: Donald.russell@carleton.ca

Jillian F. Beacon, Gilles Comeau, Pierre Payeur and Donald Russell have asserted their right under the Copyright, Designs and Patents Act, 1988, to be identified as the authors of this work in the format that was submitted to Intellect Ltd.