Characterizing Children's Motion Qualities: Implications for the Design of Motion Applications for Children

Aishat Aloba

Department of CISE, University of Florida, aoaloba@ufl.edu

Lisa Anthony

Department of CISE, University of Florida, lanthony@cise.ufl.edu

The goal of this paper is to understand differences between children's and adults' motions in order to improve future motion recognition algorithms for children. Motion-based applications are becoming increasingly popular among children (e.g., games). These applications often rely on accurate recognition of users' motions to create meaningful interactive experiences. Motion recognition systems are usually trained on adults' motions. However, prior work has shown that children move differently from adults. Therefore, these systems will likely perform poorly on children's motions, negatively impacting their interactive experiences. Although prior work has established that there are perceivable differences between child and adult motion, these differences are yet to be quantified. If we can quantify these differences, then we can gain new insights about how children perform motions (i.e., their motion qualities). We present 24 articulation features (11 of which we newly developed) that describe motions quantitatively; we then evaluate them on a subset of child and adult motions from the publicly available Kinder-Gator dataset to reveal differences; motions in this dataset are represented as postures, each of which is defined by 3D positions of 20 joints tracked by a Kinect at a specific time instance. Our results showed that children perform motions that are quantifiably *faster*, *more intense*, *less smooth*, and *less coordinated* as compared to adults. Based on our results, we propose guidelines for improving motion recognition algorithms and designing motion applications for children.

CCS CONCEPTS • Human-centered computing~Gestural input • Human-centered computing~Graphics input devices • Social and professional topics~Children

Additional Keywords and Phrases: motion, children, motion applications, motion recognition, articulation features, global-level features, joint-level features

ACM Reference Format:

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, and Third Author's Name, Initials, and Last Name. 2018. The Title of the Paper: ACM Conference Proceedings Manuscript Submission Template: This is the subtitle of the paper, this document both explains and embodies the submission format for authors using Word. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 10 pages. NOTE: This block will be automatically generated when manuscripts are processed after acceptance.

1 Introduction

There has been a recent surge in technologies and applications that use natural interaction modalities, such as motion and speech [10,17,25]. Specifically, within the context of motions, the prevalence of low-cost tracking sensors that can accurately track users' movements, such as the Microsoft Kinect v1 and v2 [19], and most recently, the Azure Kinect

DK [20], has increased the popularity of applications that support motion interactions. For example, these sensors have enabled the development of exertion games that translate physical movements into game commands [17,32] and assistive robots that rely on motions to facilitate human-robot collaboration [18]. Motion-based applications are also becoming increasingly popular among children as researchers and practitioners have started using these applications to target children's needs. For example, exergames such as iFitQuest [17] and Vortex Mountain [32] were explicitly designed to increase the time children spend engaged in physical activity. Researchers in the field of child-robot interaction have also designed robots that can facilitate social interaction with children (e.g., companion robots [18]).

To support motion interactions, motion-based applications usually include intelligent algorithms that can accurately recognize when the user performs the specific sets of motions that the application supports. Nijhar et al. [21] found that the precision of motion recognition algorithms is positively associated with higher levels of immersion in exergames, indicating that accurate recognition of motion sets play an important role in facilitating meaningful interactive experiences. Motion recognition algorithms are usually trained on adults' motions. However, there is enough evidence to establish that children move differently from adults. Prior work has found that naïve viewers can perceive the difference between child and adult motions at levels significantly above chance, even when the motion is abstracted from all appearance cues (e.g., height and build) [15]. Prior work has also found that children exhibit higher variance in the body parts they move when performing the same motions compared to adults [4]. Hence, motion recognition systems trained on adults' motions will likely perform poorly on children's motions, negatively impacting their interactive experiences. Although prior work has established that there are perceivable differences between child and adult motion, these differences are yet to be quantified. If we can quantify these differences, then we can gain new insights about how children perform motions (i.e., their motion qualities). Hence, the goal of this paper is to understand differences between children's and adults' motions in order to improve future motion recognition algorithms for children.

To achieve this goal, we defined a set of 24 articulation features (11 of which we newly developed) that describe motions quantitatively. An articulation feature is any measure that quantifies a specific property associated with how a user performs a motion (e.g., length, shape) [5,24,27]. We used 13 features from prior work [26]that describe motions globally based on the overall posture of the body (i.e., the position of the body at a specific time instance defined by a set of joints with positions in 3D; Figure 1) (global-level features). We also defined a set of 11 new features that characterize properties of a joint moving in 3D space (joint-level). We analyzed both these global-level and joint-level features on a subset of motions from the publicly available Kinder-Gator dataset [2] to reveal differences; motions in this dataset are represented as postures, each of which is defined by 3D positions of 20 joints at a specific time instance (Figure 1).

Supporting prior work [4], we found that children are more inconsistent in how they perform motions as compared to adults. Furthermore, we found that children's natural motions differ from adults' motions along four dimensions: *speed, intensity, smoothness,* and *coordination.* Children perform motions that are quantifiably *faster, more intense, less smooth*, and *less coordinated*. We also found that the type of the motion also plays a role in the variations children show when performing motions. Based on our results, we propose guidelines for improving motion recognition algorithms and designing motion sets for children.

The contributions of this paper are: a) proposing joint-level features for quantifying motions; b) establishing features that differentiate child motion from adult motion; c) characterizing children's natural motion qualities; and d) presenting guidelines for improving motion recognition algorithms and designing motion applications for children. We hope that designers and researchers can use these guidelines to improve children's interactive experiences in motion-based applications.

2 Related Work

2.1 Understanding Children's Motion Performance

The ability to make movements relies on motor development, which is the study of the progression in a person's ability to perform motions [22]. Motor development is usually age-related [9] and, although well-developed in adults,

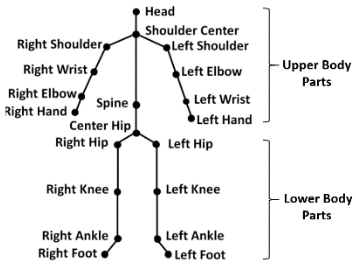

Figure 1. Kinect skeleton showing the 20 joints tracked. The skeleton also depicts a posture of a motion.

prior work in child development has outlined the ways in which children are still in the process of developing their motor skills [9,13]. For example, Hill et al. [13] noted that gross motor skills (e.g., arm movements) and postural control, which are important for the performance of complex movements [30], are still developing in children. Therefore, children will likely perform motions differently from adults. In the context of motion recognition, such differences will mean that motion recognition systems trained on adults' motions will perform poorly on children's motions, emphasizing a need to understand these differences to improve recognition performance. However, there is limited work that has sought to characterize the differences between children's and adults' motions. Aloba et al. [3] quantified the differences between children's and adults' walking and running motions using features based on gait literature and found that children move faster and with higher speed as compared to adults. In another study, Aloba et al. [4] found that children are less consistent in how they perform motions as compared to adults when comparing agreement in which joints are used to perform similar motions. However, a general understanding of how children's motions differ from adults' motions in terms of motion features that can inform motion recognition systems is still lacking. Prior work in stroke gesture research has established that features that quantitatively describe properties of the geometric paths of stroke gestures can reveal new insights about the differences between children's and adults' stroke gestures. Shaw et al. [24] analyzed children's and adults' stroke gestures using both simple features that characterize stroke gesture path properties (e.g., length) [5] and relative accuracy features that characterize deviations of a stroke gesture path from a representative articulation of the gesture (e.g., shape deviation) [27] and found significant differences only for the relative features, indicating that children are more likely than adults to exhibit variations in how they perform stroke gestures. Since stroke gestures and motions are similar in that they both involve lines, curves, and corners moving in space over time [4], our work focuses on identifying features that quantitatively describe properties of motion paths, irrespective of their motion type to reveal new insights about the differences between children's and adults' motion performance.

2.2 Improving Motion Recognition Accuracy

Prior work has noted that differences in how users perform motions can negatively impact recognition performance [4]. There has been extensive work in motion recognition research focused on defining features to promote accurate recognition of motions [7,33]. For example, Weinland et al. created Fourier features, which are view-invariant features represented in Fourier space extracted from motion history volumes [33] while Ali and Shah defined kinematic features, based on optical flow [1]. However, these features are usually designed for algorithmic interpretation and difficult for designers to reason about intuitively, so Vatavu [26] termed them *machine-readable features*. In contrast, features that are tailored for human understanding (i.e., *human-readable*) would be more informative in understanding how users actually perform motions [26]. Only Vatavu has proposed human-readable features that characterize users' motion performance [26]. The author proposed distance-based, time-based, and appearance-based (i.e., based on the composition of postures) features for characterizing motions globally, based on the overall postures that make up the motions. However, the extent to which these features can be used to understand how users actually perform motions is not yet known. Furthermore, these features focus on the overall posture, so only capture absolute characteristics of the motions as a whole. Prior work in stroke gesture research has noted that such features that capture only absolute characteristics do not have enough descriptive powers as they do not capture details about the gesture path. Hence, in our work, to identify a set of features that are human-readable and can quantitatively describe both the motions as a whole as well as properties of the motion path, we rely on Vatavu's global-level features and also define a new set of joint-level features that quantify properties of a motion path. The new features were inspired by features from stroke gesture research for quantifying stroke gesture paths [3] and movement analysis frameworks, such as Laban Movement Analysis [1]. By analyzing these features on children's and adults' motions, we aim to propose design guidelines that can be used to improve the performance of motion recognition algorithms for children's motions.

3 Articulation Features

To understand how children's motions differ from adults' motions, we defined a set of 24 articulation features (11 of which we newly developed) that describe motions quantitatively. An articulation feature is any measure that quantifies a specific property associated with how a user performs a motion (e.g., length, shape) [24,27] . Our features comprise a set of 13 **global-level features** proposed by Vatavu [26] and a newly proposed set of 11 **joint-level features**, which we developed inspired by stroke gesture research [27]. Vatavu's features describe motions globally, based on the overall posture of the body while our joint-level features describe properties of the paths of individual joints, as tracked by the motion sensor (Figure 2).

Vatavu [26] identified three categories of global features for characterizing motion performance: spatial features, kinematic features, and appearance features. **Spatial features** (7 features) capture properties related to area, volume, and amplitude of gesture movement performed by the whole body or individual body parts [26]. Examples include the volume of the 3D space in which the motion is performed (*gesture volume*) and the total amount of movement (*quantity of movement*). **Kinematic features** (2) capture properties related to time and speed [26]. These features include the time it takes to perform the motion (*performance time*) and the speed of the motion (*average gesture speed*). **Appearance features** (4) capture how motions decompose into simple units of movements [26]. Examples include the average deviation of a body posture from the centroid posture of the motion (*body posture variation)* and the maximum difference between the body postures that make up the motion (*body posture diffusion*). Table 1 shows the thirteen global-level features considered in this work.

3.1 Joint-Level Features

A limitation of global-level features is that they focus on the position of the whole body at a given point in time, so these features will not capture motion qualities that relate to subtleties of the joint articulation path, defined by a set of consecutive 3D points the joint moves through along the time-domain. These subtleties can inform an understanding of the variations in how users move their joints during motion. For example, global-level features will be useful in recognizing whether a user performed a Jump motion as opposed to a Forward Lunge motion. But they may not be as helpful in understanding whether two users lifted their feet differently when performing the Jump motion. Therefore, we propose a set of joint-level features that quantify geometric properties (e.g., shape, curvature) of the joint articulation paths necessary to perform motions. To identify the joints necessary to perform a given motion, we use the filterJoint method proposed by Aloba et al. [4], which uses standard deviation and K-means clustering [12] iteratively to select the set of joints that are actively moving during a motion.

To compute the joint-level features, we use the **template gesture task axis** method from 2D touchscreen stroke gesture research [27]. Vatavu et al. defined a gesture task axis as a representative way to articulate a stroke gesture [27]. To compute a 2D stroke gesture task axis, stroke gesture paths are first resampled to the same number of points to enable point-to-point comparison and translated so that the centroid is at the origin [27]. The template task axis is a "*canonical template form supplied to a recognizer to which articulated gestures will be compared in a template-based matching approach*" [27]. Given a dataset of stroke gestures (e.g., for the letter A), the template task axis is the gesture with the least distance to a representative gesture, computed as the average of all the gestures in the dataset [27].

Similarly, we define a **joint task axis** as a representative way to move a joint. Unlike stroke gestures in which there is only one articulation path, defined by the finger's movement, and one task axis, motions have multiple articulation paths, defined by all the joints tracked by the motion sensor. Therefore, a motion will have multiple joint

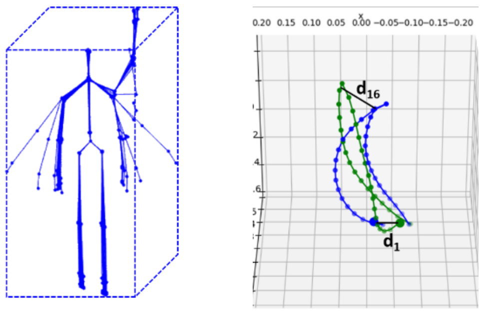

Figure 2. (left) Gesture volume computed on postures of a Raise your hand motion. (right) Shape error (sum of d's) computed by comparing two articulation paths of the wrist joint of a Raise your hand motion (blue path = joint task axis)

task axes, one for each joint necessary to perform the motion. To apply the template method, we use a leave-one-out approach inspired by the Leave-One-Out-Cross-Validation approach (LOOCV) used in recognition experiments [4,35]. In LOOCV, motions from one participant are selected for testing (i.e., the candidate user) while motions from all other participants are used for training. This process is repeated until all participants have been selected once for testing. Similarly, given a set of motions (e.g., for the motion Raise your hand), one user is selected as the candidate user. For each of their joints that is actively moving, the template joint task axis for that joint will include the corresponding joint articulation path of every other user in the set. For example, given three users (c1, c2, c3) wherein c1 is the candidate user who actively moves their right hand when raising their hand, then the task axis for the right hand joint will include the articulation paths of the right hand joint for users c2 and c3. The process is repeated until every user in the set has been selected as the candidate user once.

Next, we define a set of joint-level features that characterize the deviations of the articulation path of a joint from the joint task axis with respect to properties of the joint articulation path, such as length, shape, curvature, and time (Table 1). Because our end goal is to improve recognition of children's motions and recognition accuracy for pattern matching algorithms is fundamentally driven by consistency among users, we focus on features that can quantify such consistency in motion performance. These features were inspired by the relative accuracy features from Vatavu et al. [27], features for tracking mouse paths [14], and features from Laban Movement Analysis [16]. We defined 11 joint-level features, such as *Shape error*, which measures the average absolute deviation of the shape of a joint articulation path and the corresponding joint task axis, and *Shape variability*, which measures the uniformity of the shape errors. Most of these features rely on a concept of "error", which does not imply that the user moved the joint in the wrong way, but rather measures inconsistency with respect to the task axis [27]. Each feature requires a comparison between the joint path of the candidate user and every other users' corresponding joint path (i.e., the articulation paths in the joint task axis). Then, the average of the feature computation across all the comparisons and all the joints that the candidate user is actively moving is used to compute the feature for a given candidate user's motion instance. For example, given three users (c1, c2, c3) where c1 is the candidate user who actively raises their right hand and right elbow when raising their hand. To compute the shape error, we compare the shape of c1's right hand path to the shape of c2's right hand path and the shape of c1's right hand path to the shape of c3's right hand path and get the average (u1). We do the same for the right elbow joint to get the average (u2). Then, we take the average of u1 and u2 as the Shape error. Our set of eleven joint-level features is shown in Table 1.

3.2 Feature Computation

We computed the features on a set of motions we selected from the Kinder-Gator dataset [2], a publicly-available dataset of 10 children (ages 5 to 9) and 10 adults performing motions, forward-facing the Kinect v1. The Kinect v1 tracks the motions of 20 joints along three dimensions: x (horizontal), y (vertical), and z (depth). Although Kinder-Gator motions were collected in a lab setting, the experiment was set up similarly to typical Kinect-based games in that children stood forward-facing the Kinect and performed the motions as they would naturally in their day-to-day activities [2]. Specifically, we computed the features on the same subset of 14 motions used in prior work by Aloba et al. [4], selected to avoid overly similar motions (Table 2). Because we are comparing child and adult motion, we account for height differences: we use participant height as a normalization factor for the global-level features, estimated as the absolute difference between a user's head and foot in the y-dimension of the first frame when performing the *Raise hand* motion. This movement guarantees that the participant is standing in the first frame. For the joint-level features, first, we smoothed the joint articulation path using an exponential moving average filter to remove tracking noise, similar to prior work [4]. We then resampled the path to n = 32 points, so that different motion paths are directly comparable even at different movement speeds while maintaining the original duration of the motion. Then, we translated the centroid of each joint path to the origin, such that the distance between the positions denotes differences in the shape of the paths. Finally, we scaled the paths uniformly by the same scaling factor across all dimensions to normalize the paths with respect to the height/limb length of the participant, thus accounting for height differences. Gesture recognition research has used this scaling method uses this scaling method to ensure that different articulation paths are comparable irrespective of size [4,28].

To ensure accurate comparison of a joint articulation path and the joint task axis, we apply the following methods to address the four ways in which motions can be performed that will impact the distance between two joint articulation paths: a) Same joint (i.e., left vs right) and same direction, compare as-is. (b) Same joint but different directions (e.g., swiping right to left vs. swiping left to right), FLIP the joint articulation path 180^0 along the x-axis to change direction and then compare joint articulation path and joint task axis. (c) Different joints but same direction, REPLACE articulation path of the joint task axis with the articulation path of the joint of its opposite limb (i.e., replace left joints with right joints and vice-versa, leave middle joints as-is) then compare. (d) Different limbs and different directions, FLIP the joint articulation path, REPLACE articulation path of joint task axis, then compare. For a given candidate joint articulation path and an articulation path in the set of articulation paths in the corresponding joint task axis, the comparison with the least Euclidean distance is used to compute all the joint-level features for this pair of joint paths.

3.3 Analysis

To quantify the differences between how children and adults perform motions, we analyzed the computed feature values statistically using ANOVA. Because none of our data satisfied the requirements for normality, we used a non-parametric version of the ANOVA test, known as the Aligned Rank Transform (ART) [34]. For each of the features, we ran a two-way repeated measures ANOVA with a between-subjects factor of *age group* (child, adult) and a within-subjects factor of *motion type* (14 motions: Table 2). Since we want to understand how child motion differs from adult motion, we are only interested in the main effect of age group and the interaction effect between age group and motion type (different motion types will be guaranteed to exhibit different feature values). All post-hoc analysis was done using the Tukey method.

3.4 Results

For the global-level features, we found a significant difference in 6 of the 13 features, indicating that children's motions in the Kinder-Gator dataset differed from adults' motions with respect to the specific property being measured by each of these six features (starred in Table 1). For the joint-level features, we found significant differences in all but 2 of the 11 features (starred in Table 1). Since the joint-level features compare articulation paths, this finding indicates that children move more inconsistently than adults for each of the specific properties being measured. In the following results, we only present the features (GL for global-level features and JL for joint-level features) that showed a significant difference between children's and adults' motions. We include computation formulae for our new JL features below; GL features from prior work can be found in Vatavu et al. [2] and formulae for all features are provided in our paper's supplemental materials.

Table 1: List of Articulation Features. * means the feature was significant for age group at p < 0.05.

Global-Level Features (13)		Joint-Level Features (11)	
Gesture Volume *	Performance Time*	Shape Error*	Efficiency
Gesture Area*	Average Gesture Speed*	Shape Variability*	Time Error
Quantity of Movement	Body Pose Variation	Bend Error*	Speed Error*
Quantity of Upper Movement	Body Pose Diffusion	Bend Variability*	Acceleration Error*
Quantity of Lower Movement	Body Pose Density*	Length Error*	Jerk Error*
Difference of Movement	Body Pose Rate*	Size Error*	
Ratio of Movement*			

3.4.1 Gesture Volume (GV) and Gesture Area (GA) [GL].

GV measures the volume of the 3D space where the motion is performed and is computed as the product of the difference between the maximum and minimum positions of the body in the x, y, and z dimensions. GA measures the area of the 2D space in front of the motion sensor and is computed the same way as GV without the z dimension [26]. We found a significant main effect of age group for both features (GV: $F_{1,18} = 6.44$, $p < 0.05$, GA: $F_{1,18} = 20.18$, $p < 0.001$) with children requiring a larger 3D and 2D space (GV: mean (M) = 0.20m^3, standard deviation (SD) = 0.11, median = 0.18, GA: M = 0.55m^2, SD = 0.17, median = 0.52) as compared to adults (GV: M = 0.16m^3, SD = 0.12, median = 0.12, GA: M = 0.44m^2, SD = 0.13, median = 0.41), when normalized for height. We also found a significant interaction effect between age group and motion type for both features (GV: $F_{13,234} = 3.28$, $p < 0.001$, GA: $F_{13,234} = 4.68$, $p < 0.001$). Post-hoc tests for GV revealed that children and adults require similar 3D space for all motion types except Jump, Lift leg to side, and Touch toes, wherein children required a larger 3D space. For GA, children required a larger 2D space for the motions Jump, Swipe screen, and Raise hand. Therefore, in general, children use more space (proportionally) to perform motions as compared to adults.

3.4.2 Ratio of Movement (RM) [GL].

RM measures the ratio of the quantity of movement (QM) in the upper and lower body where QM measures the amount of movement, computed as the cumulative pairwise Euclidean distance between corresponding joints of time-consecutive frames [26]. We found a significant main effect of age group ($F_{1,18} = 35.82$, $p < 0.001$), with children having a lower ratio of upper to lower body movement (M = 2.91, SD = 2.04, median = 2.41) as compared to adults (M = 3.77, SD = 3.45, median = 2.29). Although we found a significant interaction effect ($F_{13,234} = 4.68$, $p < 0.001$), post-hoc tests revealed no significant difference between interaction pairs after Tukey correction. These findings indicate that for the motion types we considered, children move their upper body more in comparison to their lower body.

3.4.3 Performance Time (PT) and Average Gesture Speed (AGS) [GL].

PT measures the time a user takes to perform a motion while AGS is the ratio of the quantity of movement to performance time [26]. We found a significant effect of age group for both features (PT: $F_{1,18} = 8.08$, $p < 0.01$, AGS: $F_{1,18} = 29.37$, $p < 0.001$) with children moving faster (M = 3.20s, SD = 1.17, median = 2.90) and at a higher speed (M = 0.16m/s, SD = 0.09, median = 0.09) as compared to adults (PT: M = 3.76s, SD = 1.19, median = 0.14, AGS: M = 0.12m/s, SD = 0.07, median = 0.09). We also found a significant interaction effect between age group and motion type for both features (PT: $F_{13,234} = 2.26$, $p < 0.01$, AGS: $F_{13,234} = 2.83$, $p < 0.01$). Post-hoc tests revealed that children moved at similar speeds as adults for all motion types except Bend knee, Jump, Raise hand, and Swipe screen (Figure 3). Like ratio of movement, post-hoc tests for performance time showed no significant difference between interaction pairs after Tukey correction. These findings indicate that children generally move faster than adults, corroborating findings from prior work [3].

3.4.4 Body Posture Density (BPDe) and Body Posture Rate (BPR) [GL]

BPDe is the ratio of the Body pose variation (BPV) to the gesture volume (GV). BPV measures the variability of body postures and is computed as the average Euclidean distance between the position of all postures of the motion and the average posture [26]. BPR is the ratio of BPV to the performance time [26]. We found a significant effect of age group

Table 2. Subset of motions used, selected from the Kinder-Gator dataset from Aloba et al. [4]. Terms in parentheses indicate the abbreviated motion name.

Gestures	
Touch your toes (Touch toes)	Do a forward lunge (Forward lunge)
Point at the camera (Point at camera)	Lift your leg to one side (Lift leg to side)
Raise your hand (Raise hand)	Jump (Jump)
Raise your arm to one side (Raise arm)	Kick a ball as hard as you can (Kick ball hard)
Bend your knee (Bend Knee)	Throw a ball as far as you can (Throw ball far)
Put your hands on your hip and lean to one side (Put hands on hips)	Swipe across an imaginary screen in front of you (Swipe screen)
Punch (Punch)	
Bow (Bow)	

for both features (BPDe: $F_{1,18} = 10.74$, $p < 0.01$, BPR: $F_{1,18} = 18.99$, $p < 0.01$), with children having lower posture densities (M = 6.70m^{-2}, SD = 3.19, median = 5.89) but higher posture rates (M = 0.42m/s, SD =0.26, median = 0.37) as compared to adults (BPDe: M = 8.05m^{-2}, SD = 4.28, median = 6.87, BPR: M = 0.32m/s, SD =0.21, median = 0.24). We also found a significant interaction effect between age group and motion type for both features (BPDe: $F_{13,234} = 3.16$, $p < 0.001$, BPR: $F_{13,234} = 1.91$, $p < 0.05$). Children and adults had similar posture densities for all motion types except Lift leg to side and had similar posture rates for all motion types except Bend knee, Jump, and Put hands on hips. These findings suggest that children's and adults' motions will differ in appearance (i.e., the distribution of their body postures) when the space and time they require to perform the motion are considered.

3.4.5 Shape Error (ShE) and Shape Variability (ShV) [JL].

ShE measures the average absolute deviation of the shape of a joint path in a motion instance from the shape of the same joint path in the joint task axis and is computed as the Euclidean distance between the 3D points of both joint paths. ShV measures how uniform the shape errors are along a joint path and is computed as the standard deviation of the distances between the joint paths. The formulae for computing the features are:

$$ShE = \frac{1}{n}\sum_{i=1}^{n}\left\|j_{c(i)} - j_{t(i)}\right\|, ShV = \frac{1}{n-1}\sum_{i=1}^{n}\left\|j_{c(i)} - j_{t(i)}\right\| - ShE$$

We found a significant effect of age group for both features (ShE: $F_{1,18} = 137.20$, $p < 0.0001$, ShV: $F_{1,18} = 29.37$, $p < 0.001$) with children having higher shape errors (M = 0.31, SD = 0.08, median = 0.31) and shape variabilities (M = 0.17, SD = 0.06, median = 0.17) as compared to adults (ShE: M = 0.20, SD = 0.08, median = 0.21, ShV: M = 0.11, SD = 0.06, median = 0.10). We also found a significant interaction effect between age group and motion type for both features (ShE: $F_{13,234} = 4.70$, $p < 0.0001$, ShV: $F_{13,234} = 4.41$, $p < 0.0001$). Children had higher shape errors than adults for the motions Bend Knee, Bow, Forward lunge, Jump, Lift leg to side, Point at camera, Put hands on hips, Raise arm to side, Swipe screen, and Throw ball far (Figure 3). They had higher shape variability than adults for the motions Forward lunge, Kick ball hard, Point at camera, Swipe screen and Throw ball far. Therefore, children show more variations in how they move their body parts as compared to adults for most motion types, thus making them more inconsistent.

3.4.6 Bend Error (BE) and Bend Variability (BV) [JL].

BE measures users' tendency to bend or curve the joint path and is computed as the absolute difference between the turning angles of the 3D points of a joint path in a motion instance and a joint path in the joint task axis. Given a 3D point p, its turning angle is the angle between p and the previous point (p-1) and p and the next point (p+1) [27]. BV measures how uniform the bend errors are along the joint path and is computed as the standard deviation of the differences between the turning angles. The formulae for computing the features are:

$$BE = \frac{1}{n-1}\sum_{i=1}^{n}\left\|\theta_{c(i)} - \theta_{t(i)}\right\|, BV = \frac{1}{n-1}\sum_{i=1}^{n}\left\|\theta_{c(i)} - \theta_{t(i)}\right\| - BE, \theta_i = < \overline{J_{\iota-1}J_\iota}.\overline{J_\iota J_{\iota+1}}$$

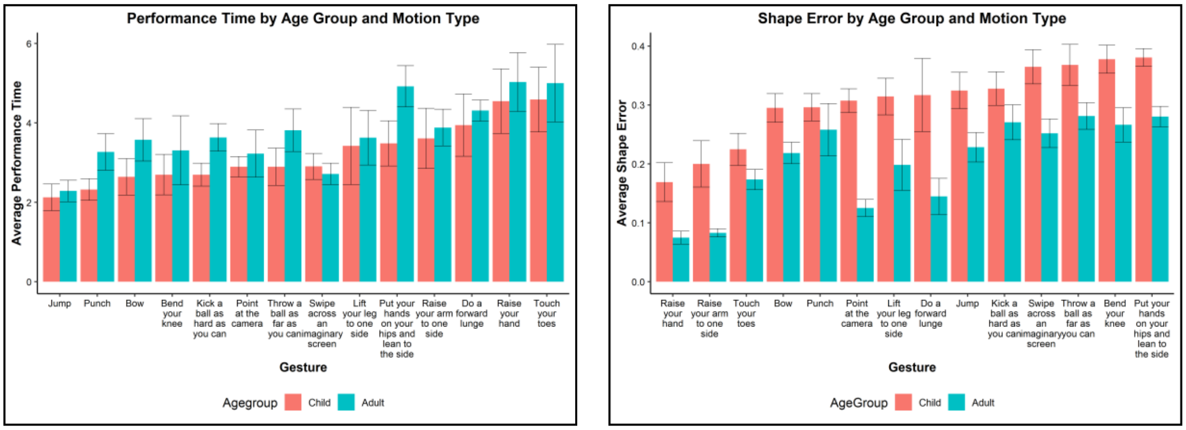

Figure 3. (left) Performance Time (GL) by age group and motion type. (right) Shape Error (JL) by age group and motion type.

We found a significant effect of age group for both features (BE: $F_{1,18} = 78.61$, $p < 0.0001$, BV: $F_{1,18} = 36.50$, $p < 0.0001$) with children having higher bend errors (M = 0.31, SD = 0.09, median = 0.30) and bend variabilities (M = 0.41, SD = 0.09, median = 0.39) as compared to adults (BE: M = 0.24, SD = 0.09, median = 0.23, BV: M = 0.35, SD = 0.10, median = 0.26). We also found a significant interaction effect between age group and motion type for both features (BE: $F_{13,234} = 7.35$, $p < 0.0001$, BV: $F_{13,234} = 6.81$, $p < 0.0001$). Children had higher bend errors than adults for the motions Lift leg to side, Raise hand, and Swipe screen, and had higher bend variability than adults for motions Raise hand and Swipe screen. Hence, children vary in how they bend their body parts to perform motions across instances of the same motion.

3.4.7 Length Error (LE) and Size Error (SE) [JL].

LE measures a user's tendency to stretch their joint articulation path and is computed as the absolute difference between the path lengths of a joint path in a motion instance and a joint path in the joint task axis. SE measures a user's tendency to stretch the joint path with respect to the gesture volume (GV) and is computed as the absolute difference between the volumes of the smallest bounding box encompassing each joint path. The formulae for computing the features are:

$$LE = |L(j_c) - L(j_t)|, L(j) = \sum_{i=1}^{n} \|j_i - j_{i-1}\|, SE = |V(j_c) - V(j_t)|, V(j) = \prod_{d \,\epsilon\, \{x,y,z\}} max(p_d) - min(p_d)$$

We found a significant effect of age group for both features (LE: $F_{1,18} = 35.28$, $p < 0.0001$, SE: $F_{1,18} = 37.44$, $p < 0.0001$) with children having higher length errors (M = 0.44, SD = 0.27, median = 0.39) and size errors (M = 0.13, SD = 0.06, median = 0.13) as compared to adults (LE: M = 0.32, SD = 0.17, median = 0.29, SE: M = 0.10, SD = 0.07, median = 0.09). We also found a significant interaction effect between age group and motion type for both features (LE: $F_{13,234} = 6.42$, $p < 0.0001$, SE: $F_{13,234} = 3.21$, $p < 0.0001$). Children had higher length errors than adults for the motions Lift leg to side, Point at camera, Raise hand, and Forward lunge, and had higher size errors than adults for the Forward lunge motion. The length error findings corroborate the shape error findings in that children are inconsistent in how they move their body parts. The size error results indicate that not only do children require more space to perform motions (see section 3.4.1), but they also vary in the amount of space they require to perform motions across instances of the same motion.

3.4.8 Speed Error (VE), Acceleration Error (AE), and Jerk Error (JE) [JL].

VE, AE, and JE measure the average difference between the speed, acceleration, and jerk of a joint path in a motion instance and a joint path in the joint task axis, respectively. Speed, acceleration, and jerk are computed as the ratio of quantity of movement, speed, and acceleration to performance time, respectively. The formulae for computing the features are:

$$VE = \frac{1}{n}\sum_{i=1}^{n}\left|S\left(j_{c(i)}\right) - S\left(j_{t(i)}\right)\right| \quad AE = \frac{1}{n}\sum_{i=1}^{n}\left|A\left(j_{c(i)}\right) - A\left(j_{t(i)}\right)\right| \quad JE = \frac{1}{n}\sum_{i=1}^{n}\left|H\left(j_{c(i)}\right) - H\left(j_{t(i)}\right)\right|$$

$$S = \begin{cases} \dfrac{X(j_{i+1}) - X(j_i)}{t_{i+1} - t_i} & i = 1 \\ \dfrac{X(j_i) - X(j_{i-1})}{t_i - t_{i-1}} & i = n \\ \dfrac{X(j_{i+1}) - X(j_{i-1})}{t_{i+1} - t_{i-1}} & Other \end{cases} \quad A = \begin{cases} \dfrac{S(j_{i+1}) - S(j_i)}{t_{i+1} - t_i} & i = 1 \\ \dfrac{S(j_i) - S(j_{i-1})}{t_i - t_{i-1}} & i = n \\ \dfrac{S(j_{i+1}) - S(j_{i-1})}{t_{i+1} - t_{i-1}} & Other \end{cases} \quad H = \begin{cases} \dfrac{A(j_{i+1}) - A(j_i)}{t_{i+1} - t_i} & i = 1 \\ \dfrac{A(j_i) - A(j_{i-1})}{t_i - t_{i-1}} & i = n \\ \dfrac{A(j_{i+1}) - A(j_{i-1})}{t_{i+1} - t_{i-1}} & Other \end{cases}$$

$$X(j_i) = arc - length = \sum_{a=2}^{i} \|j_a - j_{a-1}\|$$

We found a significant effect of age group for all features (VE: $F_{1,18} = 23.05$, $p < 0.001$, AE: $F_{1,18} = 80.67$, $p < 0.0001$, JE: $F_{1,18} = 70.07$, $p < 0.0001$) with children having higher speed errors (M = 0.30, SD = 0.12, median = 0.27), acceleration errors (M = 0.28, SD = 0.13, median = 0.25), and jerk errors (M = 2.85, SD = 1.72, median = 2.41) as compared to adults (VE: M = 0.27, SD = 0.10, median = 0.21, AE: M = 0.19, SD = 0.11, median = 0.17, JE: M = 1.69, SD = 1.72, median = 2.41). We also found a significant interaction effect between age group and motion type for all features (VE: $F_{13,234} = 3.51$, $p < 0.0001$, AE: $F_{13,234} = 5.01$, $p < 0.0001$, JE: $F_{13,234} = 5.92$, $p < 0.0001$). Children had higher speed errors than adults for the motions Lift leg to side and Swipe screen; they had higher acceleration errors for the motions Punch, Raise arm to side, and Swipe screen; and they had higher jerk errors for the motions Bow, Lift leg to side, Punch, Put hands on hips, Raise arm to side, Swipe screen, and Throw ball far. Our speed error results indicate that even though children move consistently faster than adults (i.e., because there was no significant effect of age group for Time error), they vary in the rate at which they move their body parts over the duration of the motion. Prior work has also noted that high levels of jerk indicate that motions are performed quickly, with more urgency and less smoothness [31]. Since the global-level features showed us that children move faster and with higher speeds, taken with these results, we can also deduce that children move with significantly higher acceleration and jerk as compared to adults. Hence, children's motions will be less smooth in general as compared to adults' motions.

4 Children's Motion Qualities

Overall, our results showed that children move differently from adults in ways that can be quantified with specific posture- and joint-motion-based articulation features. For example, children move faster and are more inconsistent in how they move their body parts to perform motions, as compared to adults (Figure 3). Next, we looked across the features to identify themes, in which a theme is an inference from the result (e.g., children require more space compared to adults). We grouped these themes to identify dimensions along which children's motions differ from adults' motions:

4.1 Speed

This dimension relates to how fast children perform motions. **Children move faster than adults**. Children are consistently faster than adults when performing motions (based on features like performance time and time error). This finding echoes and deepens prior work, which showed that children move faster than adults for walking and running motions [3]. Our results also show that children move with higher speeds but exhibit more variation in the speeds they use to perform motions (average gesture speed, speed error). Since speed measures the rate of joint movement and time is consistent, these variations mean that children are inconsistent in how they move their body parts to perform motions. Prior work in stroke gesture research has shown that gestures articulated at faster speeds have higher inconsistencies compared to gestures articulated at slower speeds [27]. Hence, there is a speed-accuracy trade-off relating to children's motion performance; they move quickly but are more inconsistent (i.e., less accurate) in how they perform motions. This trade-off emerges because children are still developing their motor abilities [9].

4.2 Intensity

This dimension relates to the amount of effort used to perform motions. **Children perform more exaggerated motions, therefore more intense, motions as compared to adults**. Children and adults differ in the appearance of their motions when space and time are considered. Children's postures are less dense and are completed faster than adults (Body pose density, Body pose rate) since they require more space and move faster (Gesture volume, Performance time). Hence, children's motions will appear more exaggerated as compared to adults, thus requiring more effort [6,8]. Prior work in biomechanics supports this finding as it noted that exaggerated postures require more energy [23]. In addition, prior work in exercise motions also noted that exaggerated motions require less time [8]; an assertion evidenced by our Performance Time and Body pose rate findings. Like speed, intensity also plays a role in the inconsistencies children show when performing motions. Since children are still developing their motor skills [9] and move fast, the additional effort they use to perform motions could result in loss of control over body parts during movements, resulting in loss of balance (a behavior that leads to jerky motions).

4.3 Smoothness

This dimension relates to how well children move each body part that is necessary to performing a motion. **Children's motions are less smooth as compared to adults' motions**. Children jerk inconsistently when performing motions (jerk error) but move with higher levels of jerk in their motions (average gesture speed, performance time). Therefore, children are more likely than adults to make jerky motions, which explains why children are inconsistent in how they move body parts (shape error). Furthermore, the inconsistency in the uniformity of their shape errors means that children are also inconsistent in the ways in which they are inconsistent (shape variability). Prior work has noted that children in the age range we considered (i.e., 5 to 9) are still developing their motor abilities [9]. Therefore, we believe that they perform motions less smoothly as compared to adults because they have less expertise controlling body parts to perform motions. As mentioned in the previous section, speed and intensity also play a role in the smoothness of motions. Children are more likely to lose control over body parts when the motion is performed with high speed and forceful intensity, especially for complex motions involving the whole-body (Forward lunge and Jump). Our Shape Error finding validates this idea as it showed that children are not as consistent as adults when performing both motions.

4.4 Coordination

This dimension relates to how well children move body parts relative to each other. This quality is closely related to smoothness, but coordination involves multiple body parts. **Children make less well-coordinated multi-limb movements as compared to adults**. Motions often require a lateral shift of an individual's center of gravity and body balance due to postural changes [23]. Motions involving movements of lower body parts (e.g., Kick ball hard, Lift leg to side, and Bend Knee) shift a user's center of gravity once they lift their foot. We know from our other themes that children move with high speeds and forceful intensity, behaviors that result in a shift in body balance [23]. However, children are still developing their postural stability [13], indicating that they will often overestimate the speed and intensity at which a motion should be performed and falter due to loss of balance. Prior work has noted that children move their arms when performing lower body movements to maintain balance [13], which may explain why children move their upper body more than their lower body when performing movements, as compared to adults (Ratio of Movement). Therefore, we see that children move extra body parts that are not necessary to motion performance, in order to stabilize their body once they begin to falter. Prior work further noted that how users account for a shift in body balance will determine their degree of coordination [23], with well-coordinated movements requiring control over body parts. Since children move these extra body parts in response to a loss in balance, they will not move these extra body parts as intentionally (i.e., with full awareness) as the key body parts they are actively moving. Therefore, children will have little to no conscious control over how these body parts move (e.g., direction of movement). Combined with the smoothness dimension, the above statements indicate that children not only find it difficult controlling one body part in isolation but will also find it difficult coordinating multiple body parts.

Prior work further noted that the movement of extra body parts results in inconsistencies in how children perform motions [4]. The authors computed the degree of agreement among the actively moving joints when performing lower body motions [4], in which the degree of agreement is defined as the total number of unique joint combinations used to perform a motion. They found that children had a lower degree of agreement, indicating higher inconsistency, for lower-limb motions due to some children moving upper body parts (e.g., arms) even though such body parts were not necessary for motion performance.

5 Design Implications

Our results contribute the first in-depth understanding of the quantifiable differences between child and adult motions. Based on our results, we propose guidelines both for designing motion-based applications and for developing motion recognition systems for children.

Favor simpler motions. Children are less consistent when performing complex movements that involve the whole-body. Therefore, designers of motion applications should favor simple motions that children are familiar with (e.g., upper-body motions requiring a single limb) instead of more complex whole-body motions (e.g., lunges). However, favoring simple motions may not always be feasible (e.g., exercise games require some complex exercise motions to make ensure achieve moderate-to-vigorous physical activity [17,32]). Designers should provide opportunities for children to practice complex motions to get them more familiar with how the motion is performed.

Be flexible about space requirements. Children require more space to perform motions and are inconsistent in the space they require to move body parts. Therefore, designers of whole-body applications should customize space requirements to ensure that children have enough space (with respect to area and volume). Furthermore, designers should program the application to verify proper space allowances (e.g., using the depth camera to detect hazards) to prepare for scenarios in which children make exaggerated motions that will require even more space. Space verification is important to prevent injury [11].

Favor looser pointing approaches. Prior work has adapted template-based stroke gesture recognizers that use one-to-one matching to motions [4]. However, children make jerky motions and are inconsistent in how they move their body parts to perform motions (shape error). Therefore, we recommend designers of motion recognition systems for children to use less stringent point-based approaches. For example, designers can use something similar to the many-to-one approach used in the $P+ stroke gesture recognizer [29]. This approach accounts for variations along an articulation path by choosing the best point in a template path that matches a given point in the candidate path, such that multiple points in the template path can be assigned to one point in the candidate path during recognition. This approach will be less affected by shape errors and shape variabilities, thus improving recognition.

Account for extra limb movements. We saw children often moved extra limbs to perform motions, which resulted in inconsistencies in their motion performance. Therefore, designers of motion recognition systems should consider only the joints that users are actively moving intentionally. Prior work in motion recognition with the filterJoint method found an increase in recognition accuracy when only the minimum sufficient subset of the joints tracked by the motion sensor is considered [4]. For example, designers can use the filterJoint method to select all actively moving joints and use only these joints during recognition to improve accuracy.

6 Conclusion and Future Work

Although prior work had established that there are perceivable differences between children's and adults' motions, these differences had yet to be quantified. In this work, we presented a set of 24 global-level and joint-level features and evaluated them on a subset of motions from the publicly-available Kinder-Gator dataset [2] to illuminate differences. We found that children's natural motions are *less smooth* and *less coordinated*, and are performed at *faster speeds* with *higher intensities*, than adults' motions. Future work can also use our results to evaluate whether generated children's gestures (e.g., in Embodied Conversational Agents) resemble actual children's gestures for use in relevant applications. One limitation of our work is that motions in the Kinder-Gator dataset were tracked with the Kinect v1

[19], which is less accurate in comparison to high-precision mocap sensors. We used this dataset due to the lack of publicly available datasets with more precise tracking of children's motions, and collecting new data was out of scope for our project. While our work did include approaches to mitigate tracking noise, e.g., the filterJoint method [4], future work can replicate our study on datasets tracked with higher-precision sensors (when available) to further validate our findings. Another limitation is that we only considered between-user consistency because Kinder-Gator only includes one example per user per motion. Future work can consider using our articulation features to analyze within-user consistency in one of two ways: (1) periodic motions, in which the same postures occur multiple times over a brief time period, such as multiple jumping jacks; or (2) multiple separate repetitions of the same motion type, at different time points. While Kinder-Gator does include some periodic motions which we did not analyze here, to the best of our knowledge, there is currently no dataset of children's motions that include multiple separate repetitions. Based on our results, we proposed guidelines for designing motion recognition systems and motion applications for children. We hope that designers and researchers can adopt these guidelines to tailor motion recognition systems to children's motion qualities for more accurate recognition and improve children's interactive experiences in motion-based applications.

ACKNOWLEDGMENTS

This work is partially supported by National Science Foundation Grant Award #IIS-1552598. Opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect these agencies' views. Special thanks to Dr. Jaime Ruiz, Dr. Eakta Jain, and Dr. Angelos Barmpoutis for their feedback on the research methodology.

REFERENCES

<bib id="bib1"><number>[1]</number>Saad Ali and Mubarak Shah. 2010. Human action recognition in videos using kinematic features and multiple instance learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 2: 288–303. https://doi.org/10.1109/TPAMI.2008.284
<bib id="bib2"><number>[2]</number>Aishat Aloba, Gianne Flores, Julia Woodward, Alex Shaw, Amanda Castonguay, Isabella Cuba, Yuzhu Dong, Eakta Jain, and Lisa Anthony. 2018. Kinder-Gator: The UF kinect database of child and adult motion. In *Eurographics (Short Papers)*, 13–16. https://doi.org/10.2312/egs.20181033</bib>
<bib id="bib3"><number>[3]</number>Aishat Aloba, Annie Luc, Julia Woodward, Yuzhu Dong, Rong Zhang, Eakta Jain, and Lisa Anthony. 2019. Quantifying Differences between Child and Adult Motion based on Gait Features. In *International Conference on Human-Computer Interaction (HCII '19)*, 385–402.</bib>
<bib id="bib4"><number>[4]</number>Aishat Aloba, Julia Woodward, and Lisa Anthony. 2020. FilterJoint: Toward an Understanding of Whole-Body Gesture Articulation. In *International Conference on Multimodal Interaction (ICMI '20)*, 213–221. https://doi.org/10.1145/3382507.3418822</bib>
<bib id="bib5"><number>[5]</number>Lisa Anthony, Radu-Daniel Vatavu, and Jacob O. Wobbrock. 2013. Understanding the consistency of users' pen and finger stroke gesture articulation. In *Proceedings of the Graphics Interface Conference (GI '13)*, 87–94. https://doi.org/10.5555/2532129.2532145</bib>
<bib id="bib6"><number>[6]</number>Leslie Bishko. 2014. Animation principles and laban movement analysis - movement frameworks for creatinc empathic character performances. *Nonverbal Communication in Virtual Worlds: Understanding and Designing Expressive Characters*: 177–203.</bib>
<bib id="bib7"><number>[7]</number>Aaron F. Bobick and James W. Davis. 2001. The recognition of human movement using temporal templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*: 257–267. https://doi.org/http://dx.doi.org/10.1109/34.910878</bib>
<bib id="bib8"><number>[8]</number>William Bruce. 1910. Exercise in Education and Medicine. *The Lancet* 175, 4521: 1165–1166. https://doi.org/10.1016/S0140-6736(01)14222-4</bib>
<bib id="bib9"><number>[9]</number>Frances Cleland-Donnelly, Suzanne S Mueller, and David L Gallahue. 2017. *Developmental Physical Education for All Children: Theory Into Practice*. Human Kinetics.</bib>
<bib id="bib10"><number>[10]</number>Diane J. Cook and Sajal K. Das. 2007. How smart are our environments? An updated look at the state of the art. *Pervasive and Mobile Computing* 3, 2: 53–73. https://doi.org/10.1016/j.pmcj.2006.12.001</bib>
<bib id="bib11"><number>[11]</number>Shannon E. Gray and Caroline F. Finch. 2015. The causes of injuries sustained at fitness facilities presenting to Victorian emergency departments - identifying the main culprits. *Injury Epidemiology* 2, 1: 1–8. https://doi.org/10.1186/s40621-015-0037-4</bib>
<bib id="bib12"><number>[12]</number>John A. Hartigan and Manchek A. Wong. 2006. Algorithm AS 136: A K-Means clustering algorithm. *Journal of the Royal Statistical Society Series C (Applied Statistics)* 29, 1: 100–108. https://doi.org/10.2307/2346830</bib>
<bib id="bib13"><number>[13]</number>Mathew W. Hill, Maximilian M. Wdowski, Adam Pennell, David F. Stodden, and Michael J. Duncan. 2019. Dynamic postural control in children: Do the arms lend the legs a helping hand? *Frontiers in Physiology* 9: 1932–1932. https://doi.org/10.3389/fphys.2018.01932</bib>
<bib id="bib14"><number>[14]</number>Zaher Hinbarji, Rami Albatal, and Cathal Gurrin. 2015. Dynamic user authentication based on mouse movements curves. In *International Conference on Multimedia Modeling (MMM '15)*, 111–122. https://doi.org/10.1007/978-3-319-14442-9_10</bib>
<bib id="bib15"><number>[15]</number>Eakta Jain, Lisa Anthony, Aishat Aloba, Amanda Castonguay, Isabella Cuba, Alex Shaw, and Julia Woodward. 2016. Is the Motion of a Child Perceivably Different from the Motion of an Adult? *ACM Transactions on Applied Perception* 13, 4: 1–17. https://doi.org/10.1145/2947616</bib>
<bib id="bib16"><number>[16]</number>Rudolf Laban and Lisa Ullman. 1971. *The Mastery of Movement.* </bib>
<bib id="bib17"><number>[17]</number>Andrew Macvean and Judy Robertson. 2012. iFitQuest: a school based study of a mobile location-aware exergame for adolescents. *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '12)*: 359. https://doi.org/10.1145/2371574.2371630</bib>

<bib id="bib18"><number>[18]</number>Ilias El Makrini, Kelly Merckaert, Dirk Lefeber, and Bram Vanderborght. 2017. Design of a collaborative architecture for human-robot assembly tasks. In *IEEE International Conference on Intelligent Robots and Systems (IROS '17)*, 1624–1629. https://doi.org/10.1109/IROS.2017.8205971</bib>
<bib id="bib19"><number>[19]</number>Microsoft. Kinect for Windows. Retrieved January 31, 2019 from https://developer.microsoft.com/en-us/windows/kinect</bib>
<bib id="bib20"><number>[20]</number>Microsoft. Azure Kinect DK. Retrieved May 25, 2021 from https://azure.microsoft.com/en-us/services/kinect-dk/</bib>
<bib id="bib21"><number>[21]</number>Jasmir Nijhar, Nadia Bianchi-Berthouze, and Gemma Boguslawski. 2012. Does Movement Recognition Precision affect the Player Experience in Exertion Games? In *International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN '12)*, 73–82. https://doi.org/10.1007/978-3-642-30214-5_9</bib>
<bib id="bib22"><number>[22]</number>V. Gregory Payne and Larry D. Isaacs. 2011. *Human Motor Development: A LifeSpan Approach.* https://doi.org/10.4324/9781315213040</bib>
<bib id="bib23"><number>[23]</number>R C. Shaefer. 1987. Body Alignment, Postures, and Gait. In *Clinical biomechanics: musculoskeletal actions and reactions*. Wiliams & Wilkins. Retrieved May 24, 2021 from https://www.chiro.org/ACAPress/Body_Alignment.html</bib>
<bib id="bib24"><number>[24]</number>Alex Shaw and Lisa Anthony. 2016. Analyzing the Articulation Features of Children's Touchscreen Gestures. In *ACM International Conference on Multimodal Interaction (ICMI '16)*, 333–340. https://doi.org/10.1145/2993148.2993179</bib>
<bib id="bib25"><number>[25]</number>Manuel Silverio-Fernández, Suresh Renukappa, and Subashini Suresh. 2018. What is a smart device? - a conceptualisation within the paradigm of the internet of things. *Visualization in Engineering* 6, 3: 1–10. https://doi.org/10.1186/s40327-018-0063-8</bib>
<bib id="bib26"><number>[26]</number>Radu-Daniel Vatavu. 2017. Beyond Features for Recognition: Human-Readable Measures to Understand Users' Whole-Body Gesture Performance. *International Journal of Human Computer Interaction* 33, 9: 713–730. https://doi.org/http://dx.doi.org/10.1080/10447318.2017.1278897</bib>
<bib id="bib27"><number>[27]</number>Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2013. Relative accuracy measures for stroke gestures. In *International Conference on Multimodal Interaction (ICMI '13)*, 279–286. https://doi.org/10.1145/2522848.2522875</bib>
<bib id="bib28"><number>[28]</number>Radu-Daniel Vatavu, Lisa Anthony, and Jacob O Wobbrock. 2012. Gestures as point clouds: A $P recognizer for user interface prototypes. In *ACM International Conference on Multimedia Interaction (ICMI '12)*, 273–280. https://doi.org/10.1145/2388676.2388732</bib>
<bib id="bib29"><number>[29]</number>Radu Daniel Vatavu. 2017. Improving gesture recognition accuracy on touch screens for users with low vision. In *SIGCHI Conference on Human Factors in Computing Systems (CHI '17)*, 4667–4679. https://doi.org/10.1145/3025453.3025941</bib>
<bib id="bib30"><number>[30]</number>Evi Verbecque, Luc Vereeck, and Ann Hallemans. 2016. Postural sway in children: A literature review. *Gait & Posture* 49: 402–410. https://doi.org/10.1016/J.GAITPOST.2016.08.003</bib>
<bib id="bib31"><number>[31]</number>Federico Visi, Esther Coorevits, Rodrigo Schramm, and Eduardo Reck Miranda. 2017. Musical instruments, body movement, space, and motion data: Music as an emergent multimodal choreography. *Human Technology* 13, 1: 58–81. https://doi.org/10.17011/ht/urn.201705272518</bib>
<bib id="bib32"><number>[32]</number>Diane Watson, Regan L. Mandryk, and Kevin G. Stanley. 2013. The design and evaluation of a classroom exergame. In *Proceedings of the International Conference on Gameful Design, Research, and Applications (Gamification '13)*, 34–41. https://doi.org/10.1145/2583008.2583013</bib>
<bib id="bib33"><number>[33]</number>Daniel Weinland, Remi Ronfard, and Edmond Boyer. 2006. Free viewpoint action recognition using motion history volumes. *Computer Vision and Image Understanding* 104, 2–3: 249–257. https://doi.org/10.1016/j.cviu.2006.07.013</bib>
<bib id="bib34"><number>[34]</number>Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for nonparametric factorial analyses using only ANOVA procedures. In *Conference on Human Factors in Computing Systems (CHI '11)*, 143–146. https://doi.org/10.1145/1978942.1978963</bib>
<bib id="bib35"><number>[35]</number>Tzu Tsung Wong. 2015. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition* 48, 9: 2839–2846. https://doi.org/10.1016/j.patcog.2015.03.009</bib>