AWARD NUMBER: W81XWH-15-1-0445

TITLE: Automated Assessment of Postural Stability (AAPS)

PRINCIPAL INVESTIGATOR: Iyad Obeid, PhD

CONTRACTING ORGANIZATION: Temple University
1801 N Broad St
Philadelphia PA 19122

REPORT DATE: DECEMBER 2019

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE DECEMBER 2019	**2. REPORT TYPE** Final	**3. DATES COVERED** 15 SEP 2015 – 14 SEP 2019
4. TITLE AND SUBTITLE Automated Assessment of Postural Stability (AAPS)		**5a. CONTRACT NUMBER** W81XWH-15-1-0445
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Iyad Obeid, PhD; Carole Tucker, PhD		**5d. PROJECT NUMBER**
		5e. TASK NUMBER
E-Mail: iobeid@temple.edu		**5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) TEMPLE UNIVERSITY-OF THE COMMONWEALTH SY 1801 N BROAD ST PHILADELPHIA PA 19122-6003		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012		**10. SPONSOR/MONITOR'S ACRONYM(S)**
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This is the final project report for "Automated Assessment of Postural Stability (AAPS)." AAPS and its successor xAAPS use commodity motion-capture tools and custom-built software to create computerized tests of balance and postural stability. These tests are designed specifically for use in austere military environments and are administered without the need for medical or technical experts. All project goals have been successfully met and have been disseminated through numerous peer-reviewed conference and journal publications.

15. SUBJECT TERMS

motion tracking, balance assessment, Microsoft Kinect, concussion assessment

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** USAMRMC
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	Unclassified	12	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

Introduction

The Automated Assessment of Postural Stability (AAPS) is an automated system for measuring balance deficits in warfighters in austere field environments. Its intended purpose is to assess balance deficits following concussion or lower extremity injuries and to provide context in making return-to-duty assessments. The system uses inexpensive commodity hardware to detect body movements, combined with a custom designed software suite that carefully measures those movements, compares them to ideal norms, and scores the results. AAPS was successfully designed to be operable by any minimally trained warfighter and does not require any medical or technical expertise to use. An extended AAPS (xAAPS) was also created that evaluates the quality of kinetic movements (such as lunges and high steps) instead of static balance poses. This work contributes to a growing body of knowledge in which inexpensive and ubiquitous off the shelf sensors can be used in developing field-usable tools that produce actionable biomedical data.

The AAPS project began on September 15, 2015 and ran for four years (including a one-year no-cost extension). The award amount was for $1.36M. All work was performed at Temple University as a collaborative effort between the departments of Electrical & Computer Engineering and Physical Therapy. The goals of the project were all met and have been disseminated by a number of publications.

Keywords

motion tracking, balance assessment, Microsoft Kinect, concussion assessment

Accomplishments

What were the major goals and objectives of the project?

The purpose of this project was to create a portable system for assessing balance in armed forces personnel that could be administered in austere field conditions by personnel with minimal training. Although there are many reasons for assessing an individual's sense of balance, our project focused on balance deficits caused by concussion, traumatic brain injury, and musculoskeletal injury since these are especially relevant to fitness for duty.

Our methodology was to use commodity off-the-shelf motion tracking technology to computerize a well-known balance test called the Balance Error Scoring System (BESS). The motion tracking data is analyzed by a custom designed software tool running on a dedicated conventional Windows PC computer. Through an intelligent self-calibration process, the system is functionally "plug-and-play" meaning that the operator only has to turn on the camera and run the software in order to start the test; adjustments for lighting, background clutter, pixel resolution, camera tilt, and distance are all made automatically. The system prompts the subject to hold various poses and then it tabulates the resulting balance deficits. The BESS test requires subjects to hold a pose for 20 seconds at a time while a trained therapist counts infractions such as excessive sway, foot and hand movement, etc. There are three poses, each done once on firm ground and once again on a foam pad; error scores are capped at ten for each of the six pose-ground combinations.

Using similar technological methods, an expanded system was built for tracking subjects during dynamic movements (in contrast to the static poses of the BESS test). This system tracks subjects as they perform movements associated with a standardized physical fitness test known as the Functional Movement Screen (FMS). The computer was trained to compare motion-captured movements against ideal benchmarks and to produce the same FMS score that is typically computed manually by a trainer or physical therapist.

A significant part of this work was quantifying how accurately the commodity system could track motion. This information was necessary for eliminating systemic sensor biases during balance testing as well as quantifying uncertainty in our results. The calibration studies were performed by simultaneously tracking motion with our system (Kinect, Microsoft) and a state-of-the-art research-grade motion capture system (Vicon).

The ability to use inexpensive commodity sensors to provide actionable biomedical information in austere settings has direct military relevance. One example is that personnel in the field need smart tools to assess warfighters for fitness to return to duty following injury. Other examples may include assessment tools for measuring mental and physical acuity under various forms of duress such as combat/deployment traumatic stress, and deprivation of food, water, and sleep. The literature shows a clear focus on better understanding these issues; techniques such as the ones we have developed and promoted have a certain role in moving the field forward in the future.

The work performed under this grant was organized into four Specific Aims. All aims, goals, and milestones were successfully met, and the work has been disseminated through numerous peer-reviewed publications.

Aim 1: Develop the baseline AAPS system

- Write image processing code in C/C++
- Develop user interface
- Develop AAPS specifically for field use

Aim 2: Calibrate AAPS and perform baseline evaluation

- Healthy subject evaluation
- Concussed subject evaluation
- Musculoskeletal injury subject evaluation

Aim 3: AAPS field evaluation

- Evaluate use by non-clinician operators
- Evaluate AAPS in field conditions

Aim 4: Develop expanded xAAPS test

- Determine dynamic movements to be measured by xAAPS
- Modify existing AAPS software to handle dynamic movements
- Evaluate xAAPS

To maintain military relevance, our team received feedback from a Military Advisory Panel comprised primarily of directors and instructors of the Temple University Reserve Officer Training Corps. The study called for data collection from 145 human subjects: 50 healthy, 50 concussed, and 35 mild musculoskeletal injury AAPS subjects, and 10 healthy xAAPS subjects.

What was accomplished under these goals?

All proposed goals have been achieved. The AAPS and xAAPS systems have been built, calibrated, and tested extensively. Human data has been successfully collected and analyzed. And finally, new directions for military-relevant technology have been identified based on our work. The only aspect of the work that technically remains incomplete is recruitment of five male concussed subjects (out of the proposed 50). We observed that, especially among men, head trauma often goes undiagnosed and untreated, which increases the difficulty of identifying qualified human subjects. The reasons for this are well understood. Injured male subjects are often young and often lack health insurance, which discourages them from seeking medical care. Others seek to minimize the seriousness of their injuries in order to maintain a personal reputation of toughness. And still others seek to avoid the inconvenience of a medical interaction that won't yield any perceived meaningful treatment.

The initial three-year project duration was extended by one year in order to accommodate the outstanding human tests as well as a change in the motion capture sensor. In 2017, Microsoft discontinued production of the Kinect for Xbox. Although Microsoft replaced it with Azure Kinect, that device is primarily intended for cloud-based applications and is not properly suited for standalone applications such as AAPS and xAAPS. Fortunately, we were able to identify a third-party motion capture and skeleton tracking device (Astra, Orbbec) to integrate into our system. The use of standardized body tracking application program interfaces (APIs) simplified this process.

Details of the research performed under this grant can be found in the Products appendix at the end of this report. Our peer-reviewed journal and conference publications present this work in its most detailed form.

The project was completed on budget, costing $1.36M. A cost breakdown includes 50% labor, 10% fringe, 5% non-compensation, and 35% overhead.

What opportunities for training and professional development has the project provided?

This project proudly supported the technical development of a number of trainees. Specifically, a substantial portion of the work was performed by two postdoctoral fellows, Drs. Stephen Glass and Alessandro Napoli, both of whom performed technical and management tasks. Dr. Glass is now on the faculty of Physical Therapy at Radford University. Dr. Napoli is the lead rehabilitation engineer at Jefferson University Hospital in Philadelphia. This work also supported a doctoral student, Christian Ward, as well as numerous Masters' and undergraduate level students. Dr. Ward is a staff scientist at Los Alamos National Labs, and the various other students have moved on to positions that include data scientists and defense contractors. In all cases, the technical development came primarily in the form of significant project responsibilities which often required attaining new skills such as programming, data analysis, and project management. Students were personally mentored by PIs Obeid and Tucker. Students were also exposed to research conferences such as the Military Health System Research Symposium (MHSRS) and the Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) as part of their technical development.

How were the results disseminated to communities of interest?

We have disseminated our results through public talks, journal publications, and conference presentations (see the "Products" section below). We have been regular attendees at the MHSRS meeting in Orlando where we supplement our formal dissemination products with informal discussions with colleagues.

What do you plan to do during the next reporting period to accomplish the goals and objectives?

Nothing to report.

Impact

What was the impact on the development of the principal discipline(s) of the project?

This work has contributed to a growing body of knowledge concerning the application of inexpensive commodity computer tools to military-facing biomedical issues. Nominally, the work we performed used the Microsoft Kinect to computerize the Balance Error Scoring System (the AAPS) and a 3D graphics engine to computerize the Functional Movement Screen (the xAAPS). However, more broadly speaking, this work has shown that any number of fitness or mental acuity screens can be computerized and enhanced, allowing them to be deployed in austere field conditions and operated by servicemembers without specialized medical training. In particular, this work has led to an understanding of how networks of biometric sensors (e.g. gait, motion capture, ECG, and galvanic skin response) can be integrated with immersive virtual reality to create multitask testing paradigms. Tests of this nature can simultaneously assess cognitive function, visual acuity, and decision-making ability under realistically stressful conditions; degree of difficulty can be controlled in real-time using biometric feedback. Although tests of such scope have been investigated by others, they are either performed under minimally realistic environments (marching in place in front of a television screen) or in large immobile lab environments such as the Computer-Assisted Rehabilitation Environment (CAREN). The work we have accomplished under this grant shows a potentially superior middle ground. The ubiquity of versatile third-party biometric sensors can be exploited to create tools that generate actionable results for military decision-makers.

What was the impact on other disciplines?

Nothing to report.

What was the impact on technology transfer?

Nothing to report.

What was the impact on society beyond science and technology?

Our goal has always been to create tools that allow the scientific community to better understand concussion and to help warfighters and athletes alike manage their concussive symptoms. Through our extensive data collection efforts, we have demonstrated the AAPS and xAAPS to several hundred individuals. Our observation has been that there is a great deal of enthusiasm for systems like ours that can be used to bring quantifiable performance results directly to the end-user as opposed to requiring a lab or clinic setting.

Changes/Problems

Changes in approach and reasons for change.

Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents.

Nothing to report.

Actual or anticipated problems or delays and actions or plans to resolve them.

Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents.

Nothing to report.

Changes that have a significant impact on expenditures.

Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents.

Nothing to report.

Products

Published

Glass, S. M., Napoli, A., Thompson, E. D., Obeid, I., & Tucker, C. A. (2019). Validity of an Automated Balance Error Scoring System. *Journal of Applied Biomechanics*, 1–16. https://doi.org/10.1123/jab.2018-0056

Napoli, A., Glass, S. M., Tucker, C. A., & Obeid, I. (2018). The Expanded Automatic Assessment of Postural Stability: the xAAPS. In *Military Health System Research Symposium (MHSRS)*. Orlando, FL.

Glass, S. M., Napoli, A., Obeid, I., & Tucker, C. A. (2018). Effects of Concussion History on Center of Mass Motion During Modified Balance Error Scoring System (BESS) Testing in Women. In *Military Health System Research Symposium (MHSRS)*. Orlando, FL.

Glass, S. M., Napoli, A., Obeid, I., & Tucker, C. A. (2017). Inverse Kinematics Using Portable, Low-Cost Sensor Technology. In *Gait & Clinical Movement Analysis Society Annual Meeting*. Salt Lake City, UT.

Napoli, A., Glass, S., Ward, C., Tucker, C., & Obeid, I. (2017). Performance analysis of a generalized motion capture system using Microsoft Kinect 2.0. *Biomedical Signal Processing and Control*, *38*, 265–280. https://doi.org/10.1016/J.BSPC.2017.06.006

Napoli, A., Glass, S. M., Tucker, C., & Obeid, I. (2017). The Automated Assessment of Postural Stability: Balance Detection Algorithm. *Annals of Biomedical Engineering*. https://doi.org/10.1007/s10439-017-1911-8

Napoli, A., Ward, C. R., Glass, S. M., Tucker, C., & Obeid, I. (2016). Automated Assessment of Postural Stability System. In *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 6090–6093). IEEE. https://doi.org/10.1109/EMBC.2016.7592118

In Development

Glass, S. M., Napoli, A., Obeid, I., & Tucker, C. A. (TBD) Evaluation of Static and Dynamic Balance Outcomes Using Field-Expedient Methods

Napoli, A., Glass, S. M., Tucker, C. A., Obeid, I. (TBD) Design and Validation of a Dynamic Movement and Balance Assessment Tool

Napoli, A., Glass, S. M., Tucker, C. A., Obeid, I. (TBD) Predicting Concussion from Static Balance Assessment Data

Participants & Other Collaborating Organizations

What individuals have worked on the project?

Name:	Iyad Obeid, PhD
Project Role:	co-Principal Investigator
Identifier:	https://orcid.org/0000-0002-5796-843X
Person-Months:	11.3
Contribution:	Dr. Obeid contributed to project design and management, analyzed data, supervised data marshalling, wrote quarterly reports, and contributed to all technical publications.

Name:	Carole Tucker, PhD
Project Role:	co-Principal Investigator
Identifier:	https://orcid.org/0000-0002-9408-5898
Person-Months:	10.2
Contribution:	Dr. Tucker contributed to project design and management, IRB preparation, human subject protocol design, data collection, and analysis, and all technical publications.

Name:	Alessandro Napoli
Project Role:	Postdoctoral Fellow
Identifier:	https://orcid.org/0000-0002-4061-3747
Person-Months:	35.2
Contribution:	Was responsible for managing all aspects of the software organization and development and contributed heavily to actual software creation. He managed the graduate RAs and the undergraduates, contributed to data collection and analysis, and took a leading role on all technical publications.

Name:	Stephen Glass
Project Role:	Postdoctoral Fellow
Identifier:	https://orcid.org/0000-0001-6263-527X
Person-Months:	27
Contribution:	Was responsible for managing all aspects of data planning, collection and analysis, including IRB development. He managed junior students and took a leading role in all technical publications.

Name:	Christian Ward
Project Role:	Graduate Research Assistant
Identifier:	https://orcid.org/0000-0001-5394-8135
Person-Months:	19
Contribution:	Provided software development and data analytics support; contributed to management of undergraduate students.

Name:	Nicholas Satterthwaite
Project Role:	Graduate Researcher
Person-Months:	11.8
Contribution:	Code development and documentation

Name: Victor Espinoza
Project Role: Graduate Researcher
Person-Months: 4.2
Contribution: Code development and documentation

Name: Anirvan Majumdar
Project Role: Graduate Research Assistant
Person-Months: 2
Contribution: Code development and documentation

Name: Zachary Kane
Project Role: Undergraduate Researcher
Person-Months: 4.1
Contribution: Code development and documentation

Name: Evan Stecco
Project Role: Undergraduate Researcher
Person-Months: 3.5
Contribution: Code development and documentation

Name: Orlena Roe
Project Role: Undergraduate Researcher
Person-Months: 3.3
Contribution: Contributed to software testing and development of the graphical interface.

Name: Elizaveta Ibeme
Project Role: Undergraduate Researcher
Person-Months: 2.1
Contribution: Contributed to software testing and wrote software tools for data analysis.

Name: Bhautik Amin
Project Role: Undergraduate Researcher
Person-Months: 1.2
Contribution: Code development and documentation

Name: Paula Oliveira
Project Role: Undergraduate Researcher
Person-Months: 0.4
Contribution: Code development and documentation

Name: Lillian Veloso
Project Role: Undergraduate Researcher
Person-Months: 0.9
Contribution: Code development and documentation

Name: Alexander Wroblewski
Project Role: Undergraduate Researcher
Person-Months: 0.3
Contribution: Code development and documentation

Name: Von Kaukeano
Project Role: Undergraduate Researcher
Person-Months: 0.6
Contribution: Code development and documentation

Name: Chad Martin
Project Role: Undergraduate Researcher
Person-Months: 0.6
Contribution: Code development and documentation

Has there been a change in the other active support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nothing to report.

What other organizations have been involved as partners?

Nothing to report.

Special Reporting Requirements

See Quad Chart in the Appendix

Appendices

Quad Chart

Products

Automated Assessment of Postural Stability (AAPS)

Log Number: MR141272
Award Number: W81XWH-15-1-0445

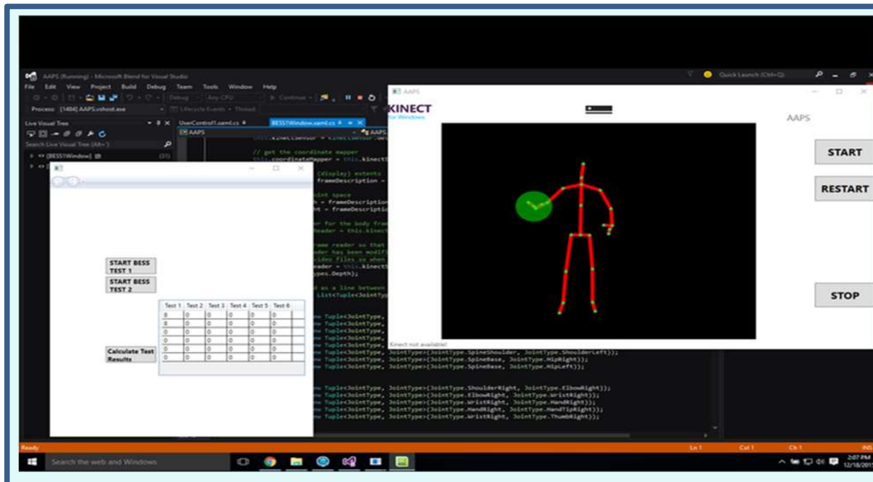
PI: Iyad Obeid & Carole A. Tucker **Org:** Temple University **Award Amount:** $1.36M

Study/Product Aim(s)

- Develop a fully functional proof-of-concept system (AAPS), featuring a complete software suite for automatically administering the Balance Error Scoring System (BESS) test.
- Calibrate the AAPS on healthy, concussion, and musculoskeletal injury subjects.
- Fully field test AAPS to ensure use by non-medical technicians.
- Expansion of AAPS to include dynamic postural tasks.

Approach

We aim to develop, calibrate, and field test a system for quantifying the impact postural and balance injuries using the Microsoft Kinect, an inexpensive motion capture system. The system will administer and score the BESS in field conditions without requiring a medically trained operator. We will expand the BESS to include dynamic tasks (lunge, squat, etc.) to better assess readiness for return to active military duty post mild TBI.



Screenshot showing development of C# code, prototype proof-of-concept GUI interface, and live image capture skeleton.

Timeline and Cost

Activities CY	15	16	17	18
AAPS system development				
Calibrate AAPS (n=50 subjects)				
Field test AAPS				
Expand AAPS – Dynamic tasks				
Estimated Budget ($k)	**$200**	**$500**	**$500**	**$200**

Updated: 12 February 2020

Goals/Milestones

CY15 Goals – System development
✓ Port existing system from Matlab to C/C++ *[100%]*
✓ Develop user interface for automatic test administration *[100%]*

CY16 Goal – Calibration and Field Testing
✓ Determining reference scores for healthy, concussion, and musculoskeletal injury subjects *[100%]*
✓ Comparing performance to gold standard benchmarks *[10%]*
✓ Optimizing design for use by non-medical technicians *[100%]*

CY17 Goal – System expansion
✓ Determining optimal dynamic tasks for assessment *[100%]*
✓ Updating software to handle dynamic task tracking *[100%]*

CY18 Goal – System optimization
✓ Complete expansion & optimize software via beta testing *[100%]*

Comments/Challenges/Issues/Concerns

- none

Budget Expenditure to Date

Projected Expenditure: ~$1.36M
Actual Expenditure: ~$1.36M

Automated Assessment of Postural Stability System

Alessandro Napoli*, Christian R. Ward*, Stephen M. Glass**, Carole Tucker**, and Iyad Obeid*

Abstract — **The Balance Error Scoring System (BESS) is one of the most commonly used clinical tests to evaluate static postural stability deficits resulting from traumatic brain events and musculoskeletal injury. This test requires a trained operator to visually assess balance and give the subject a performance score based on the number of balance "errors" they committed. Despite being regularly used in several real-world situations, the BESS test is scored by clinician observation and is therefore (a) potentially susceptible to biased and inaccurate test scores and (b) cannot be administered in the absence of a trained provider. The purpose of this research is to develop, calibrate and field test a computerized version of the BESS test using low-cost commodity motion tracking technology. This 'Automated Assessment of Postural Stability' (AAPS) system will quantify balance control in field conditions. This research goal is to overcome the main limitations of both the commercially available motion capture systems and the standard BESS test. The AAPS system has been designed to be operated by a minimally trained user and it requires little set-up time with no sensor calibration necessary. These features make the proposed automated system a valuable balance assessment tool to be utilized in the field.**

I. Introduction

Traumatic Brain Injury (TBI) is defined as brain damage generated by an external mechanical force. Such forces can be caused by rapid acceleration or deceleration, blast waves, crush, or impact or penetration by a projectile. TBI can lead to temporary or permanent impairment of cognitive, physical and physiological functions. TBI contributes to a substantial number of deaths and permanent disability and it is a contributing factor to a third of all injury-related deaths in the United States. About 75% of TBIs that occur each year are concussions or other forms of mild traumatic brain injury (mTBI) [1]. TBI can cause a vast series of symptoms and it can be challenging to diagnose given the number of confounding factors involved. This is especially true for traumatic events that need to be diagnosed in the field, such as in military or athletic play scenarios [2].

Concussion diagnostic and management decisions are based on many elements, including symptom presentation, physical examinations and specialized tests designed to detect deficits resulting from concussive injuries [3]–[5]. Consequently, concussion assessment requires a multidisciplinary team of professionals and a series of specialized tests [6]. With the increased attention on and recognition of concussive injuries, there is a need for new assessment tools that, combined with more traditional techniques, will help evaluate concussion injuries more accurately [7], [8]. For some time, balance testing has been used in clinical settings as a reliable and valuable assessment tool to evaluate neurological functioning in subjects suffering from concussion and musculoskeletal injuries [9], [10]. There is a need to move concussion testing from the clinic onto the field, into the locker room, or out into the military theater [10], [11]. In these cases, computerized systems are often limited by size, weight, portability, ease of setup and use, and ease of calibration. Furthermore, there is a demonstrated need for such systems to be operable by non-medical or non-expert personnel, such as coaches and ordinary military corpsmen. We have developed a system that addresses these issues: a computerized system for administering and scoring the BESS test in a wide array of non-clinic locations using an inexpensive commodity motion-capture system.

II. Material and Methods

A. The BESS Test

The Balance Error Scoring System (BESS) is one of the most commonly used clinical tests for assessing postural stability following concussion [12]. It measures standing posture and balance related impairments [11]–[13]. Patients hold three different poses on two different surfaces (firm and foam), for 20 seconds each with their hands on hips and eyes closed. The three stances are: double leg stance with feet flat on testing surface; single leg stance, with the subject standing on the non-dominant leg with the contralateral limb held at approximately 20° of hip flexion, 45° of knee flexion; tandem stance, with one foot placed in front of the other with heel of the anterior foot touching the toe of the posterior foot. The subject's non-dominant leg is in the posterior position. A specially trained clinician observes the patient during each pose and records the number of balance "errors". These errors include:

- Moving the hands off the hips
- Opening the eyes
- Step, stumble or fall
- Abduction or flexion of the hip beyond 30°
- Lifting the forefoot or heel off the testing surface
- Remaining out of the proper testing position for longer than 5 seconds

Because the BESS test is scored by clinician observation, it is potentially susceptible to biased and inaccurate test scores that could lead to inappropriate return to duty or play before adequate recovery from the traumatic event [8]. The BESS is further limited by the need for properly trained clinicians to simultaneously administer, score and interpret the test and to ensure patient safety. Such an expert examiner may not be readily available in field situations. Finally, the BESS only focuses on static postural balance tasks and lacks assessment of more dynamic postural tasks. Testing only for static stability

* Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122, USA

** Department of Physical Therapy, Temple University, Philadelphia, PA, 19122, USA

A. Napoli (corresponding author; e-mail: a.napoli@temple.edu).

may not capture other important domains of balance, including dynamic or cognitive aspects [14].

B. The Automated Assessment of Postural Stability System

The purpose of this research is to develop, calibrate and field test an Automated Assessment of Postural Stability (AAPS) system to quantify balance control in military field conditions. The AAPS' objective is to evaluate postural and balance deficits due to concussion and musculoskeletal injuries commonly seen in active duty military personnel for return-to-duty assessment. The AAPS system has been developed in the C# programming language in the Microsoft Visual Studio 2015 programming environment and the .NET framework. The system includes a comprehensive graphical interface (GUI) to guide the operator and the subjects through the BESS test. It also provides user controls, data management features, a real-time display of the detected body and intuitive visual feedback on the AAPS tracking capabilities. Furthermore, the GUI has been designed to be user-friendly and its use only requires minimal training and experience. The current AAPS GUI is shown in Fig. 1.

C. System Design

The AAPS system utilizes an inexpensive Microsoft Kinect v2.0 motion capture sensor [15], [16] and a custom designed software suite for Microsoft Windows to objectively track kinemtics during the Balance Error Scoring System (BESS) test. Fig.2 displays the system workflow. Set-up time is minimal, and no additional calibration time is required. These features make the AAPS a valuable balance assessment tool to be utilized in the field. It only requires the patient to assume the right stance and the system starts extracting tracking information from the sensor data stream and uses a GUI to display the tracked skeleton. By using some simple visual feedback features, the system provides the user with information regarding the subject's position, joints and eyes detection. This helps the non-trained user in positioning the subject for the test and guarantees that all the necessary parameters are tracked correctly before starting the test. During the test, in each frame, the AAPS extracts two types of data from the Kinect sensor: infrared camera videos, which will be saved in the computer memory for optional off-line analysis; and the subject's joint location and eye data. These will be used, frame-by-frame, to build arrays containing the three metrics that will be then used to evaluate errors, namely relative joint distances, segment angles and eye state. After test completion, these metrics will be used to compute balance errors as explained in more detail in the next section.

III. AAPS BALANCE ERROR ASSESSMENT

Once the subject is in position, the test is started and the AAPS guides the user through the set of six stances and then automatically computes balance errors. The AAPS does not require baseline recordings for balance assessment. It only uses a few seconds worth of calibration data that are acquired right before the start of each BESS stance, with the patient standing still in the right test position. The calibration data are used to learn subject dependent features, such as body shape and postural characteristics. Furthermore, calibration data are necessary to compensate for any small fluctuations in joint tracking that may occur. This is necessary, since even when a subject stands as still as possible, the Kinect typically shows joint locations fluctuating to some small degree. In order to not erroneously flag such flutter as balance errors, a brief calibration recording is used to determine the mean and standard deviation of each joint location. During actual testing, only movements exceeding five standard deviations are flagged as balance errors. In order to detect a subject's postural changes, three biometric measurements are used: joint distances, segment angles and eye status. These variables are computed in each video frame but error detection is performed on the output of a moving average filter with a window time duration of five frames. This filter was implemented to further stabilize the system error detection by reducing high frequency artifacts. The average of these metrics during calibration are compared to their corresponding actual values during testing to detect changes in postural stability. Equation 1 shows how the average distance calibration values are computed.

$$\hat{d}_{cal} = \frac{1}{N_{cal}} \sum_{fr=1}^{N_{cal}} \sqrt{\left(joint_{1_x}(fr) - joint_{2_x}(fr)\right)^2 + (joint_{1y}(fr) - joint_{2y}(fr))^2} \quad (1)$$

Equation 1 provides one calibration value for each biometric variable used. N_{cal} is the number of calibration frames used, fr is the frame index, and $joint_{1_x}$ and $joint_{1y}$ are the 2-D coordinates of a joint of interest.

After calibration, during postural testing, the metric variables are computed once per frame and the resulting

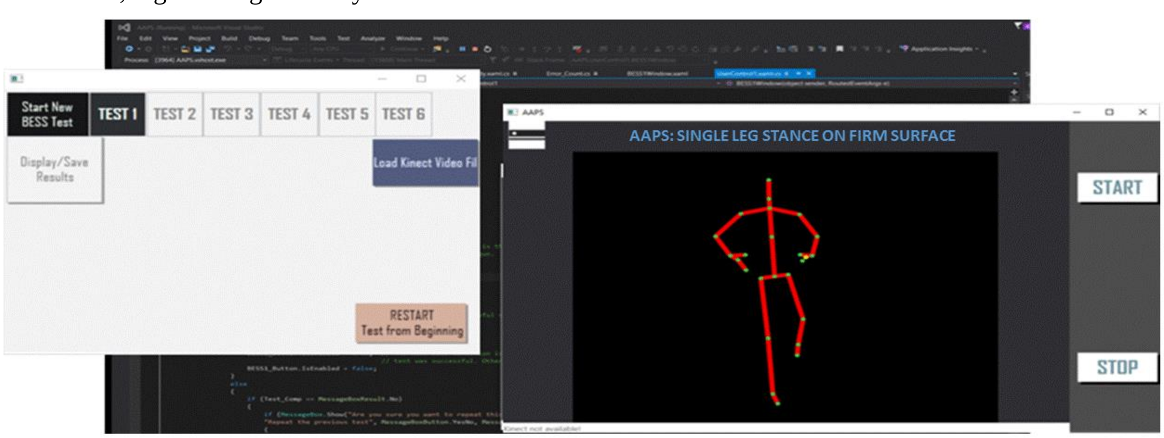


Figure 1: The AAPS Graphical User Interface

sequences $\{d_{fr}\}_{fr=1}^{N}$ stored in memory arrays (See Equation 2). Error calculation is performed by comparing the average calibration values to the output of a moving average filter with a sliding window duration of five frames (see Equation 3).

Sequence: $\{d_{fr}\}_{fr=1}^{N}$ where fr is the video frame index

$$d_{fr} = \sqrt{\begin{matrix}\left(joint_{1_x}(fr) - joint_{2_x}(fr)\right)^2 + \\ (joint_{1y}(fr) - joint_{2y}(fr))^2\end{matrix}} \tag{2}$$

where $\{\widehat{d}_i\}_{i=1}^{N-n+1}$ is the moving average output window, where n is the number of frames in the moving average (default is 5).

$$\widehat{d}_i = \frac{1}{n}\sum_{fr=1}^{1+n-1} d_{fr} \tag{3}$$

If the changes in distances and angles between the calibration and the actual values are larger than five times the standard deviation in the calibration data, then the AAPS system will detect an error. The error detection criterion is shown in Equation 4.

$$e_i = \left|\hat{d}_i - \hat{d}_{cal}\right| > 5 * \sigma_{cal} \tag{4}$$

where

$$\sigma_{cal} = \sqrt{\frac{1}{N_{cal}}\sum_{fr=1}^{N_{cal}}(d_{fr} - \hat{d}_{cal})^2} \tag{5}$$

At the end of a complete BESS test, the errors with respect to each condition and the error types are stored in a file together with patient, test information and video files corresponding to the scored tests.

The automated system's BESS scoring is based on the detection of four types of errors that are generated when: 1) the hands are off the hips; 2) the foot distance changes; 3) the angle between the center of the shoulders and the center of the hips becomes larger than 30°; 4) the subjects open their eyes.

IV. MAIN CHALLENGES OF A KINECT-BASED AUTOMATED BALANCE SYSTEM

As a motion tracking system, the Kinect sensor has some clear limitations as compared to professional or clinical grade digital motion capture systems that make use of more sophisticated sensors and body markers. In this section, some of the main challenges are introduced:

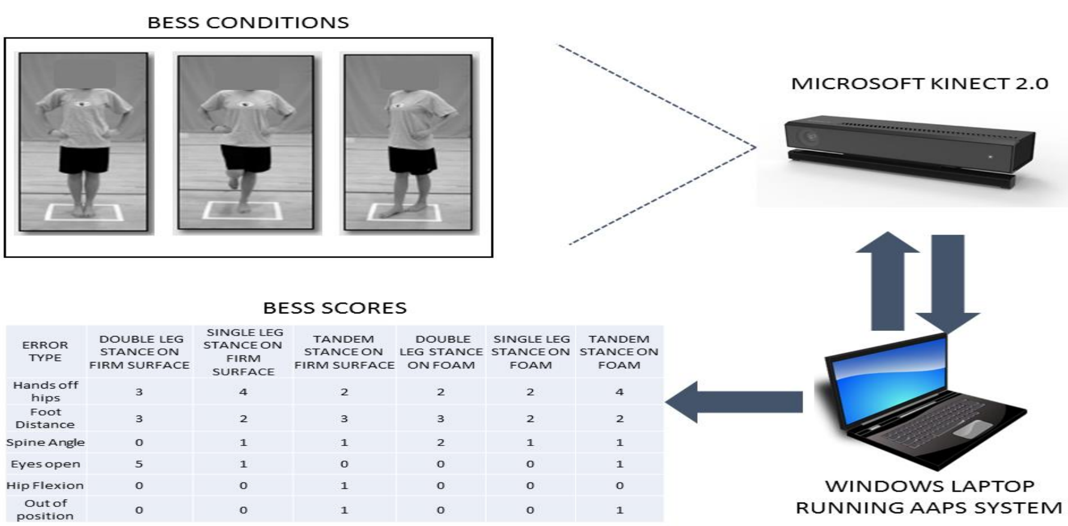
1) The Kinect sensor cannot stably detect foot positions on certain surfaces, especially during static testing. This is likely caused when the floor interferes with the infrared depth signal projected by the Kinect. This issue is being addressed by determining the optimal sensor height, angle, and distance from subject.

2) The Kinect sensor can successfully detect subject's trunk lateral movements from the vertical position, in the frontal plane. However, the sensor is not as effective when tracking trunk changes that occur in the transverse plane. The AAPS will account for this limitation by combining other joint positions to estimate 3-dimensional trunk movements.

3) During the Tandem Stance, with the subject facing the Kinect, the sensor and the body tracking system have issues detecting the back leg that is hidden behind the other leg. Even though the information on the back leg is not available to the AAPS, our system will infer changes in the back leg position by using data derived from other joints.

4) The eye gaze sensor tracking capabilities yield the best results when the subject's eyes are aligned with the sensor's color camera. This can be challenging when the sensor is placed at the height of the subject's trunk. This issue can be solved by performing a series of face orientation computations to guarantee maximal eye detection accuracy during the test.

An important implementation detail relates to the Kinect sensor, which provides a number of different data streams, such as both infrared and color videos, depth information, tracking data, and gesture and face recognition. These data require intensive real-time signal and image processing that is carried out on the computer video card using Microsoft proprietary algorithms. While the real-time data are valuable for the AAPS to perform accurately, the resulting computational burden makes the system performance unpredictable and unstable. A performance bottle neck arises as the computer handles this Kinect data in real-time while simultaneously running other system processes. Microsoft's



ERROR TYPE	DOUBLE LEG STANCE ON FIRM SURFACE	SINGLE LEG STANCE ON FIRM SURFACE	TANDEM STANCE ON FIRM SURFACE	DOUBLE LEG STANCE ON FOAM	SINGLE LEG STANCE ON FOAM	TANDEM STANCE ON FOAM
Hands off hips	3	4	2	2	2	4
Foot Distance	3	2	3	3	2	2
Spine Angle	0	1	1	2	1	1
Eyes open	5	1	0	0	0	1
Hip Flexion	0	0	1	0	0	0
Out of position	0	0	1	0	0	1

Figure 2: The AAPS operational diagram

Kinect libraries respond to limitations in computing power by automatically reducing the amount of information coming in from the various Kinect sensors by continually adjusting the framerate during run-time. In ideal conditions, the sensor returns data streams at 30 frames/second. However, during our testing, it was not unusual to see the frame rate drop below 10 frames/second, even though the AAPS was tested on high-end laptops with late model quad-core processors and dedicated video cards. In order to account for this variability in frame rate, the AAPS is implemented using a series of timer variables that allow it to keep track of real-time system performance. Furthermore, the variable and non-controllable frame rate will affect the operation of the moving average, the filter time constant will depend on the frame rate. To compensate for such variability, the moving average size can be chosen dynamically and adapted to the instantaneous frame rate to keep the filter cut-off frequency constant. The system real-time performance does not appear to be a main concern when performing static testing, but may become critical in the future when implementing dynamic balance testing.

V. Future work

In future work, the AAPS system will be further expanded to introduce new functionalities. Namely, the AAPS will not only perform the BESS test, but it will also provide kinematic metrics that will better capture the examinee's balance deficits by using time measures of performance and dynamic testing of performance during fundamental movements. Furthermore, the AAPS software suite will be expanded by integrating physiological data recorded from the examinee in real time, into the concussion battery of tests, while performing the automated tests. These data will be used to create personalized test scoring procedures and thus increase the system capability to evaluate cognitive and behavioral elements of the examinee. Moreover, assessing the examinee's physical engagement and states will greatly improve the AAPS ability to account for suboptimal effort levels, that is a major concern in computer-based testing for concussion assessment [10].

VI. Conclusion

The proposed research presents an automated system for quantifying balance control deficits due to traumatic brain injury. The AAPS system consists of two components: a Microsoft Kinect motion sensor and a Windows laptop running a specifically developed software suite. The AAPS main goal is to perform the BESS test in filed conditions, with no prior calibration and no medical or experienced personnel needed. The AAPS provides reliable and repeatable balance assessment results that are important in managing the return to duty determination after traumatic brain events.

This work has been demonstrated the validity of the AAPS for performing automated BESS tests. In this view, the set-up simplicity, robustness, test repeatability and the user-friendly approach of the AAPS make the system the perfect platform to develop the next generation of in the field concussion evaluation tests.

Acknowledgment

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs through the Neurosensory and Rehabilitation Research Award Program under Award No. W81XWH-15-1-0045. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense.

References

[1] C. V. Faul M, Xu L, Wald MM, "Traumatic brain injury in the United States: emergency department visits, hospitalizations, and deaths," *Centers Dis. Control Prev. Natl. Cent. Inj. Prev. Control*, pp. 891–904, 2010.

[2] C. J. Bryan, "Multiple traumatic brain injury and concussive symptoms among deployed military personnel.," *Brain Inj.*, vol. 27, no. 12, pp. 1333–7, 2013.

[3] C. M. Marshall, "Sports-related concussion: A narrative review of the literature.," *J. Can. Chiropr. Assoc.*, vol. 56, no. 4, pp. 299–310, 2012.

[4] W. Erdman, John, M. Oria, and L. Pillsbury, *Nutrition and traumatic brain injury: improving acute and subacute health outcomes in military personnel*. National Academies Press, 2011.

[5] T. S. Fu, R. Jing, S. R. McFaull, and M. D. Cusimano, "Health & Economic Burden of Traumatic Brain Injury in the Emergency Department," *Can. J. Neurol. Sci. / J. Can. des Sci. Neurol.*, pp. 1–10, 2016.

[6] T. Covassin, R. Moran, and K. Wilhelm, "Concussion symptoms and neurocognitive performance of high school and college athletes who incur multiple concussions.," *Am. J. Sports Med.*, vol. 41, no. 12, pp. 2885–9, 2013.

[7] A. Napoli, M. Barbe, K. Darvish, and I. Obeid, "Assessing traumatic brain injuries using EEG power spectral analysis and instantaneous phase.," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2012, pp. 4692–5, Jan. 2012.

[8] P. Dave, I. Obeid, and C. Tucker, "Automated System for Balance Error Scoring," in *Biodevices*, 2014, pp. 329–333.

[9] P. Mccrory, W. H. Meeuwisse, M. Aubry, B. Cantu, R. J. Echemendia, L. Engebretsen, K. Johnston, J. S. Kutcher, M. Raftery, A. Sills, B. W. Benson, G. A. Davis, R. G. Ellenbogen, K. Guskiewicz, S. A. Herring, G. L. Iverson, and B. D. Jordan, "Consensus statement on concussion in sport : the 4th International Conference on Concussion in Sport held in Zurich , November 2012," *J. Sports Med.*, no. 47, pp. 250–258, 2013.

[10] S. P. Broglio, M. S. Ferrara, S. N. Macciocchi, T. A. Baumgartner, and R. Elliott, "Test-Retest Reliability of Computerized Concussion Assessment Programs," *J. Athl. Train.*, vol. 42, no. 4, pp. 509–514, 2007.

[11] K. M. Guskiewicz, S. E. Ross, and S. W. Marshall, "Postural Stability and Neuropsychological Deficits After Concussion in Collegiate Athletes," *J. Athl. Train.*, vol. 36, no. 3, pp. 263–273, 2001.

[12] D. R. Bell, K. M. Guskiewicz, M. A. Clark, and D. A. Padua, "Systematic Review of the Balance Error Scoring System," *Sport. Heal. A Multidiscip. Approach*, vol. 3, no. 3, pp. 287–295, 2011.

[13] B. L. Riemann and K. M. Guskiewicz, "Effects of Mild Head Injury on Postural Stability as Measured Through Clinical Balance Testing," *J. Athl. Train.*, vol. 35, no. 1, pp. 19–25, 2000.

[14] L. A. King, F. B. Horak, M. Mancini, D. Pierce, K. C. Priest, J. Chesnutt, P. Sullivan, and J. C. Chapman, "Instrumenting the balance error scoring system for use with patients reporting persistent balance problems after mild traumatic brain injury," *Arch. Phys. Med. Rehabil.*, vol. 95, no. 2, pp. 353–359, 2014.

[15] S. Choppin and J. Wheat, "The potential of the Microsoft Kinect in sports analysis and biomechanics," *Sport. Technol.*, vol. 6, no. 2, pp. 78–85, 2013.

[16] X. Xu and R. W. McGorry, "The validity of the first and second generation Microsoft Kinect for identifying joint center locations during static postures," *Appl. Ergon.*, vol. 49, pp. 47–54, 2015.

INVERSE KINEMATICS USING PORTABLE, LOW-COST SENSOR TECHNOLOGY

Stephen M. Glass[1], Alessandro Napoli[1], Iyad Obeid, and Carole A. Tucker[2]
[1] Temple University, Philadelphia, PA, USA
E-mail: stephen.glass@temple.edu

INTRODUCTION

The Microsoft Kinect 2.0™ is a portable platform that allows for low-cost motion capture in field-based and clinical settings [1]. The sensor performs remarkably well as an out-of-the-box technology, particularly considering its limitations as a single-camera, markerless system. These achievements notwithstanding, data from the Kinect™ are not readily comparable to conventional kinematics. Specifically, the sensor generates joint displacement time histories without rigid body definitions or range of motion constraints. These measurement errors tend to have large effects on any joint angles calculated by the user. It is possible, therefore, that the quality of Kinect™ data could be improved for research purposes through the introduction of modeling constraints commonly applied in inverse kinematics. Our purpose is to demonstrate the performance of Kinect-based inverse kinematics solutions in comparison with similar data acquired using standard laboratory technology.

CLINICAL SIGNIFICANCE

Applying modeling constraints to motion capture data acquired using low-cost sensors may increase the accuracy of clinical measurement techniques. These methods could support diagnostics and clinical decision-making by enabling the collection of higher quality data.

METHODS

Eight participants (4 males/4 females, 25.6 ± 2.5 years, age, 170.8 ± 10.1 cm, 66.2 ± 12.8 kg) performed trials of a series of commonly used clinical movement tests as data were captured using 1) a Qualisys Oqus motion capture system (Q), and 2) a Microsoft Kinect 2.0™ (K). Inverse kinematics solutions were computed in OpenSim 3.3 using marker (Q) or joint center (K) trajectories. Both analyses were conducted using the Vicon Plug-In-Gait model with segment/trajectory relationships (i.e. the “marker set”) modified to match the source of the data. Sagittal plane hip and knee angles were analyzed to compare performance between the two systems. After resampling the raw data from the Kinect sensor (which samples at a variable rate), cross-correlation and RMS error were calculated on the synchronized signals.

DEMONSTRATION

Test	**DS**	**HS**	**FSTS**	**ABL**	**MIP**
Hip cor.	0.89 (0.27)	0.90 (0.06)	0.98 (0.01)	0.87 (0.13)	0.80 (0.11)
Hip err.	22.89 (7.51)	14.82 (5.17)	23.40 (6.65)	22.98 (18.02)	23.25 (9.50)
Knee cor.	0.89 (0.25)	0.89 (0.06)	0.99 (0.01)	0.90 (0.08)	0.86 (0.10)
Knee err.	12.41 (8.64)	11.95 (3.45)	14.70 (18.62)	11.70 (5.16	16.00 (6.14)

OHS = Overhead Squat, HS = Hurdle Stepping, FSTS = 5x Sit-to-Stand, ABL = Alternating Barbell Lunge, MIP = Marching in Place, corr. = correlation (*r*), err. = RMS error (deg.)

Figure 1 shows hip flexion time series derived from both sensors during a representative trial of the Five Times Sit-to-Stand test. Kinect tracking of the sagittal plane hip and knee angles using inverse kinematics performs favorably in comparison with a gold-standard system. We do note, however, that whereas temporal features of the signal are well preserved, the magnitude of angular displacements can be overestimated by the Kinect.

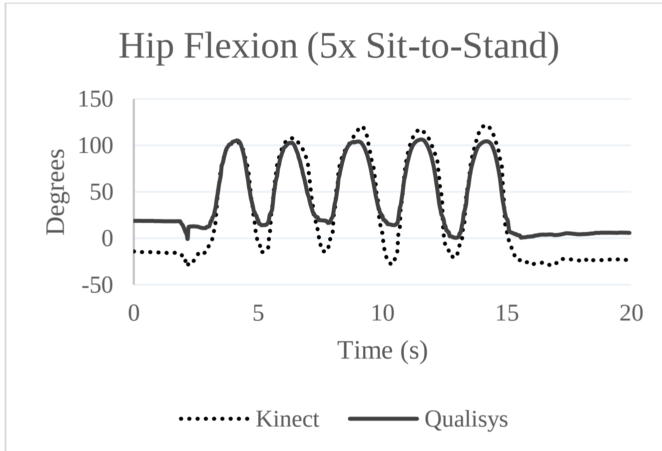


Figure 1

SUMMARY

Our laboratory has recently shown that Kinect 2.0™ data is suitable for instrumenting simple field-expedient clinical tests [2]. With the present work, we expand on our automated scoring algorithm research to improve the quality of this sensor as a robust motion capture tool. These data show that accuracy of certain angular kinematics from low-cost sensors such as the Microsoft Kinect 2.0™ may approach gold-standard criteria with commonly used inverse kinematic modeling techniques. The most substantial benefits likely derive from rigid body definitions and joint-specific range of motion limits, neither of which are applied to Kinect™ data.

Our approach is implemented with openly available software (OpenSim) using a modified version of a familiar kinematic model, the Vicon Plug-In-Gait model. This workflow could greatly benefit clinics and mobile laboratories requiring high quality data without the time and expense typical of multicamera systems. In future work, we will demonstrate the performance of a Kinect-based inverse kinematics analysis in tracking complex, multiplanar movement during a variety of dynamic posture tasks.

REFERENCES

1. Clark, R. A., et al. (2015). Gait Posture 42(2): 210-213.
2. Napoli et al. (2016) EMBC Proceedings, 6090-6093

ACKNOWLEDGMENTS

This work was supported by DoD/USAMRMC Award # W81XWH-15-1-0445.

DISCLOSURE STATEMENT

The authors have no conflicts of interest to disclose.

Annals of Biomedical Engineering, Vol. 45, No. 12, December 2017 (© 2017) pp. 2784–2793
DOI: 10.1007/s10439-017-1911-8

The Automated Assessment of Postural Stability: Balance Detection Algorithm

Alessandro Napoli,[1] Stephen M. Glass,[2] Carole Tucker,[1,2] and Iyad Obeid[1]

[1]Department of Electrical & Computer Engineering, Temple University, Philadelphia, PA 19122, USA; and [2]Department of Physical Therapy, Temple University, Philadelphia, PA 19140, USA

(Received 17 July 2017; accepted 27 August 2017; published online 30 August 2017)

Associate Editor Karol Miller oversaw the review of this article.

Abstract—Impaired balance is a common indicator of mild traumatic brain injury, concussion and musculoskeletal injury. Given the clinical relevance of such injuries, especially in military settings, it is paramount to develop more accurate and reliable on-field evaluation tools. This work presents the design and implementation of the automated assessment of postural stability (AAPS) system, for on-field evaluations following concussion. The AAPS is a computer system, based on inexpensive off-the-shelf components and custom software, that aims to automatically and reliably evaluate balance deficits, by replicating a known on-field clinical test, namely, the Balance Error Scoring System (BESS). The AAPS main innovation is its balance error detection algorithm that has been designed to acquire data from a Microsoft Kinect® sensor and convert them into clinically-relevant BESS scores, using the same detection criteria defined by the original BESS test. In order to assess the AAPS balance evaluation capability, a total of 15 healthy subjects (7 male, 8 female) were required to perform the BESS test, while simultaneously being tracked by a Kinect 2.0 sensor and a professional-grade motion capture system (Qualisys AB, Gothenburg, Sweden). High definition videos with BESS trials were scored off-line by three experienced observers for reference scores. AAPS performance was assessed by comparing the AAPS automated scores to those derived by three experienced observers. Our results show that the AAPS error detection algorithm presented here can accurately and precisely detect balance deficits with performance levels that are comparable to those of experienced medical personnel. Specifically, agreement levels between the AAPS algorithm and the human average BESS scores ranging between 87.9% (single-leg on foam) and 99.8% (double-leg on firm ground) were detected. Moreover, statistically significant differences in balance scores were not detected by an ANOVA test with alpha equal to 0.05. Despite some level of disagreement between human and AAPS-generated scores, the use of an automated system yields important advantages over currently available human-based alternatives. These results underscore the value of using the AAPS, that can be quickly deployed in the field and/or in outdoor settings with minimal set-up time. Finally, the AAPS can record multiple error types and their time course with extremely high temporal resolution. These features are not achievable by humans, who cannot keep track of multiple balance errors with such a high resolution. Together, these results suggest that computerized BESS calculation may provide more accurate and consistent measures of balance than those derived from human experts.

Keywords—Mild traumatic brain injury, Concussion detection, Field-expedient balance test, Automated BEES, Automatic balance error scoring detection, Kinect, Return-to-duty evaluation, On-field automatic balance detection.

INTRODUCTION

The incidence of mild traumatic brain injury (mTBI), concussion, and musculoskeletal injuries has increased in the patient population of the Department of Defense (DoD) and the Veterans Health Administration (VHA) as a result of injuries in military and combat operations.[9,17,34] Such injuries cause a substantial number of deaths and can lead to temporary or permanent disability. Despite their clinical relevance, many injuries are still unreported and this matter is further complicated by the limited sensitivity and reliability of current on-field clinical tests.[23,31] With the increased attention on recognition of neuromusculoskeletal injuries, there is a strong need for new assessment tools to help evaluate these injuries onsite, in non-clinical environments, more effectively and in a timely manner.

In on-field situations, balance is a commonly used indicator of mild traumatic brain injury (mTBI), concussion, and musculoskeletal injury.[14,16,32] To measure balance, a number of standardized screening tools are

Address correspondence to Alessandro Napoli, Department of Electrical & Computer Engineering, Temple University, Philadelphia, PA 19122, USA. Electronic mail: a.napoli@temple.edu

becoming prevalent in sideline/on-field balance assessment, replacing routine physical and clinical exams.[6,15,29] These on-field evaluations aim to provide a relatively brief assessment for determining whether a potentially injured service member or athlete is suitable to return to duty

The most commonly used clinical balance assessment tool following concussion is the Balance Error Scoring System (BESS).[16,28] The BESS test measures static postural stability and it is typically administered by trained medical personnel who must observe and count on a 0–10 scale, specific behaviors corresponding to deficits in postural control while simultaneously spotting the subject to prevent falls. The subject under test is required to maintain balance with eyes closed and hands on hips in three stance conditions: double-leg, single-leg and tandem stance. Each stance is performed on two surface types, hard ground (DS, SS, TS) and on a foam pad (DF, SF, TS). The standardized BESS defines the subject's balance errors, which must be counted:

- Moving the hands off the hips.
- Opening the eyes.
- Step, stumble or fall.
- Abduction or flexion of the hip beyond 30°.
- Lifting the forefoot or heel off the testing surface.
- Remaining out of the proper testing position for longer than 5 s.

Although fast and inexpensive, the BESS test presents a series of limitations that are intrinsically related to its subjective and manual scoring method. The BESS has been reported to have modest and widely ranging test sensitivity due to scoring inaccuracies and observer bias.[11,13] For instance, in Ref. 11 it has been reported that the inter-rater and intra-rater minimum detectible change for the total BESS score were respectively 9.4 and 7.3 points. These changes are in the same range as BESS score differences between baseline and testing in concussed subjects. It has been found that the average BESS score after concussion is 17 errors (range 15–19 errors), compared with ten errors at baseline (range 8.4–12.7 errors).[5] Further BESS limitations are the need for properly trained medical personnel to administer the test and its susceptibility to fatigue and practice effects.[5,18]

Given these limitations, numerous research efforts have aimed at improving balance evaluation for the BESS test. Such efforts can be divided into two main groups: (1) modifying and optimizing the human-based version of the test to make it more sensitive and to reduce the effects of fatigue and practice[3,18,19,33]; (2) Instrumenting the BESS, using either inertial measurement units[2,20,21] or force platforms,[1,8] to make the test more reliable and accurate, and reduce variability due to human bias.

MATERIALS AND METHODS

In order to overcome the BESS test limitations, we developed the automated assessment of postural stability (AAPS) system to quantify balance control in military field conditions. The AAPS' objective is to evaluate postural deficits due to concussion and musculoskeletal injuries commonly seen in active duty military personnel for return-to-duty assessment. The AAPS system utilizes an inexpensive Microsoft Kinect v2.0 motion capture sensor[5,18] and a custom designed software suite for Microsoft Windows to objectively track kinematics during Balance Error Scoring System (BESS) testing.

The AAPS system has been developed in the C# programming language in the Microsoft Visual Studio 2015 programming environment and the .NET 4.5 framework. The system requires minimal set-up time and no dedicated calibration time. Furthermore, it includes a comprehensive graphical user interface (GUI) to guide the operator and the subjects through the BESS test. The GUI also provides user controls, data management features, a real-time display of the detected body and intuitive visual feedback on the AAPS tracking capabilities. The system is user-friendly and its use only requires minimal training and experience. These characteristics facilitate the AAPS' integration and deployment in military practices. In addition to collecting information regarding the subject's joint positions and eye status (open/closed), the system provides real-time visual feedback to the operator. These characteristics help non-medical operators to properly position the subject in the field of view and guarantee that the necessary parameters are tracked correctly before starting the test.[27] The eye status tracking feature has been implemented using a face tracking library developed by Microsoft and available *via* the Kinect SDK 2.0. The face tracking combines HD color and infrared video streams to detect the location and status of the subject's eyes.[25]

This paper focuses on the balance error detection algorithm that has been implemented in the AAPS system to evaluate postural stability and provide a reliable and automated BESS score starting from raw Kinect sensor data. The algorithm has been designed to track balance errors as they are defined in the BESS standard.

This research was approved by the Temple University Institutional Review Board. All subjects provided written, informed consent prior to participating. A total of 15 healthy subjects (7 male, 8 female) each performed two complete BESS tests. Every trial was simultaneously tracked by a Kinect 2.0 sensor and a professional-grade motion capture system (Qualisys AB, Gothenburg, Sweden). High definition (HD) videos with BESS trials were scored off-line by three experienced observers. Further details regarding the experimental setup can be found in Ref. 26.

The AAPS system can be divided into three operating blocks: the Kinect sensor that detects human bodies using the depth data stream, the Microsoft proprietary skeletal tracking algorithm that converts the 3D camera images into 3D body joint coordinates, and the AAPS software that processes these coordinates and records balance errors. A top level block diagram of the system is shown in Fig. 1.

The Microsoft Kinect™ sensor is a low-cost, portable, and marker-less motion tracking system developed for video game applications. Despite being a mass-produced commodity, this 3D depth camera combined with the Microsoft proprietary skeletal tracking algorithm has the potential to be used as an alternative to laboratory-grade motion tracking systems.

Over the past few years, numerous performance comparisons between the Kinect and professional motion tracking systems have been carried out.[7,10,22,24,30,35] It is not surprising that such an inexpensive off-the-shelf commodity item cannot reach the levels of performance of professional-grade systems. Consequently, some important signal processing challenges must be overcome when using Kinect in applications such as the AAPS. For instance, the Kinect sensor frequently shows inaccuracies and oscillations when tracking body joints in both static and dynamic conditions. Such inaccuracies are affected by various parameters such as room conditions and geometry, the subject's body type, their distance from the sensor, and/or their clothing. Other types of tracking error can be due to quantization noise or missing information in the sensor data stream. Furthermore, the Kinect skeletal tracking presents two types of inaccuracy due to: (1) relatively small levels of white noise caused by detection imprecisions; and (2) temporary spikes in noise levels caused by joint tracking inaccuracies on a frame-by-frame basis. In real-world applications, as a result of these challenges, skeletal tracking can be carried out with precision levels in the range of a few centimeters.[12,26] Thus, without a strict rigid body model to be superimposed onto the Kinect raw data, it is not possible to compensate completely for the sensor-related errors.[4] This work demonstrates how dedicated signal processing techniques can mitigate these errors.

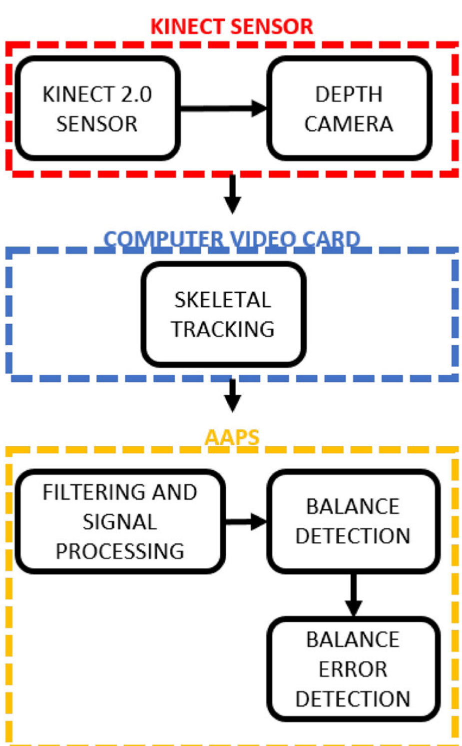


FIGURE 1. AAPS top level block diagram.

The first task of the AAPS software is to visualize and store the Kinect sensor raw data output, which is composed of HD video, infra-red video, three-dimensional depth data, joint position and orientation. Subsequently, as shown in Fig. 2, the AAPS extracts human body joint coordinates and locates the floor plane in real-time. The floor plane is used to identify the position and tilt of the sensor with respect to the subject. The joint coordinates are multiplied by a rotation matrix to compensate for sensor tilt and positioning. Next, the data frame rate is set to a constant value of 30 frames per second using linear interpolation. This is necessary because the Kinect provides data at a variable frame rate that depends on the instantaneous operating conditions of the acquisition computer (hardware/software) and data collection environment conditions such as lighting, room geometry, type and number of objects in the sensor field of view. To further account for the potentially large variability in the Kinect sensor frame rate (5–30 fps), the AAPS software was designed to perform real-time frame rate checks. If during a trial, the instantaneous frame rate drops below a certain value (10 fps in this application), an error message is displayed and the user is notified that the acquisition needs to be repeated. This is a fundamental feature in an automated system to guarantee acceptable performance levels in any condition. Based on our data collection sessions with the AAPS system, the ideal value of 30 fps tends to drop to 15 fps a few times per minute, while lower values are less frequent and usually occur once every 50 trials.

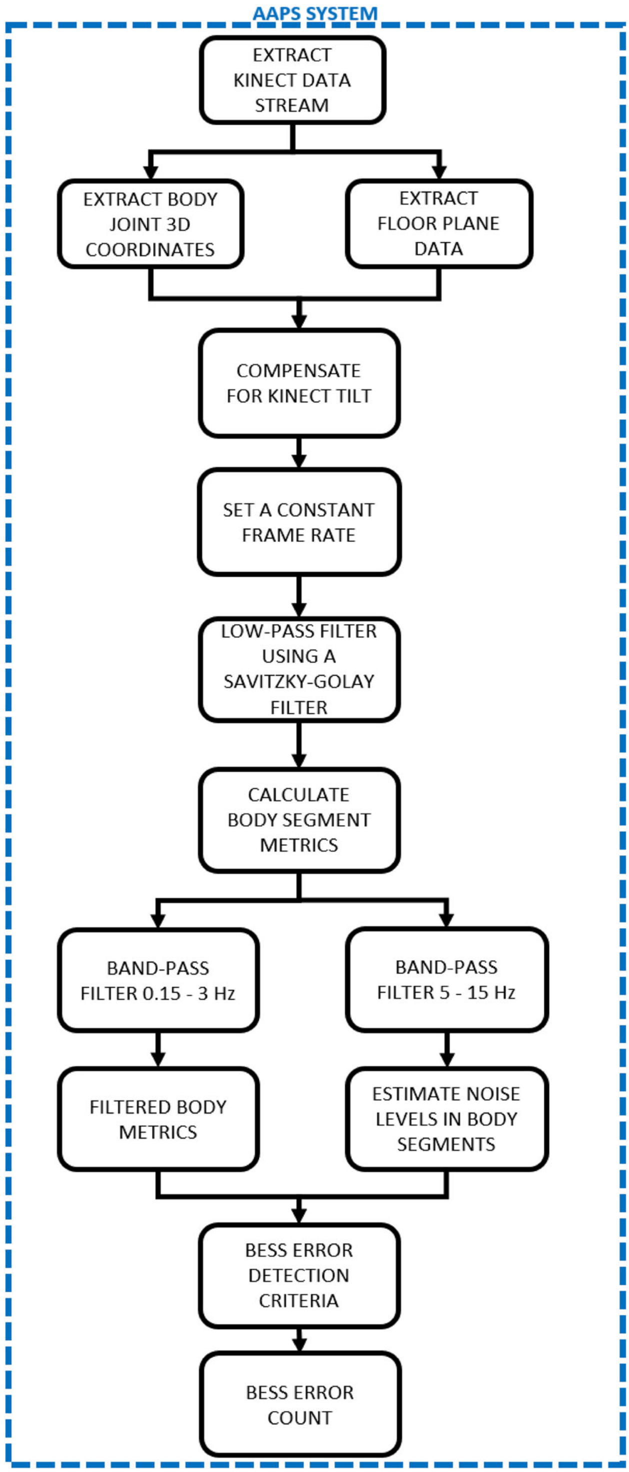


FIGURE 2. Detailed block diagram of the AAPS error detection algorithm.

Next the extracted body joint 3D coordinates are filtered using a Savitzky–Golay filter. This is a smoothing filter with minimal signal distortion that operates by fitting low-order polynomial approximations to consecutive signal time windows using a least-squares approach. A filter with a third order polynomial approximation and a time window duration of 0.166 s was used. At a constant sampling frequency of 30 Hz, such a window length corresponds to selecting five data points for each step of the least-squares approximation. As discussed above, filtering Kinect data with a smoothing filter is necessary to attenuate the effects of the Kinect inaccuracy and variability in estimating the joint positions of a tracked human body, even when subjects stand perfectly still in the sensor field of view. With the signal adequately smoothed, body metrics are calculated on a frame-by frame basis. The metrics that have been used in the AAPS algorithm to detect balance errors during BESS trials are listed in Table 1.

In order to detect errors in the subject's pose during balance trials, the algorithm uses a 1 s calibration window to estimate the reference subject's stance and the current levels of noise in the Microsoft skeletal tracking algorithm. The calibration is necessary to assess data variability due to changes in both subject-specific poses and sensor-specific body estimations. Subsequently, the metrics are bandpass filtered (using a second order Butterworth filter) between 0.15 and 3 Hz to emphasize signal components that are related to subject motion and to minimize other sources of variability (noise).

Additionally, sensor tracking inaccuracy is estimated by measuring the standard deviation of the noise in the calibration window. Specifically, the raw metrics are band-pass filtered with a second order band-pass Butterworth filter with passband set to 5–15 Hz. This frequency range was selected to emphasize the signal components that are mainly due to measurement noise.

During the 20-s long BESS trials, the estimated calibration stance and the current subject's position are continuously compared. The comparison is carried out using a threshold that is set using the estimated standard deviation of the noise and the mean of the metric obtained during calibration. Balance errors are flagged each time the metrics cross such a threshold. Specifically, in the ith frame, a balance error E_i is detected if the absolute difference between the calibration metric M_{cal} and the current metric $\hat{M}_i$ exceeds the threshold, set to ϵ times the estimated standard deviation σ_{cal} of the noise. The list of the kinematic metrics (M) that have been used to calculate the respective BESS errors (E_i) is presented in Table 1.

Mathematically, three categories of balance errors are detected:

(1) *Unilateral single threshold errors* estimated from low-noise and unilateral metrics.

$$E_i = \left(\left|\hat{M}_{n_i} - M_{n_{\text{cal}}}\right| > \epsilon_n \times \sigma_{n_{\text{cal}}}\right),$$

where the subscripts i, n and cal indicate respectively the frame number, the type of metric and the calibration window.

(2) *Bilateral errors* estimated from low-noise bilateral metrics. An error is detected if the threshold is crossed on either side of the body.

$$E_i = \begin{array}{c}\left(\left|\hat{M}_{n_{\text{left}_i}} - M_{n_{\text{left}_{\text{cal}}}}\right| > \epsilon_n \times \sigma_{n_{\text{left}_{\text{cal}}}}\right)\\ \text{OR}\\ \left(\left|\hat{M}_{n_{\text{right}_i}} - M_{n_{\text{right}_{\text{cal}}}}\right| > \epsilon_n \times \sigma_{n_{\text{right}_{\text{cal}}}}\right)\end{array}$$

where the subscripts left and right indicate from which side of the body the metrics were derived.

(3) *Double threshold errors* to improve detection performance, errors, that are estimated using low-accuracy metrics, are detected using two correlated metrics and corresponding thresholds. An error is detected only if both metrics cross the threshold.

$$E_i = \left(\left|\hat{M}_{n_i} - M_{n_{\text{cal}}}\right| > \epsilon_n \times \sigma_{n_{\text{cal}}}\right)$$
$$\text{and}\quad \left(\left|\hat{M}_{m_i} - \hat{M}_{m_{\text{cal}}}\right| > \epsilon_m \times \sigma_{m_{\text{cal}}}\right)$$

where the subscripts n and m indicate different metrics.

The above error types can be combined for improved balance detection precision. The different error types detected on a frame-by-frame basis are then converted into BESS scores, namely the total error count per trial, with two important caveats: (1) at most one error type can be detected within a pre-defined

TABLE 1. Calculated metrics extracted from Kinect raw data that are tracked during BESS tests.

Joints of interest	Metric (M)	Detected balance error (E_i)
Left hand—left hip	3D distance	Hands off hips
Right hand—right hip	3D distance	Hands off hips
Left elbow—left hip	3D distance	Hands off hips
Right elbow—right hip	3D distance	Hands off hips
Left knee—right knee	3D distance	Foot movement
Left hip—left ankle	3D distance (single-leg stance)	Hip flexion
Right hip—right ankle	3D distance (single-leg stance)	Hip flexion
Ankles	3D position (tandem stance & single-leg stance)	Foot movement
Frontal plane spine angle	Angle	Spine frontal motion
Sagittal plane spine angle	Angle	Spine sagittal motion

time window (set to 2 s); (2) a BESS error is recorded only if the infraction remains above the threshold for a pre-defined time duration (set to 110 ms). A detailed block diagram of the AAPS algorithm is shown in Fig. 2.

In order to validate the results of the error detection algorithm, we simultaneously collected data using a Kinect sensor and a 12-Camera Qualisys system. Qualisys data have been post-processed using Opensim with a modified plug-in-gait model. After running inverse kinematics on the trajectory data, three-dimensional body joint positions were derived. The Kinect and Qualisys derived joint coordinate time series were time-synchronized using a large movement performed at the beginning of each trial and then fed into the BESS error detection algorithm as described above. Finally, scores obtained from the two systems were compared against scores from three human experts reviewing video footage of the BESS tests.

RESULTS

The AAPS algorithm was tested, using data derived from both Qualisys and Kinect systems, on 15 healthy subjects, each performing the BESS test twice. These subjects' balance was also evaluated by three expert observers using the gold standard BESS method. In the algorithm performance analysis, the average human scores have been chosen as ground truth (Reference) for the correct error count.

Figure 3 shows the differences in the scores obtained using the different evaluation techniques: AAPS vs. Reference, Qualisys vs. Reference, AAPS vs. Qualisys, Human 1 vs. Reference, Human 2 vs. Reference, and Human 3 vs. Reference. The comparison of AAPS vs. Qualisys was carried out to investigate potential differences in performance due to the two different optical acquisition systems. Variations in scores have been quantified by calculating the signed average difference between each technique and the reference. Differences can range between -10 and 10 points, where low error levels are indicated by values close to zero. Standard deviations are presented as error bars.

Table 2 reports the overall level of agreement for the different groups, where values close to 100% (high agreement) correspond to differences in BESS scores close to zero. The values in the table are calculated by taking the percentage complement of the normalized absolute average differences in the scores. The absolute differences were normalized using the BESS full scale (10 points per trial).

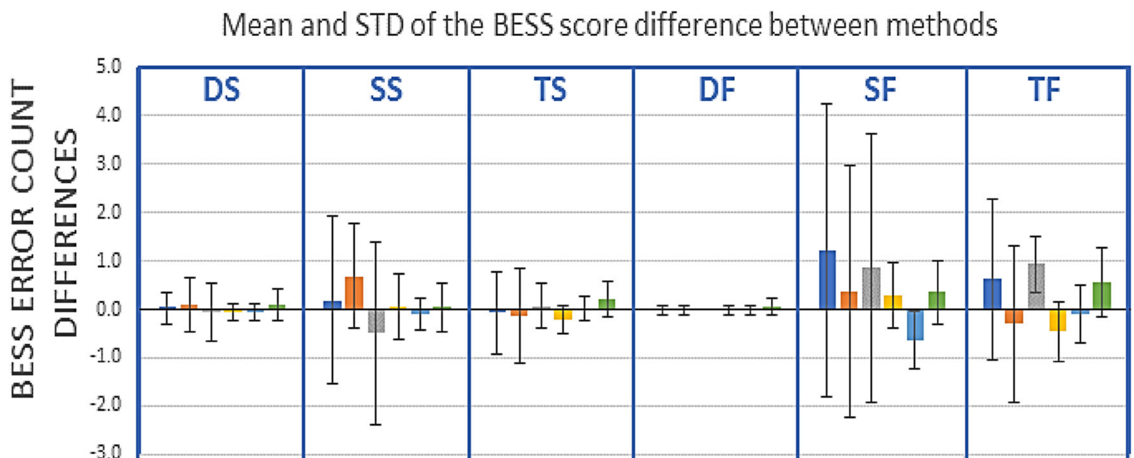
To evaluate the statistical significance of the observed score variations a multiple comparison one-way ANOVA test was implemented (α set to 0.05). The results are shown in Fig. 4, where the means (filled circles) and 95% confidence intervals (horizontal lines) of condition-based balance scores are presented. The gray vertical dotted lines represent the 95% confidence intervals with respect to the Reference group. No statistically significant differences were found between any of the balance scoring methods and the Reference (average human scores, in blue). The multiple comparison ANOVA results emphasize that although differences in the scores are non-significant, the Kinect-based AAPS reaches its lowest performance in the single-leg on foam condition, as also highlighted by the lowest agreement levels reported in Table 2.

Although single-leg on foam was identified as the condition with lowest agreement levels between observers, there was no significant difference in performance between the Kinect-based and the Qualisys-based AAPS. This finding suggests that the AAPS software algorithm provides satisfactory performance levels using raw data from both motion capture systems; BESS error detection performance is not significantly affected by the acquisition hardware.

DISCUSSION

As shown in the RESULTS section, the ANOVA analysis did not reveal any significant difference in the scores. It is worth noting that the lowest AAPS performance levels are detected in single-leg and tandem stances on foam. In such conditions, despite the Qualisys-based AAPS system performing more closely to humans than the Kinect-based one, statistical analysis shows no significant difference in performance. This result demonstrates that the AAPS, built around an inexpensive, general-purpose 3D single-camera sensor, is viable for use in on-field applications.

The lowest agreements between both the AAPS systems and human observers are seen in the single-leg and tandem stances on foam condition. We hypothesize that lower agreement levels might be due to the higher levels of subjective evaluation that this condition requires to detect BESS errors. Specifically, we identified three main factors. First, the presence of the foam complicates balance evaluation, because the foot on which the subjects stand is partially obscured by the foam. Secondly, this condition is arguably the most challenging, and consequently more motion is expected. This results in multiple errors and subjects having more difficulty to find and maintain their balance when trying to go back into the right position. In these cases, we found that human observers tend to use their "judgment" to count errors rather than strictly relying upon the BESS rules for balance error count. Finally, in single-leg on foam conditions, the auto-



	DS	SS	TS	DF	SF	TF
AAPS vs. Ref	0.0	0.2	-0.1	0.0	1.2	0.6
Qual vs. Ref	0.1	0.7	-0.1	0.0	0.4	-0.3
AAPS vs. Qual	-0.1	-0.5	0.1	0.0	0.9	0.9
H1 vs. Ref	0.0	0.0	-0.2	0.0	0.3	-0.5
H2 vs. Ref	0.0	-0.1	0.0	0.0	-0.6	-0.1
H3 vs. Ref	0.1	0.0	0.2	0.0	0.4	0.5

FIGURE 3. Means and standard deviations of the score differences calculated for each balance scoring method and grouped by stance condition. Bottom: mean error values for each group and condition. The tested stance conditions are: double leg (DS), single leg (SS) and tandem stance (TS) on firm ground; double leg (DF), single leg (SF) and tandem stance (TF) on foam pad. The blue, orange, grey, yellow, light blue and green bars represent different balance evaluations derived respectively for AAPS vs. Reference, Qualisys vs. Reference, AAPS vs. Qualisys, Human 1 vs. Reference, Human 2 vs. Reference, and Human 3 vs. Reference.

TABLE 2. Average differences expressed as percentage of agreement between different balance evaluation systems in detecting BESS scores, grouped by condition.

Percentage of agreement						
Condition	AAPS vs. Ref	Qual vs. REF	AAPS vs. Qual	H1 vs. Ref	H2 Vs. Ref	H3 vs. Ref
DS	99.8	99.0	99.3	99.5	99.5	99.0
SS	98.1	93.1	95.0	99.5	99.0	99.5
TS	99.3	98.6	99.3	97.9	100.0	97.9
DF	99.8	99.8	100.0	99.8	99.8	99.5
SF	87.9	96.4	91.4	97.1	93.6	96.4
TF	93.8	96.9	90.7	95.5	99.0	94.5

matic system seems to be operating at the limits of agreement between humans and AAPS systems because of the low sensitivity of the BESS test. This limitation has been reported in previous studies in which the modest sensitivity of the BESS is explained by the large variance in performance during the stances on foam. Over 53% of the variance in errors can be attributed to the single-leg and tandem conditions on foam.[21]

The BESS only focuses on static postural control tasks and lacks assessment of more dynamic postural tasks. Thus, the choice of filtering the kinematic metrics between 0.15 and 3 Hz to emphasize relevant data was deemed appropriate. The Kinect, and consequently the AAPS capabilities will be tested at their operational limit when introducing dynamic testing with the aim of capturing ''faster'' human movements. In such conditions, although the motion of large human body segments rarely exceeds a few Hertz, the filter high cut-off frequency needs to be increased to avoid signal's distortion and artifacts. However, based on our preliminary data during dynamic trials, the AAPS seems to perform at acceptable levels when compared to the Qualisys lab-grade performance.

Testing only for static stability may not capture other important domains of balance, including dy-

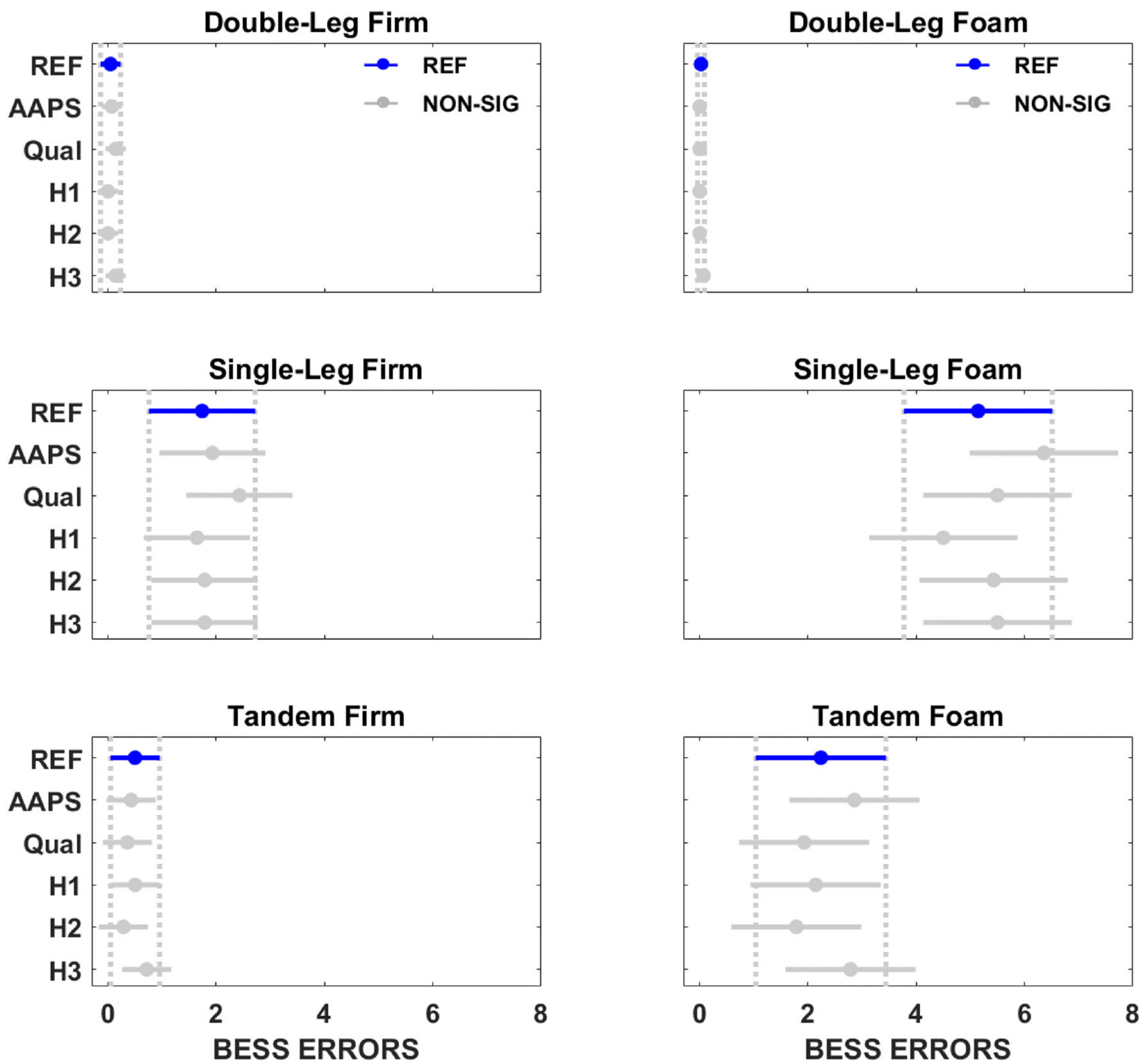


FIGURE 4. Results of a multiple comparison ANOVA test on the BESS scores. BESS errors derived using the AAPS, the Qualisys and three different human observers are compared to the average human scores, used as reference. Reference groups are in blue; vertical dotted lines are 95% confidence intervals for the Reference group. None of the differences with respect to Reference are statistically significant.

namic or multi-task postural control aspects.[13] It is worth noting that the AAPS capability of detecting balance deficits had to be reduced to a single error count number per trial for the purpose of the comparison presented here.

These limitations derive from human administration of such testing protocols, wherein some information (e.g., error type, time, and magnitude) must be sacrificed in order to accommodate the capacity of a human observer. We hypothesize that an improved automated balance test, in which dynamic conditions and more reliable proxy kinematic variables are used, can be readily implemented by exploiting the existing capabilities of the AAPS system. The use of such a system to detect, track and quantify balance deficits in the field will provide the opportunity to go beyond traditional balance testing protocols that only rely on human visual observations reported with manual annotations. This will facilitate more informed and data-driven clinical decision making in non-clinical settings.

Despite some level of disagreement between human and AAPS-generated scores, the use of an automated system yields important advantages over currently available human-based alternatives. A computer scoring system is by definition deterministic, meaning that it eliminates variability during repeated evaluations, the same criterion does not apply to human scoring. Moreover, the AAPS can record specific error types with extremely high temporal resolution, it can detect multiple error types on a frame-by-frame basis and record their time course progression. These features are not achievable by humans, who cannot keep track of all those variables with such a high time resolution. Together, these results suggest that computerized BESS calculation may provide more accurate and consistent measures of balance than those derived from human experts.

Our results show that the AAPS error detection algorithm presented here can accurately and precisely detect balance deficits with performance levels that are comparable to those of experienced medical personnel. Specifically, our results show agreement levels between the AAPS algorithm and the human average BESS scores ranging between 87.9% (single-leg on foam) and 99.8% (double-leg on firm ground). In addition, statistically significant differences were not detected by an ANOVA test with significance level set to 0.05. Moreover, significant performance deficits were not detected when the less expensive, portable and markerless AAPS was compared to a lab-grade system, with agreement levels between the two different motion capture systems ranging between 90.7% (tandem on foam) and 100% (double-leg on foam). These results underscore the value of using the Kinect-based AAPS, which can be quickly deployed in the field and/or in outdoor settings with minimal set-up time.

In future work, we plan on expanding the AAPS with new features, such as introducing criteria to account for balance error characteristics and fine-grained evaluation of dynamic and static postural control strategies using kinematic variables rather than trying to capture complex motion performance with an arbitrary summary scale. Such a system will also implement functional dynamic protocols that can be customized to a specific subject and application. These new dynamic posture screening tools combined with the ability to derive real-time meaningful postural metrics will help us develop innovative automated tools for more effective and comprehensive on-field postural strategy assessment. Furthermore, the AAPS capabilities will be tested in clinical populations, such as individuals suffering from low-extremity injuries and concussion.

ACKNOWLEDGMENTS

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs through the Neurosensory and Rehabilitation Research Award Program under Award No. W81XWH-15-1-0045. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense. The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702-5014 is the awarding and administering acquisition office.

REFERENCES

[1]Alsalaheen, B. A., J. Haines, A. Yorke, K. Stockdale, and S. P. Broglio. Reliability and concurrent validity of instrumented balance error scoring system using a portable force plate system. *Phys. Sports Med.* 43(3):221–226, 2015.

[2]Amick, R. Z., S. D. Jansen, A. Chaparro, N. A. Hakansson, J. A. Patterson, and M. J. Jorgensen. Comparison of the SWAY balance mobile application to the abbreviated balance error scoring system. *Athl. Train. Sport Health Care* 7(3):89–96, 2015.

[3]Azad, A. M., S. Al Juma, J. A. Bhatti, and J. S. Delaney. Modified balance error scoring system (M-BESS) test scores in athletes wearing protective equipment and cleats. *BMJ Open Sport Exerc. Med.* 2(1):e000117, 2016.

[4]Azimi, M. Skeletal Joint Smoothing White Paper. In: Microsoft Developer Network. https://msdn.microsoft.com/en-us/library/jj131429.aspx. [Accessed: 01-Jan-2017].

[5]Bell, D. R., K. M. Guskiewicz, M. A. Clark, and D. A. Padua. Systematic review of the balance error scoring system. *Sport Health A* 3(3):287–295, 2011.

[6]Broglio, S. P., S. N. Macciocchi, and M. S. Ferrara. Sensitivity of the concussion assessment battery. *Neurosurgery* 60(6):1050–1057, 2007.

[7]Clark, R. A., K. J. Bower, B. F. Mentiplay, K. Paterson, and Y. H. Pua. Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables. *J. Biomech.* 46(15):2722–2725, 2013.

[8]Clark, R. A., A. L. Bryant, Y. Pua, P. McCrory, K. Bennell, and M. Hunt. Validity and reliability of the Nintendo Wii balance board for assessment of standing balance. *Gait Posture* 31(3):307–310, 2010.

[9]Clinical Practice Guideline. Management of Concussion/mild Traumatic Brain Injury. The Management of Concussion/mTBI Working Group. VA/DoD Evidence Based Practice. April, 2009.

[10]Dolatabadi, E., B. Taati, and A. Mihailidis. Concurrent validity of the Microsoft Kinect for Windows v2 for measuring spatiotemporal gait parameters. *Med. Eng. Phys.* 38:952–958, 2016.

[11]Finnoff, J. T., V. J. Peterson, J. H. Hollman, and J. Smith. Intrarater and interrater reliability of the balance error scoring system (BESS). *PM R* 1(1):50–54, 2009.

[12]Galna, B., G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester. Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. *Gait Posture* 39(4):1062–1068, 2014.

[13]Giza, C. C., J. S. Kutcher, S. Ashwal, J. Barth, T. S. D. Getchius, G. A. Gioia, G. S. Gronseth, K. Guskiewicz, S. Mandel, G. Manley, D. B. McKeag, D. J. Thurman, and R. Zafonte. Summary of evidence-based guideline update: evaluation and management of concussion in sports. *Neurology* 80(24):2250–2257, 2013.

[14]Guskiewicz, K. M. Postural stability assessment following concussion: one piece of the puzzle. *Clin. J. Sport Med.* 11(3):182–189, 2001.

[15]Guskiewicz, K. M., B. L. Riemann, D. H. Perrin, and L. M. Nashner. Alternative approaches to the assessment of mild head injury in athletes. *Med. Sci. Sports Exerc.* 29(7 Suppl):S213–S221, 1997.

[16]Guskiewicz, K. M., S. E. Ross, and S. W. Marshall. Postural stability and neuropsychological deficits after concussion in collegiate athletes. *J. Athl. Train.* 36(3):263–273, 2001.

[17]Helmick, K. M., C. A. Spells, S. Z. Malik, C. A. Davies, D. W. Marion, and S. R. Hinds. Traumatic brain injury in the

US military: epidemiology and key clinical and research programs. *Brain Imaging Behav*. 9:358–366, 2015.
[18]Hunt, T. N., M. S. Ferrara, R. A. Bornstein, and T. A. Baumgartner. The reliability of the modified balance error scoring system. *Clin. J. Sport Med*. 19(6):471–475, 2009.
[19]Iverson, G. L., and M. S. Koehle. Normative data for the modified balance error scoring system in adults. *Brain Inj*. 27(5):596–599, 2013.
[20]King, L. A., F. B. Horak, M. Mancini, D. Pierce, K. C. Priest, J. Chesnutt, P. Sullivan, and J. C. Chapman. Instrumenting the balance error scoring system for use with patients reporting persistent balance problems after mild traumatic brain injury. *Arch. Phys. Med. Rehabil*. 95(2):353–359, 2014.
[21]King, L. A., M. Mancini, P. C. Fino, J. Chesnutt, C. W. Swanson, S. Markwardt, and J. C. Chapman. Sensor-based balance measures outperform modified balance error scoring system in identifying acute concussion. *Ann. Biomed. Eng*. 2017. doi:10.1007/s10439-017-1856-y.
[22]Lim, D., C. Kim, H. Jung, D. Jung, and K. Chun. Use of the microsoft kinect system to characterize balance ability during balance training. *Clin. Interv. Aging* 10:1077–1083, 2015.
[23]McCrory, P., W. Meeuwisse, M. Aubry, B. Cantu, J. Dvořák, R. Echemendia, L. Engebretsen, K. Johnston, J. Kutcher, M. Raftery, A. Sills, B. Benson, G. Davis, R. Ellenbogen, K. Guskiewicz, S. A. Herring, G. Iverson, B. Jordan, J. Kissick, M. McCrea, A. McIntosh, D. Maddocks, M. Makdissi, L. Purcell, M. Putukian, K. Schneider, C. Tator, and M. Turner. In: Consensus Statement on Concussion in Sport. The 4th International Conference on Concussion in Sport held in Zurich, November 2012, *Phys. Ther. Sport*, 14(2):250–258, 2013.
[24]Mentiplay, B. F., L. G. Perraton, K. J. Bower, Y. H. Pua, R. McGaw, S. Heywood, and R. A. Clark. Gait assessment using the Microsoft Xbox One Kinect: concurrent validity and inter-day reliability of spatiotemporal and kinematic variables. *J. Biomech*. 48(10):2166–2170, 2015.
[25]Microsoft. Face Tracking. In: Kinect for Windows SDK 2.0—Programming Guide. https://msdn.microsoft.com/en-us/library/dn782034.aspx. [Accessed: 01-Jan-2017].
[26]Napoli, A., S. Glass, C. Ward, C. Tucker, and I. Obeid. Performance analysis of a generalized motion capture system using microsoft kinect 2.0. *Biomed. Signal Process. Control* 38:265–280, 2017.
[27]Napoli, A., C. R. Ward, S. M. Glass, C. Tucker, and I. Obeid. In: Automated assessment of postural stability system. Proceeding of Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6090–6093, 2016.
[28]Onate, J. A., B. C. Beck, and B. L. van Lunen. On-field testing environment and balance error scoring system performance during preseason screening of healthy collegiate baseball players. *J. Athl. Train*. 42(4):446–451, 2007.
[29]Peterson, C. L., M. S. Ferrara, M. Mrazik, S. Piland, and R. Elliott. Evaluation of neuropsychological domain scores and postural stability following cerebral concussion in sports. *Clin. J. Sport Med*. 13(4):230–237, 2003.
[30]Pfister, A., A. M. West, S. Bronner, and J. A. Noah. Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. *J. Med. Eng. Technol*. 1902(August):1–7, 2014.
[31]Putukian, M. The acute symptoms of sport-related concussion: diagnosis and on-field management. *Clin. Sports Med*. 30(1):49–61, 2011.
[32]Riemann, B. L., K. M. Guskiewicz, and E. W. Shields. Relationship between clinical and forceplate measures of postural stability. *J. Sport Rehabil*. 8(2):71–82, 1999.
[33]Starling, A. J., D. F. Leong, J. M. Bogle, and B. B. Vargas. Variability of the modified balance error scoring system at baseline using objective and subjective balance measures. *Concussion* 1:1–8, 2015.
[34]Stroupe, K. T., B. M. Smith, T. P. Hogan, J. R. St, T. Andre, M. L. Pape, E. Steiner, Z. Huo Proescher, and C. T. Evans. Healthcare utilization and costs of veterans screened and assessed for traumatic brain injury. *J. Rehabil. Res. Dev*. 50(8):1047–1068, 2013.
[35]Yeung, L. F., K. C. Cheng, C. H. Fong, W. C. C. Lee, and K.-Y. Tong. Evaluation of the Microsoft Kinect as a clinical assessment tool of body sway. *Gait Posture* 40(4):532–538, 2014.

Performance Analysis of a Generalized Motion Capture System Using Microsoft Kinect 2.0

Alessandro Napoli, PhD[1]

Stephen Glass, PhD[2]

Christian Ward[1]

Carole Tucker, PhD[2]

Iyad Obeid, PhD[1]

[1]Department of Electrical & Computer Engineering, Temple University, Philadelphia, PA 19122

[2]Department of Physical Therapy, Temple University, Philadelphia, PA 19140

Abstract

This work presents a fine-grained analysis of the performance and limitations of the Microsoft Kinect sensor for tracking human movement in the context of biomechanical research and clinical applications. Earlier work in this field has focused on scalar summary measures or ad-hoc metrics with respect to specific movements that do not generalize well across clinical applications. In this work, the performance of the Microsoft Kinect is compared to motion tracking from a concurrently sampled professional grade Qualisys motion capture system. Subjects performed a range of clinically relevant tasks such as Sit-to-Stand and Timed Up-and-Go. Captured data included both three-dimensional joint center displacements and joint angles as recorded from both systems. Kinect performance was measured using cross correlation coefficients (CCR), root mean squared error (RMSE) relative to the Qualisys gold-standard and a new summary metric (SM) that combines both. Our results show that the Kinect-based system provides adequate performance when tracking joint center displacements in time, with overall CCR=0.78, RMSE=3.35cm and SM=1.21. On the contrary, lower accuracy was measured when tracking joint angles, with CCR=0.58, RMSE=24.59 degrees, and SM=3.76. Although performance differences for various movements and motion planes have been found, the results suggest that the Kinect is a viable tool for general biomechanical research, with specific limits on what levels of performance can be expected under various conditions.

1 Background

Capturing three-dimensional movement (or *kinematics*) is a central laboratory technique in the study of human movement. Kinematic studies have played an instrumental role in the study of joint pathology, mild traumatic brain injury, ergonomics, and athletic performance. Despite the importance of this technique, standard data acquisition methods are subject to considerable limitations. Stereophotogrammetry and electromagnetic motion tracking, for instance, require expensive, stationary equipment and time-consuming procedures for system calibration and post-processing of data. In contrast, low-cost depth sensing cameras (also known as time-of-flight or RGB-D cameras) are available off-the-shelf and may represent viable alternatives to more complex and expensive 3D camera setups. These cameras are capable of capturing RGB color images augmented with depth data at each pixel, thus providing 3D images. Such images can be used to track human motion in real-time. Among these cameras, the Microsoft Kinect™ 2.0 provides a low-cost, portable, user-friendly alternative which holds the potential to substantially increase the accessibility of kinematic data. The main advantage of using the Kinect sensor over the commercially available alternatives lies in the proprietary Microsoft algorithm that performs body and joint detection in real-time and that can be exploited using the Microsoft Software Development Kit (SDK) [1] available to .NET developers.

The performance of the Kinect™ 2.0 as a tool to evaluate kinematic variables, as compared to current standard methods, is a subject of great interest. Although several studies have been published in this area, the general trend has been to compare motion capture (MOCAP) systems based on scalar summary measures a selected metric. Examples from previous work include excursion range [2]–[5], mean or peak displacement [4], [6]–[11] or timing of discrete signal events [4], [5], [9], [11]–[15]. While such metrics are commonly studied in biomechanics, they do not adequately quantify the temporal structure of the signals under comparison, and are thus limited in terms of the generalizability of their results. To date, three laboratories have presented a more thorough treatment of Kinect 2.0 time series data in comparison to an existing standard. These studies still present certain limitations which restrict broader generalizability [16]–[19].

The first study [19] was based on a composite signal, namely, total body center of mass, which represents a weighted sum of body segments derived from the Kinect joint center time dynamics. As a weighted sum, this output suppresses the variability inherent to the underlying time series data. Another study [17] reports measures of signal agreement measured as Intraclass Correlation Coefficients (ICC) between Kinect 2.0 sensors and an OptoTrak system (Northern Digital Inc., Waterloo, Canada). This investigation analyzed the consistency between the systems in each dimension for nearly all the joint center data natively exported from the Kinect. While this approach offers an appropriate comparison of the two systems, the results are specific to gait, a primarily sagittal plane movement pattern, and therefore represent a relatively narrow range of the human motion repertoire. Finally, a pair of studies [16], [18], by a third group, quantifies Kinect signal error by its 3D L^2 norm distances from a ground truth signal as given by a 3D professional-grade MOCAP system. While they were able to classify individual data points as outliers vs. inliers based on error magnitude, information regarding the dimension and direction of signal offset is lost when using the 3D distance. As a result, this approach is not suitable for identifying systematic, direction-specific errors such as those noted by other investigators[6]. Additionally, their approach, which collapsed analyses across six tested movements, may obscure any relationships between the Kinect and ground truth signals that are specific to a given experimental condition or movement.

Considering these limitations and the expanded use of Kinect based systems in quantitative kinematic studies, a more thorough evaluation of Kinect 2.0 raw data performance as a MOCAP tool and its validation against a gold-standard 3D system is needed. The overarching goal of this research is the development of a system to collect reliable, valid kinematic data using low-cost sensors. The applications of such a technology are wide-reaching and may involve physical medicine clinics, athletics settings, and home entertainment, as well as other research domains in which kinematic data are not commonly acquired owing to prohibitive costs. The specific aim of this study was to identify limitations (and, ultimately, corrective measures) in Kinect 2.0 performance as an off-the-shelf technology for a flexible and multi-purpose MOCAP system. To that end, we have validated the Kinect against a professional three-dimensional motion capture system with 12 IR-cameras (Qualisys AB, Gothenburg, Sweden) over a range of dynamic movements and clinical tests that can be used as broad indicators of functional movement. In addition to this Kinect-vs-gold standard comparison, we present raw data from a second Kinect 2.0 sensor positioned alongside the first. These data provide an indication of reliability between Kinect 2.0 sensors. We acquired data from four healthy subjects and calculated results for point kinematics and joint angles, the latter of which are derived independently for the Kinect data using both quaternions and trigonometry applied to the joint positions.

2 Body Segment Orientation

The Microsoft Kinect 2.0 senses depth using an infrared camera sensor. A proprietary on-board algorithm locates bodies within the depth image and extracts parameters that describe the positions of up to six bodies in three-space in real time. For each tracked body, Kinect produces two data streams. The first is "joint location", which tracks the three-dimensional coordinates of 25 joints. The Kinect estimates three dimensional coordinates on a frame-by-frame basis using a probabilistic model that compares data from the depth image to a comprehensive database of human poses [20] [21]. These measurements are in meters and are measured relative to an origin that is represented by the sensor camera itself. The second stream is "body segment orientation," in which the orientation and rotation of each segment relative to its parent, (e.g. forearm relative to upper arm) can be represented numerically by a *quaternion*. These real-time data streams are both complicated by the fact that the sampling rate varies between 5 and 30 frames per second according to instantaneous demands on the computer's processor.

A quaternion is a 4-tuple that represents the orientation and rotation of an object in three dimensions relative to some parent coordinate axis. Specifically, quaternion u can be expressed as

$$u = u_0 + u_x i + u_y j + u_z k = M\cos(\alpha) + M\boldsymbol{u}\sin(\alpha) = Me^{\boldsymbol{u}\alpha}$$

If v is some other quaternion, then v can be rotated around unit quaternion u (eg. $M = 1$) by 2α radians using the following transform:$v_{rot} = uvu^*$. Although the Kinect produces a stream of "body segment orientations", these measurements must be numerically manipulated to yield clinically relevant kinematic data.

In some cases, this calculation is straightforward. For example, elbow angle can be calculated by simply calculating the angle between the quaternions of the upper arm and forearm as

$$\theta = \cos^{-1}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{|\boldsymbol{u}||\boldsymbol{v}|}\right)$$

In other cases, the transformation from quaternion orientations to clinical kinematic data requires projecting body segments into the three cardinal planes (mediolateral, vertical, and anteroposterior).

Kinect quaternion mathematics are complicated by two main factors. The first is that all Kinect quaternions are defined with respect to their "parent segment" quaternion, and the second is that quaternions do not describe anatomically significant angles. Specifically, each Kinect quaternion is defined so that its y-axis points to its "child segment" quaternion, while the z-axis is normal to both the y-axis and the body segment. The x-axis is normal to both the previous axes. Since the root joint is the lower spine, all relative orientations can be re-referenced to this initial orientation by consecutive parent/child multiplication along the quaternion body chain, using Hamilton products, as follows:

$$\begin{aligned}
q3_0 &= q1_0 * q2_0 - q1_z * q2_z - q1_y * q2_y - q1_x * q2_x \\
q3_x &= q1_0 * q2_x + q1_z * q2_y - q1_y * q2_z + q1_x * q2_0 \\
q3_y &= q1_0 * q2_y - q1_z * q2_x + q1_y * q2_0 + q1_x * q2_z \\
q3_z &= q1_0 * q2_z + q1_z * q2_0 + q1_y * q2_x - q1_x * q2_y
\end{aligned}$$

where q1 and q2 are the parent and child quaternions, respectively, and q3 is the quaternion that represents the orientation of the child segment.

The second step necessary for deriving meaningful joint angles is that segment orientations expressed using quaternions must be converted into Euler angles. Specifically, the position of a limb in three-space may be considered as the result of one or more rotations in each of the cardinal planes. The values of the rotation angles and the accuracy of the conversion relative to the original quaternion depends on the order of rotation as well as the joint in question and even the movement being performed [25]. The conversion can be performed using each one of 12 possible combinations of the three axes of rotation, also known as *rotation sequences*. In this work, we chose the rotation sequences for each movement and joint that are most commonly used in biomechanics [22]–[24].

In addition to computing joint angles from the Kinect's quaternion stream, they can also be derived directly from the three dimensional joint locations. Specifically, the location of two joints in 3D space defines a body segment orientation. Following standard practice [26], [27], the angle of each body segment is calculated relative to the normal of the floor, giving what is defined as an *absolute angle*. The absolute angles can then be differenced to compute joint angles (also called *relative angles*).

$$\Theta_{joint} = \Theta_{parent\ segment} - \theta_{child\ segment}$$

where each segment's absolute angles in the frontal and sagittal planes can be calculated respectively as:

$$\Theta_{\text{segment}_{frontal}} = \text{atan}\left(y_{distal} - y_{proximal}/x_{distal} - x_{proximal}\right)$$

$$\Theta_{\text{segment}_{sagittal}} = \text{atan}\left(z_{distal} - z_{proximal}/y_{distal} - y_{proximal}\right)$$

where x, y, and z are the coordinates of the joint centers that define a body segment. We compared joint angles computed using this method with those derived from the quaternion measurements to determine whether there were any systematic biases or errors across subjects or movements.

3 Methods

Four subjects (three males, mean age 23) performed a series of movements while being simultaneously tracked by two immediately adjacent Kinect 2.0s and a professional-grade motion capture system (Qualisys AB, Gothenburg, Sweden). The two Kinects were used to evaluate inter-unit accuracy, and motion tracking results from both systems were compared to the gold standard Qualisys system. The movements and the recording paradigm were specifically designed to facilitate the quantification of errors in tracking joint angles and limb locations relative to all three cardinal planes of the body axis: mediolateral, vertical, and anteroposterior.

Subjects wore tight-fitting shorts and an (optional) upper body garment that allowed for placement of reflective markers in accordance with the Qualisys MOCAP full-body plug-in-gait marker set [28]. This included 39 markers, placed on the head, arms, wrists, trunk, pelvis, legs and feet. All acquisitions were performed in a dedicated motion tracking laboratory which houses a 12-camera Oqus passive marker measurement system. Before each session, the Qualisys system was calibrated and then the two Kinect sensors were placed on tripods between 2m and 4m in front of the subject. The exact Kinect sensor locations were determined by following the manufacturer recommendations and by verifying that the subject was completely and optimally in the field-of-view, thus, optimizing the performance of the proprietary body tracking algorithm. The experimental setup is shown in Figure 1.

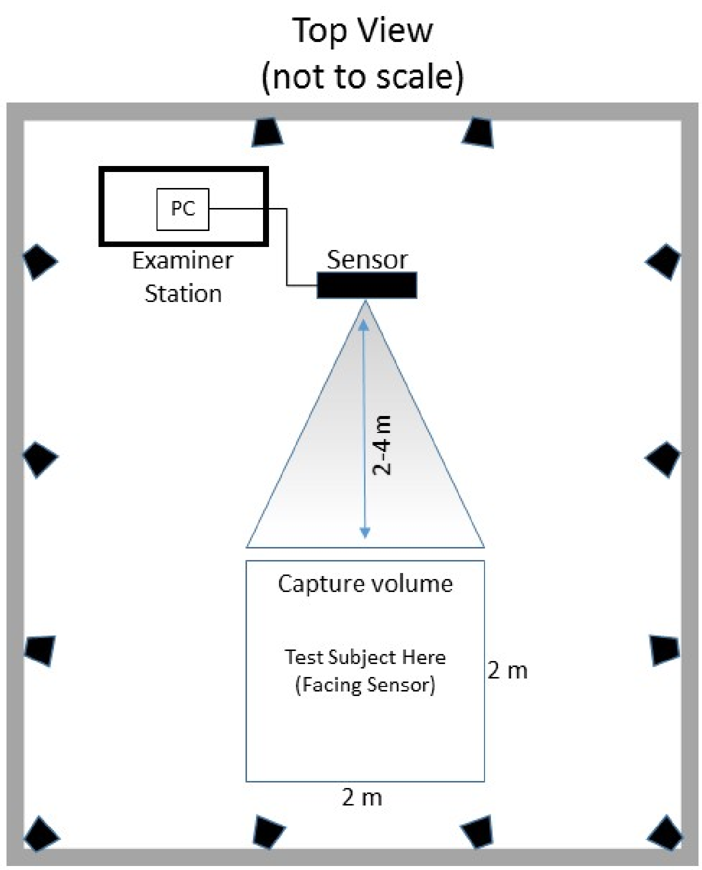


Figure 1: Experimental capture volume block diagram. Oqus cameras are represented by black boxes along the perimeter. The Kinect is labeled "Sensor."

The four subjects were asked to perform two trials each of a series of different moving postures. Each trial was preceded with a large movement such as a T-pose or an overhead reach to facilitate time-synchronization of the Qualisys and Kinect systems; these movements were not scored or otherwise included in the results. Three subjects performed a battery of *standard* clinical tests of dynamic posture, whereas the fourth subject performed the *stereotyped* postures in which movement was intentionally restricted to a single plane.

Table 1: Dynamic Postures

Standard	**Posture/Movement**	**Primary Plane**	**Notes**
	Sit-to-stand	Sagittal	• five repetitions
	Timed up-and-go	Sagittal / Transverse	• stand; walk to marked position ~3m away; walk back
	Alternating barbell lunges	Sagittal	• ten repetitions
	Overhead squats	Sagittal / Frontal	• five repetitions
	Marching in place	Sagittal	• 20 seconds
	Time to stabilization	Sagittal / Frontal	• Forward hop with unilateral landing
Stereotyped	**Posture/Movement**	**Primary Plane**	**Notes**
	Shoulder Flexion/Extension	Sagittal	• Right arm start extended at 0° • Right arm flex to 90° • Right arm flex to 180° • Right arm extend to 90° • Right arm extend to 0° • Repeat on left side
	Shoulder Ab/Adduction	Frontal	• Arms raised laterally (abducted) to “T” position • Arms raised laterally to overhead position • Arms lowered laterally (adducted) to “T” position • Arms lowered laterally to starting position • Repeat once
	Hip Ab/Adduction	Frontal	• Right leg raised laterally (hip abducted) to ~45° • Right leg lowered laterally (hip adducted) to 0° • Left leg raised laterally to ~45° • Left leg lowered laterally to 0° • Repeat once
	Combined Hip/Knee Flexion/Extension	Sagittal	• Right hip and knee flexed (e.g. “high knee” stepping) • Right hip and knee extended to starting position • Left hip and knee flexed • Left hip and knee extended to starting position • Repeat once
	Combined Arm Ab/Adduction & Elbow Flexion/Extension	Frontal / Sagittal	• Arms raised laterally to “T” position • Elbows flexed 90° to “goal post” position • Elbows flexed maximally • Elbows extended to “goal post” position • Elbows extended to achieve “T” position
	Trunk Leans	Sagittal / Frontal	• Lean left • Return to starting position • Lean right • Return to starting position • Lean backward • Return to starting position • Lean forward • Return to starting position

Body movement was tracked with the two systems simultaneously. The benchmark Qualisys was controlled via proprietary data acquisition software ("Qualisys Track Manager"). In order to convert the marker locations to angular kinematics, the trajectories were mapped offline onto a subject-specific rigid body model using the freely available OpenSim software tool. [29] In contrast, the two Kinects tracked body motions in 3D using single-camera sensors, Microsoft's proprietary body tracking algorithm [12], and a custom-designed software tool that converted the Kinect raw data into comma-separated-variable files for offline analysis. [30] These CSV files stored three-dimensional joint center trajectories, the three-dimensional orientations of 25 joints, and the Kinect-estimated floor plane quaternion. The data acquisition and processing pathway is summarized in Figure 2.

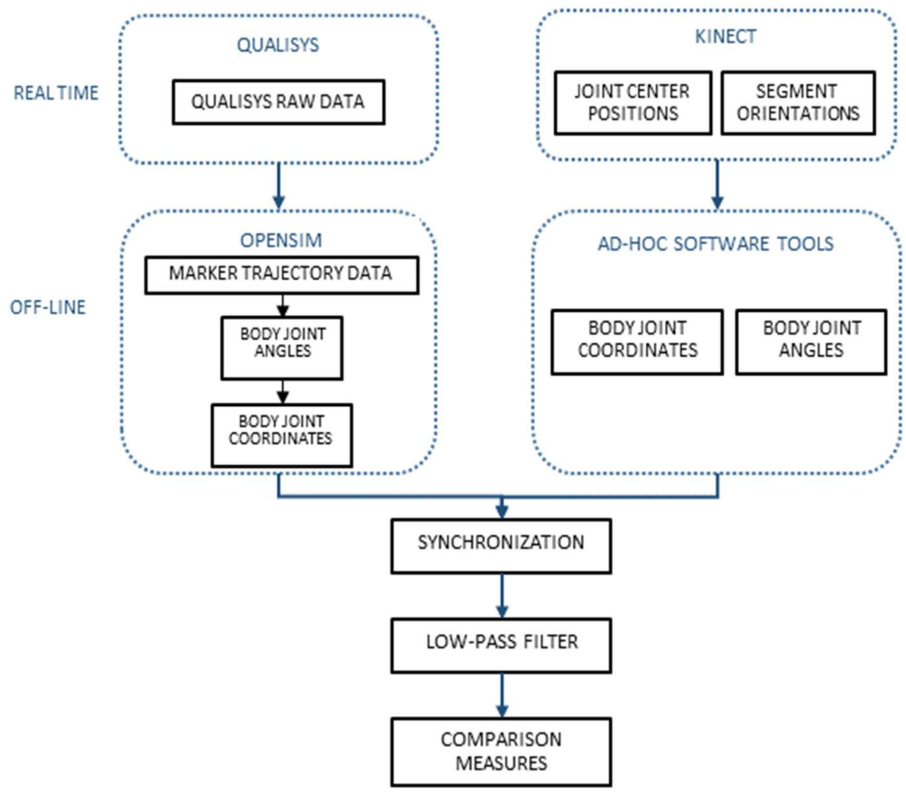


Figure 2: Signal Processing Block Diagram

Note that, while the Qualisys data were acquired at a constant frame rate of 120 fps, the Kinect data were acquired at a variable frame rate that depended heavily on the available computational and memory resources in the acquisition computer during run-time. Although the Kinect nominally collected data at 30 fps, instantaneous drops to 10-15 fps were not uncommon. For the purpose of data analysis, all Kinect data was upsampled to 120fps to match the Qualisys data.

Hardware limitations made it impossible to time-align the data streams using an electronic trigger during acquisition. Instead, data from the systems were temporally aligned by first de-trending them and then by cross-correlating them to assess the time lag that maximized their similarity [13]. To facilitate time synchronization, each pose was preceded by a large *alignment movement*, either an arms-open T-pose or an eccentric or concentric overhead press. The presence of a large alignment movement before each posture allowed the cross-correlation method to work robustly; temporal alignments were manually reviewed and found to be accurate to within a single frame. The alignment movement also served the dual purpose of initiating Kinect body tracking, which requires movement for optimal body detection.

After synchronization, data from both systems were low-pass filtered to reduce noise and acquisition artifacts. Since human movements are rarely faster than a few hundred milliseconds (eye blinks are typically 100-150ms) [31], motion capture systems do not need frequencies higher than a few Hertz. In this work, as in other manuscripts of this nature, all data was low-pass filtered using a 15th order Butterworth low-pass filter with a 3 dB cut-off frequency of 6.3 Hz [2], [32]. Finally, in order to calculate joint displacement and compare Qualisys and Kinect data, movement from each joint was zeroed relative to its initial position.

4 Results

In order to validate the Kinect-based MOCAP system, we compared its performance to a gold standard 12-camera 3D Qualisys system, that was used to concurrently evaluate subjects' movements. Subsequently, the similarity between the two sets of data was assessed by measuring both the cross-correlation values and the average absolute errors between corresponding time series, joint by joint and motion by motion, to elucidate any systematic biases that appear in some dimensions but not others. Since no statistical difference (Student's t-test, $p>0.05$) was found between measurements from the two side-by-side Kinect sensors, we report only their average measurements. Three separate error metrics were used to describe the performance of the Kinect system relative to the benchmark Qualisys system. First, we calculated cross-correlation coefficients, which effectively describe how much information one signal can yield about another, but that is generally blind to errors of constant or near-constant offset or bias. Secondly, we calculated the root-mean-squared errors (RMSEs) between the various output signals. RMSE measures constant or near constant differences between signals but is blind to signal correlation. Finally, we propose a novel "summary measure" which seeks to combine the cross-correlation and RMSE errors in a manner that is relevant to the context of Kinect-based motion tracking.

When comparing data between the two adjacent Kinect sensors, small deviations were noted between the two. However, these differences were not statistically significant (Student's t-test, $p>.05$). For brevity, only the mean values between the two sensors are presented here.

Figure 3 shows representative joint center displacements of head, the middle of the spine, left hip, and right hip, as measured for a single subject during a sit-to-stand test; data from both Qualisys and Kinect are displayed. This figure illustrates similarities between the joint displacement data from the two motion capture systems. This figure also underscores the need for multiple similarity metrics, since it is possible to have a high cross correlation but also a high root mean squared error. The poorest cross-correlation scores were generally obtained with respect to those axes along which the observed motion was negligible (mediolateral in the case of sit-to-stand). Note also that, in general, tracking between the two systems was less robust in the anteroposterior plane (representing 'depth' away from the Kinect sensor), than in the vertical plane. A more detailed summary of these results is seen in Tables 2 and 3, and in the Appendix.

Figure 4 shows representative joint angles (lumber extension, and hip flexion – left and right) for the same sit-to-stand trial as depicted in Figure 3. In the case of joint angles, we compare two sets of data to the Qualisys gold standard (blue traces). The first is joint angles derived from the Kinect quaternion data stream (red traces), whereas the second is joint angles derived directly from the Kinect's 3D joint coordinates (orange traces). We compare the Kinect-derived joint angles to each other as well as to the Qualisys benchmark. In general, the two Kinect-derived measures correlated strongly against one another and against the benchmark, although offsets and scale factors tended to negatively impact the RMSEs. Again, a more detailed summary of these results is seen in Tables 4 and 5, and in the Appendix.

A "summary metric" was devised that would combine the cross-correlation (CC) and RMSE errors into a single meaningful number. The metric was defined as the ratio of RMSE to cross-correlation value. Ideal trials with high cross-correlation (close to one) and low RMSE (close to zero) would score well (close to zero) on this scale, whereas less accurate measurements would score worse (larger values). Since the resulting metric lacks meaningful scale (which complicates interpretation), we chose to normalize the $\frac{RMSE}{CC}$ quotient by an arbitrary constant U_{ref} as follows:

$$Summary\ Metric_i = \frac{RMSE_i}{CC_i} * \frac{1}{U_{ref}}$$

where $Summary\ Metric_i$ is the $\frac{RMSE}{CC}$ quotient calculated for the i-th measurement and normalized by U_{ref}, that is given by:

$$U_{ref} = \frac{RMSE_{ref}}{CC_{ref}}$$

where $RMSE_{ref}$ was different for displacement-based and joint angle measures; and respectively chosen equal to 3 cm and 15 degrees. CC_{ref} was set to 0.75.

We chose CC_{ref} equal to 0.75 for the cross-correlation in order to make sure that the signals under observation had similar (linear) time-courses and the $RMSE_{ref}$ errors were chosen on the basis of previously reported data [33]. According to these arbitrary units and our own analysis, it was determined that summary measures ranging between 0 and 2 were ‘good’, values between 2 and 4 were approaching their limit of usability, and values greater than 4 were considered to represent measurements that Kinect cannot accurately capture.

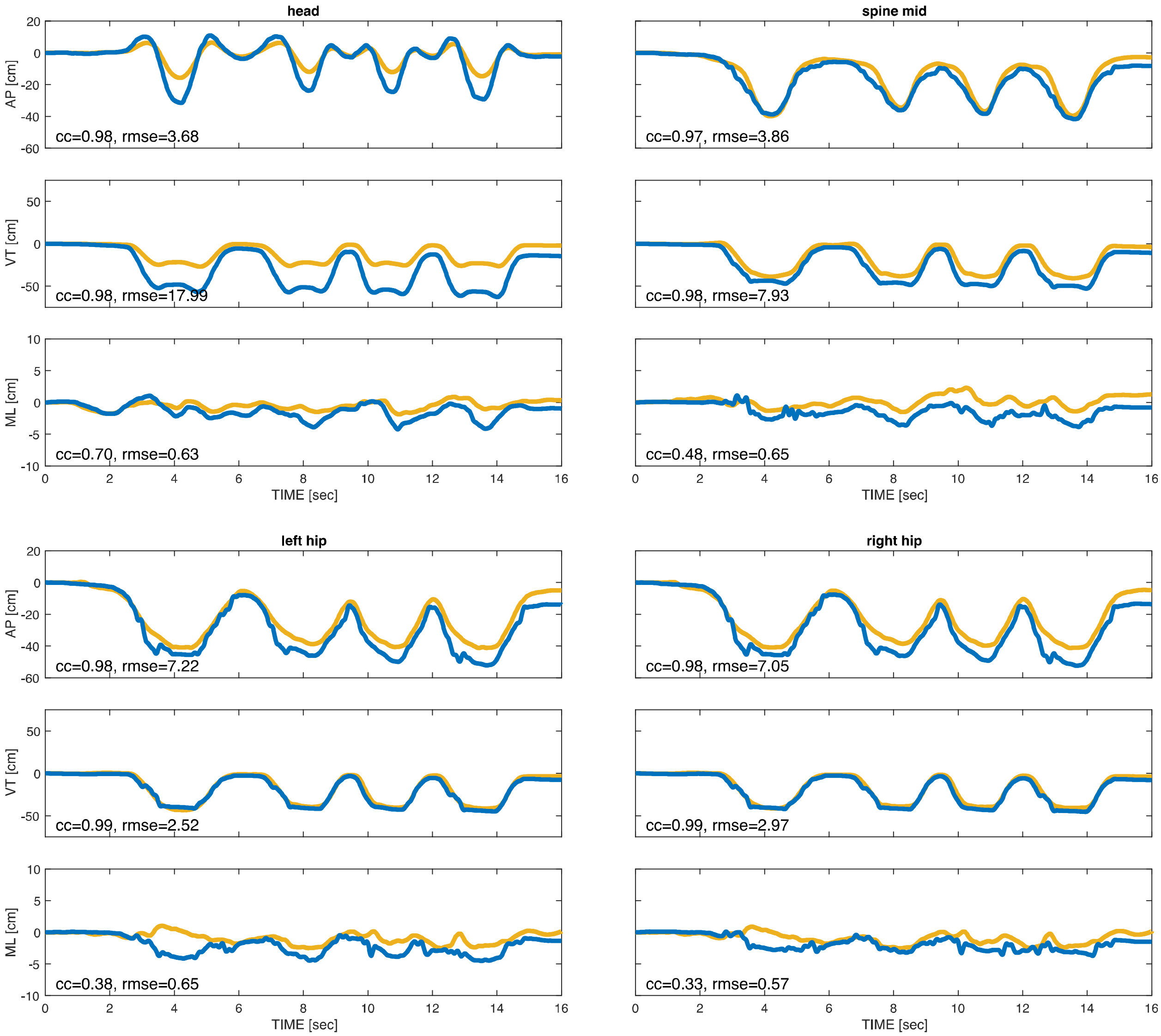


Figure 3: Head, spine middle, left and right hip joint center displacements in cm as derived from both MOCAP systems Qualisys (yellow) and Kinect (blue), during a sit to stand test. The displayed signals are for a single subject and a single trial. Values in the lower left corner of each plot show the cross-correlation coefficients and the root mean squared error calculated between the Qualisys and the Kinect displacements over time. ML = Mediolateral, VT = vertical, AP = anteroposterior

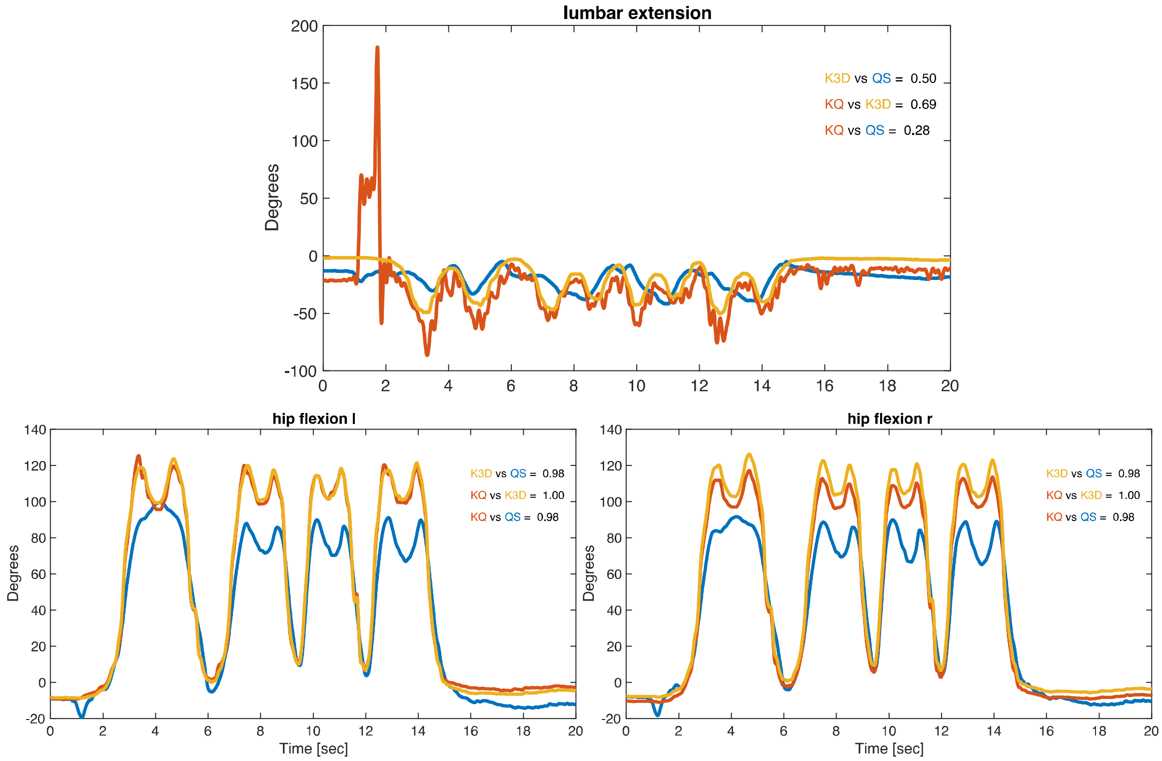


Figure 4: Joint angles for trunk extension, left and right hip flexion/extension. The blue, red and orange lines display joint angles derived using respectively Qualisys, Quaternion and coordinate data. The numbers show correlation coefficients between the different pairs of data

Table 2: Errors in tracking joint displacement. Errors are averaged over all trials of all 12 movements. AP = anteroposterior, VT = vertical, ML = mediolateral

	correlation				**error**				**summary metric**			
	AP	VT	ML	Average	AP	VT	ML	Average	AP	VT	ML	Average
head	0.93	0.98	0.93	0.95	3.85	4.99	5.58	4.80	1.59	1.32	0.72	1.21
spine top	0.73	0.78	0.68	0.73	3.17	2.42	1.75	2.45	0.68	0.69	0.35	0.57
spine mid	0.72	0.86	0.73	0.77	3.44	2.54	1.67	2.55	0.65	1.37	0.45	0.82
spine base	0.81	0.97	0.86	0.88	4.74	4.48	2.99	4.07	2.69	1.65	0.68	1.67
left shoulder	0.91	0.98	0.95	0.95	3.54	4.61	5.33	4.49	1.13	0.81	0.74	0.89
right shoulder	0.69	0.54	0.67	0.63	2.97	2.90	1.47	2.45	1.27	0.72	0.58	0.86
left elbow	0.66	0.54	0.70	0.63	3.45	3.37	1.41	2.74	1.84	1.16	0.96	1.32
right elbow	0.75	0.77	0.68	0.73	3.08	2.51	1.73	2.44	1.35	1.23	1.02	1.20
left wrist	0.94	0.71	0.88	0.84	2.57	2.12	1.27	1.99	1.02	1.28	1.54	1.28
right wrist	0.62	0.60	0.63	0.62	7.30	3.70	1.82	4.27	0.94	1.19	1.42	1.18
left hand	0.78	0.87	0.75	0.80	3.22	2.82	1.94	2.66	1.01	1.38	1.67	1.35
right hand	0.80	0.62	0.75	0.73	2.19	2.49	1.23	1.97	0.92	1.25	1.41	1.20
left hip	0.96	0.82	0.90	0.89	6.23	4.97	2.76	4.66	1.03	1.50	0.93	1.15
right hip	0.84	0.97	0.88	0.89	4.34	4.70	3.26	4.10	1.01	0.98	0.77	0.92
left knee	0.84	0.68	0.84	0.79	3.50	2.82	1.19	2.51	0.95	2.03	0.39	1.12
right knee	0.83	0.65	0.79	0.76	3.34	3.19	1.24	2.59	0.96	1.60	0.33	0.97
left ankle	0.50	0.47	0.56	0.51	5.28	2.96	1.88	3.37	0.93	1.63	0.53	1.03
right ankle	0.52	0.46	0.59	0.52	4.41	3.48	1.95	3.28	1.12	1.92	0.45	1.16
left foot	0.93	0.98	0.95	0.96	3.50	4.86	5.30	4.55	2.59	2.81	0.75	2.05
right foot	0.94	0.98	0.93	0.95	3.81	5.36	6.08	5.08	2.35	3.34	0.75	2.14

Table 3: Errors in tracking joint displacement. Errors are averaged over all trials of all 20 tracked body points. AP = anteroposterior, VT = vertical, ML = mediolateral

	correlation			error			summary metric		
Movement	AP	VT	ML	AP	VT	ML	AP	VT	ML
sit to stand	0.85	0.84	0.67	4.30	4.51	2.63	1.46	1.38	0.90
timed up and go	0.97	0.75	0.83	10.13	4.41	5.07	2.64	2.36	1.61
alternating barbell lunges	0.92	0.80	0.87	9.50	5.40	2.18	2.64	2.49	0.66
overhead squats	0.78	0.86	0.61	4.27	5.72	1.86	1.49	1.70	0.90
marching in place	0.88	0.80	0.92	2.17	2.05	1.86	0.62	0.63	0.50
time to stabilization	0.93	0.89	0.87	6.15	3.54	3.84	1.72	1.04	1.10
shoulder ab/adduction	0.60	0.75	0.65	1.24	1.98	0.98	0.66	0.93	0.34
shoulder flexion/extension	0.66	0.65	0.64	1.95	3.02	1.10	0.70	1.32	0.59
hip ab/adduction	0.71	0.78	0.95	1.85	2.61	2.85	0.84	0.88	0.75
combined hip/knee flexion/extension	0.85	0.87	0.97	1.58	2.24	3.88	0.50	0.64	1.00
combined arm ab/adduction	0.49	0.55	0.56	1.46	4.09	1.76	1.51	1.71	0.60
sagittal/frontal trunk leans	0.78	0.59	0.84	2.16	3.21	3.10	0.85	2.84	0.91

Table 4: Errors in tracking joint angles. Errors are averaged over all trials of all 12 movements. KQ = Kinect Quaternions, K3D = Kinect 3D-derived joint angles, QS = Qualisys, flxn = flexion, addn = adduction, extn = extension.

			correlation			**error**			**summary metric**		
			KQ vs QS	K3D vs QS	KQ vs K3D	KQ vs QS	K3D vs QS	KQ vs K3D	KQ vs QS	K3D vs QS	KQ vs K3D
arm	flxn	L	0.42	0.48	0.45	60.5	46.2	57.2	14.3	9.4	9.8
arm	flxn	R	0.40	0.48	0.53	77.9	54.3	52.7	13.1	8.3	6.5
arm	addn	L	0.65	0.77	0.77	41.9	22.3	26.5	4.1	1.7	2.4
arm	addn	R	0.64	0.75	0.75	47.7	27.7	29.4	5.3	2.1	2.4
elbow	flxn	L	0.64	0.38	0.36	27.8	41.7	28.5	2.3	6.0	4.7
elbow	flxn	R	0.69	0.44	0.42	27.5	41.3	25.3	2.0	6.8	10.4
hip	flxn	L	0.70	0.78	0.92	14.4	13.1	4.1	1.8	1.3	0.3
hip	flxn	R	0.71	0.71	0.98	13.7	14.4	2.6	2.5	1.9	0.2
hip	addn	L	0.36	0.35	0.64	6.2	6.8	9.4	4.4	1.5	2.7
hip	addn	R	0.30	0.30	0.68	5.0	7.9	4.5	1.6	2.7	1.5
knee	flxn	L	0.69	0.71	0.88	19.8	18.0	5.8	1.9	1.4	0.3
knee	flxn	R	0.66	0.68	0.97	21.0	19.7	4.0	2.3	2.5	0.1
knee	addn	L	0.25	0.40	0.45	22.0	10.0	20.8	2.9	1.5	2.3
knee	addn	R	0.57	0.44	0.70	12.5	12.3	18.3	1.7	1.8	1.5
lumbar	extn		0.38	0.51	0.40	38.0	19.9	25.2	6.1	4.4	4.5

Table 5: Errors in tracking joint angles. Errors are averaged over all trials of all 15 joint angles. KQ = Kinect Quaternions, K3D = Kinect 3D-derived joint angles, QS = Qualisys, alt = alternating, flex = flexion, ext = extension, sag/front = sagittal/frontal.

	correlation			**error**			**summary metric**		
Movement Task	KQ vs QS	K3D vs QS	KQ vs K3D	KQ vs QS	K3D vs QS	KQ vs K3D	KQ vs QS	K3D vs QS	KQ vs K3D
sit to stand	0.67	0.67	0.65	33.1	22.7	29.5	3.2	2.2	3.5
timed up and go	0.43	0.50	0.54	34.6	30.8	22.3	7.1	5.2	4.3
alt. barbell lunges	0.58	0.53	0.71	64.4	57.8	38.1	5.9	6.4	4.6
overhead squats	0.65	0.55	0.64	43.5	39.3	29.1	3.8	5.0	4.3
marching in place	0.62	0.61	0.62	15.4	13.3	9.7	1.5	1.8	1.4
time to stabilization	0.57	0.57	0.61	28.8	17.0	25.4	3.1	1.8	2.8
shoulder ab/adduction	0.42	0.41	0.70	14.4	14.7	13.0	4.0	4.5	7.0
shoulder flex/ext	0.33	0.36	0.73	16.1	11.6	11.0	5.0	3.6	1.3
hip ab/adduction	0.68	0.74	0.74	15.2	11.5	11.3	1.9	0.9	1.3
comb. hip/knee flex/ext	0.59	0.65	0.75	18.6	12.5	14.4	2.3	1.3	1.8
comb. arm ab/adduction	0.35	0.33	0.65	31.2	26.1	23.7	8.7	7.3	3.2
sag./front. trunk leans	0.57	0.65	0.60	33.4	27.5	24.1	6.6	2.6	4.2

For compactness sake, Tables 2-5 summarize errors by averaging over either movement or limb/joint. A full presentation of all error combinations can be found in the Appendix. Those errors are presented in the interest of investigators seeking to optimally calibrate their own Kinect-based motion tracking systems. To simplify data interpretation, the values in Tables 10 and 11 are color-coded. Green values indicate that the Kinect-based performance was good (values ranging between 0 and 2). Yellow values indicate that the performance is approaching the limit of usability, with values ranging between 2 and 4. Red values indicate those conditions where the Kinect-based MOCAP yielded poor performance, with values larger than 4.

In the tables above, for purposes of presentation, we averaged the performance metrics across movements. This allows a single value to represent how well the Kinect-based MOCAP was capable of evaluating each joint both in terms of displacements and in terms of joint angles. Additionally, in the case of the joint displacement, we averaged metrics across joints and then across the three axes to identify which axis captured the motions with the highest accuracy. Because averaging obscures performance characteristics which may be meaningful in certain applications, the data are also presented in their original, uncollapsed form in the appendices.

5 Discussion

The overall finding of this work has been that data from the Kinect compare favorably to the gold-standard Qualisys tracking system, given the limitations of the Kinect hardware. This work has quantified, for the first time, the specific limitations of a Kinect-based motion tracking system for general applications with respect to a set of representative clinical movement tests. Furthermore, whereas others have quantified Kinect's limitations either through global summary metrics or by collapsing data across movement planes, this work presents more fine-grained performance comparisons. Specifically, this data is critical for investigators who need to know the precision that can be expected when using Kinect to track motion in real-world settings.

One of the most relevant aspects of this work is to fully evaluate the potential of a Kinect-based multi-purpose MOCAP that can be useful in clinical settings. This motivated our choice to analyze system performance with respect to different clinically relevant motions, and to keep the results grouped by movement. One important finding was that some motions were better captured than others, regardless of which metrics were used in quantifying error. One hypothesis to explain the movement-based differential in tracking performance is that certain movements may be similar to those that Microsoft used when designing and calibrating the Kinect for video game play. Tables 3 and 5 show that certain motions such as hip ab/adduction, marching in place, hip/knee flexion/extension, time to stabilization and sit to stand movements yielded the highest agreement between Qualisys and Kinect. This finding suggests that a Kinect-based MOCAP system can be used more confidently when investigating the above-mentioned motions. Identifying the limitations of such a system is valuable for all those investigators and medical professionals in need of carrying out motion analysis studies using light-weight and low-cost equipment.

Another relevant observation was that there were key differences when joint angles were calculated from Kinect's quaternion stream versus being derived from the Kinect 3D coordinates. Specifically, the 3D coordinate approach was generally superior when tracking arm ab/adduction, hip flexion/extension and ab/adduction, knee flexion/extension and ab/adduction. In theory, the 3D coordinate approach should only work when the Kinect's global reference system is aligned with the subject's anatomical planes (e.g. for movements such as jumping jacks where the limbs remain in the frontal plane). To overcome this limitation, joint locations can be recomputed relative to a local rotated frame prior to computing joint angles. With this correction, only elbow flexion angles were better computed using quaternions than the 3D coordinate trigonometric method (see Tables 4 and 5).

Lower performance in the quaternion approach can be due to multiple concurrent factors. First, the Kinect sensor shows frequent errors and oscillations in evaluating joint orientations during motion. Without a strict rigid body model to be superimposed onto the Kinect raw data, it is not possible to compensate for such sensor errors. Secondly, when allowing the subjects to move in 3D space without constraining their movements to specific planes, the conventional kinematic angles cannot be easily derived from quaternions without choosing a three-axial rotation sequence. The difficulty of selecting an appropriate rotation sequence based on joint and motion type is a well-known problem in biomechanics; recommended rotation sequences for various joints have been defined [22]–[24], [34]. However, these sequences are defined relative to local reference coordinate systems that do not exist when using quaternions. It is therefore not unexpected that the same rotation sequences used as standard practice in biomechanics research are not optimal when used with Kinect quaternions. We evaluated all possible Kinect quaternion rotation sequence combinations and chose those that yielded the joint angles most closely approximating the gold-standard Qualisys kinematics. The identified rotations were then applied at a given joint irrespective of the movement being tested.

Kinect-based motion capture is not flawless, as emphasized by our results. Inaccuracies are mainly due to a combination of hardware and software limitations. From a hardware perspective, the Kinect camera is based on a combination of low-cost commodity sensors, such as a single depth sensor. With respect to software, the Microsoft body tracking algorithm is designed for gaming performance and generalization rather than tracking accuracy. The Microsoft tracking algorithm aims to detect people in the sensor field of view irrespective of their pose and body type. Consequently, for the sake of performance and usability, the Kinect system was devised to minimize the body model constraints necessary and to choose speed over accuracy. For instance, to reduce the computational time necessary to track joints, the sensor depth data stream is always processed in real-time and on a frame-by-frame basis. Although computationally efficient, it allows for predicted joint locations to jump from frame to frame depending on instantaneous sensor measurements. Consequently, it does not enforce rigid body constraints, which are a core assumption of most biomechanics studies.

We hypothesize that Kinect-based motion tracking could be improved by third-party algorithms which combine sensor data with various real-world constraints such as fixed anatomy and range of motion. Contemporary processor speeds should be in line with the types of processing demands that are necessary to implement these improvements within reasonable time frames. The combination of a low-cost, portable, expedient sensor with high-performance data modeling could yield systems whose performances are acceptable in most medical, sports, fitness and rehabilitation applications.

Although the number of subjects in this study may be a limitation, we expect that the broad extent of the movements evaluated in this work would contribute more variability to motion capture performance than would differences between subjects. Because the focus of this study was to evaluate Kinect performance for non-specific purposes, inter-subject variability was of less interest than variability in fine-grained body tracking behavior across a range of movements. In future work, we intend to build on our present findings with group-based investigations to confirm the viability of the Kinect for observing clinically relevant movement features. This will require a more narrowly focused scope with depth in sample size as opposed to experimental conditions.

6 Conclusion

The Kinect-based motion capture system can consistently track the 3D displacement of joint centers with high precision and acceptable accuracy levels. Specifically, our results show that the Kinect-based MOCAP and the Qualisys system reached high levels of agreement when tracking joint displacements, with average overall cross-correlation coefficient of 0.78, root mean squared error of 3.35 cm and a combined metric of 1.21. On the contrary, lower agreement levels were achieved when tracking joint angles with cross-correlation coefficient of 0.58, root mean squared error of 24.59 degrees and a summary metric of 3.76.

This is a promising finding, considering the ease of setup, use and cost of a Kinect-based system. The overall analysis of displacement tracking performance showed that some segments/joints are tracked with less accuracy, particularly the foot/ankle complex. Also, the accuracy levels are axis dependent, with the highest accuracy recorded along the mediolateral direction. Furthermore, the overall data analysis indicated a motion-dependent tracking accuracy, with timed up-and-go, alternating barbell lunges and combined arm ab/adduction yielding the lowest performance and largest errors. Despite some of the Kinect precision limitations, the displacement data were consistently within a threefold difference from the precision and accuracy levels that we used to normalize the performance metrics.

The data presented here suggest that Kinect-based motion capture systems may be viable alternatives to professional three-dimensional capture systems for certain applications. Our data can be used by other investigators to understand the limits of out-of-the-box Kinect motion capture accuracy with respect to various movements and planes of motion. We hypothesize that, with the addition of certain body constraints, Kinect-based tracking systems could be valuable tools in applications such as medicine, sports, rehabilitation and fitness.

Acknowledgements

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs through the Neurosensory and Rehabilitation Research Award Program under Award No. W81XWH-15-1-0045. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense. The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702- 5014 is the awarding and administering acquisition office.

7 References

[1] Microsoft, “Kinect for Windows SDK 2.0.” [Online]. Available: https://www.microsoft.com/en-us/download/details.aspx?id=44561. [Accessed: 01-Jan-2017].

[2] B. Bonnechere, B. Jansen, P. Salvia, H. Bouzahouene, L. Omelina, F. Moiseev, V. Sholukha, J. Cornelis, M. Rooze, and S. Van Sint Jan, “Validity and reliability of the Kinect within functional assessment activities: Comparison with standard stereophotogrammetry,” *Gait Posture*, vol. 39, no. 1, pp. 593–598, 2014.

[3] R. A. Clark, Y. H. Pua, K. Fortin, C. Ritchie, K. E. Webster, L. Denehy, and A. L. Bryant, “Validity of the Microsoft Kinect for assessment of postural control,” *Gait Posture*, vol. 36, no. 3, pp. 372–377, 2012.

[4] B. F. Mentiplay, L. G. Perraton, K. J. Bower, Y. H. Pua, R. McGaw, S. Heywood, and R. A. Clark, “Gait assessment using the Microsoft Xbox One Kinect: Concurrent validity and inter-day reliability of spatiotemporal and kinematic variables,” *J. Biomech.*, vol. 48, no. 10, pp. 2166–2170, 2015.

[5] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester, “Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson’s disease,” *Gait Posture*, vol. 39, no. 4, pp. 1062–1068, 2014.

[6] M. Eltoukhy, A. Kelly, C.-Y. Kim, H.-P. Jun, R. Campbell, and C. Kuenze, “Validation of the Microsoft Kinect® camera system for measurement of lower extremity jump landing and squatting kinematics.,” *Sports Biomech.*, vol. 3141, no. February, pp. 1–14, 2016.

[7] D. Lim, C. Kim, H. Jung, D. Jung, and K. Chun, “Use of the microsoft kinect system to characterize balance ability during balance training,” *Clin. Interv. Aging*, vol. 10, pp. 1077–1083, 2015.

[8] S. Phommahavong, D. Haas, J. Yu, S. Krüger-Ziolek, K. Möller, and J. Kretschmer, “Evaluating the microsoft kinect skeleton joint tracking as a tool for home-based physiotherapy,” *Curr. Dir. Biomed. Eng.*, vol. 1, no. 1, pp. 184–187, 2015.

[9] A. Pfister, A. M. West, S. Bronner, and J. A. Noah, “Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis.,” *J. Med. Eng. Technol.*, vol. 1902, no. August, pp. 1–7, 2014.

[10] E. E. Stone, M. Butler, A. McRuer, A. Gray, J. Marks, and M. Skubic, “Evaluation of the Microsoft Kinect for screening ACL injury,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 4152–4155, 2013.

[11] R. A. Clark, S. Vernon, B. F. Mentiplay, K. J. Miller, J. L. McGinley, Y. H. Pua, K. Paterson, and K. J. Bower, “Instrumenting gait assessment using the Kinect in people living with stroke: reliability and association with balance tests.,” *J. Neuroeng. Rehabil.*, vol. 12, p. 15, 2015.

[12] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, “Real-time human pose recognition in parts from single depth images,” *Stud. Comput. Intell.*, vol. 411, pp. 119–135, 2013.

[13] R. A. Clark, K. J. Bower, B. F. Mentiplay, K. Paterson, and Y. H. Pua, “Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables,” *J. Biomech.*, vol. 46, no. 15, pp. 2722–2725, 2013.

[14] X. Xu, R. W. McGorry, L. S. Chou, J. hua Lin, and C. chi Chang, “Accuracy of the Microsoft Kinect^{TM} for measuring gait parameters during treadmill walking,” *Gait Posture*, vol. 42, no. 2, pp.

145–151, 2015.

[15] E. Dolatabadi, B. Taati, and A. Mihailidis, "Concurrent validity of the Microsoft Kinect for Windows v2 for measuring spatiotemporal gait parameters," *Med. Eng. Phys.*, vol. 38, pp. 952–958, 2016.

[16] Q. Wang, G. Kurillo, F. Ofli, and R. Bajcsy, "Evaluation of Pose Tracking Accuracy in the First and Second Generations of Microsoft Kinect," no. November, 2015.

[17] D. J. Geerse, B. H. Coolen, and M. Roerdink, "Kinematic validation of a multi-Kinect v2 instrumented 10-meter walkway for quantitative gait assessments," *PLoS One*, vol. 10, no. 10, pp. 1–15, 2015.

[18] Q. Wang, G. Kurillo, F. Ofli, and R. Bajcsy, "Remote Health Coaching System and Human Motion Data Analysis for Physical Therapy with Microsoft Kinect," no. 1111965, 2015.

[19] L. F. Yeung, K. C. Cheng, C. H. Fong, W. C. C. Lee, and K.-Y. Tong, "Evaluation of the Microsoft Kinect as a clinical assessment tool of body sway.," *Gait Posture*, vol. 40, no. 4, pp. 532–8, 2014.

[20] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, 2002.

[21] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time Human Pose Recognition in Parts from Single Depth Images," *Comput. Vis. Pattern Recognit.*, pp. 1297–1304, 2011.

[22] G. Wu, S. Siegler, P. Allard, C. Kirtley, A. Leardini, D. Rosenbaum, M. Whittle, D. D. D'Lima, L. Cristofolini, H. Witte, O. Schmid, and I. Stokes, "ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion—part I: ankle, hip, and spine," *J. Biomech.*, vol. 35, no. 4, pp. 543–548, 2002.

[23] G. Wu, S. Siegler, P. Allard, C. Kirtley, A. Leardini, D. Rosenbaum, M. Whittle, D. D. D'Lima, L. Cristofolini, H. Witte, O. Schmid, I. Stokes, P. R. Cavanagh, and A. Mikaberidze, "Letter To The Editor: 'Letter to the Editor,'" *J. Biomech.*, vol. 35, no. 4, pp. 543–548, 2002.

[24] G. Wu, F. C. T. Van Der Helm, H. E. J. Veeger, M. Makhsous, P. Van Roy, C. Anglin, J. Nagels, A. R. Karduna, K. McQuade, X. Wang, F. W. Werner, and B. Buchholz, "ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion - Part II: Shoulder, elbow, wrist and hand," *J. Biomech.*, vol. 38, no. 5, pp. 981–992, 2005.

[25] A. Cappozzo, U. Della Croce, A. Leardini, and L. Chiari, "Human movement analysis using stereophotogrammetry: Part 1: theoretical background," *Gait Posture*, vol. 21, no. 2, pp. 186–196, 2005.

[26] J. Hamill and W. S. Selbie, "Three Dimensional Kinematics," in *Research Methods in Biomechanics*, D. G. E. Robertson, G. E. Caldwell, J. Hamill, G. Kamen, and S. N. Whittlesey, Eds. Champaign, IL: Human Kinetics Publishers, 2004.

[27] D. Winter, *Biomechanics and motor control of human movement.* Wiley & Sons, 2009.

[28] R. B. Davis, S. Õunpuu, D. Tyburski, and J. R. Gage, "A gait analysis data collection and reduction technique," *Hum. Mov. Sci.*, vol. 10, no. 5, pp. 575–587, 1991.

[29] D. SL, A. FC, A. AS, L. P, H. A, J. CT, G. E, and T. DG, "OpenSim: Open source to create and analyze dynamic simulations of movement," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 11, pp. 1940–1950, 2007.

[30] A. Napoli, C. R. Ward, S. M. Glass, C. Tucker, and I. Obeid, “Automated Assessment of Postural Stability System,” in *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, 2016, pp. 6090–6093.

[31] K.-A. Kwon, R. J. Shipley, M. Edirisinghe, D. G. Ezra, G. Rose, S. M. Best, and R. E. Cameron, “High-speed camera characterization of voluntary eye blinking kinematics.,” *J. R. Soc. Interface*, vol. 10, no. 85, p. 20130227, 2013.

[32] R. A. Clark, A. L. Bryant, Y. Pua, P. McCrory, K. Bennell, and M. Hunt, “Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance,” *Gait Posture*, vol. 31, no. 3, pp. 307–310, 2010.

[33] R. A. Clark, Y. H. Pua, C. C. Oliveira, K. J. Bower, S. Thilarajah, R. McGaw, K. Hasanki, and B. F. Mentiplay, “Reliability and concurrent validity of the Microsoft Xbox One Kinect for assessment of standing balance and postural control,” *Gait Posture*, vol. 42, no. 2, pp. 210–213, 2015.

[34] M. Šenk, L. Chèze, G. Wu, F. C. T. Van Der Helm, H. E. J. Veeger, M. Makhsous, P. Van Roy, C. Anglin, J. Nagels, A. R. Karduna, K. McQuade, X. Wang, F. W. Werner, and B. Buchholz, “Rotation sequence as an important factor in shoulder kinematics,” *J. Biomech.*, vol. 38, no. 5, pp. 981–992, 2005.

8 Appendix

Table 6: Average Cross-Correlation Coefficients of Joint Center Displacements measured across two separate trials from four different subjects (each subject repeated the tests twice) and grouped by movement types.

	SIT TO STAND			TIMED UP AND GO			ALTERNATING BARBELL LUNGES			OVERHEAD SQUATS			MARCHING IN PLACE			TIME TO STABILIZATION		
JOINT DISP	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML
head	0.99	0.99	0.92	0.98	0.98	0.76	0.98	0.98	0.93	0.98	0.99	0.83	0.98	0.98	0.99	0.99	0.97	0.94
spine top	0.97	0.99	0.37	0.99	0.97	0.88	0.98	0.98	0.87	0.78	0.85	0.49	0.56	0.65	0.91	0.89	0.96	0.79
right shoulder	0.56	0.39	0.65	0.95	0.21	0.85	0.83	0.17	0.65	0.72	0.58	0.33	0.96	0.90	0.92	0.89	0.73	0.91
left shoulder	0.99	0.99	0.95	0.98	0.98	0.81	0.98	0.99	0.95	0.97	0.99	0.90	0.98	0.98	0.98	0.99	0.97	0.99
right elbow	0.97	0.99	0.34	0.99	0.97	0.88	0.98	0.98	0.86	0.69	0.90	0.54	0.80	0.37	0.92	0.88	0.95	0.77
left elbow	0.71	0.40	0.68	0.97	0.31	0.90	0.79	0.14	0.89	0.71	0.62	0.48	0.96	0.93	0.95	0.83	0.87	0.75
right wrist	0.48	0.97	0.36	0.90	0.97	0.77	0.94	0.99	0.85	0.78	0.87	0.37	0.53	0.23	0.81	0.88	0.79	0.66
left wrist	0.91	0.99	0.83	0.98	0.98	0.93	0.95	0.99	0.94	0.86	0.89	0.72	0.94	0.53	0.96	0.99	0.97	0.91
right hand	0.97	0.99	0.75	0.98	0.98	0.89	0.95	0.99	0.95	0.84	0.91	0.64	0.88	0.68	0.94	0.94	0.93	0.89
left hand	0.92	0.99	0.34	0.98	0.98	0.64	0.94	0.98	0.91	0.53	0.91	0.58	0.71	0.74	0.94	0.96	0.98	0.91
spine mid	0.91	0.98	0.53	0.99	0.98	0.78	0.90	0.99	0.87	0.69	0.97	0.52	0.81	0.78	0.95	0.92	0.98	0.85
spine base	0.95	0.99	0.56	0.98	0.98	0.70	0.97	0.98	0.91	0.76	0.97	0.50	0.97	0.95	0.96	0.97	0.95	0.95
right hip	0.93	0.99	0.76	0.98	0.98	0.79	0.93	0.98	0.88	0.70	0.97	0.39	0.97	0.95	0.97	0.97	0.96	0.96
left hip	0.97	0.98	0.78	0.98	0.97	0.97	0.98	0.99	0.96	0.98	0.97	0.79	0.93	0.87	0.93	0.99	0.92	0.95
right knee	0.94	0.62	0.87	0.98	0.30	0.84	0.98	0.85	0.88	0.91	0.84	0.91	0.99	0.93	0.88	0.92	0.82	0.77
left knee	0.95	0.61	0.88	0.98	0.37	0.93	0.97	0.87	0.94	0.94	0.84	0.89	0.99	0.88	0.94	0.92	0.86	0.91
right ankle	0.54	0.45	0.55	0.90	0.11	0.81	0.71	0.17	0.62	0.43	0.59	0.42	0.82	0.80	0.67	0.89	0.53	0.77
left ankle	0.41	0.51	0.49	0.93	0.07	0.82	0.74	0.09	0.76	0.32	0.61	0.24	0.81	0.82	0.74	0.73	0.74	0.74
right foot	0.99	0.99	0.93	0.98	0.98	0.75	0.97	0.98	0.91	0.98	0.99	0.78	0.98	0.98	0.99	0.99	0.96	0.96
left foot	0.99	0.99	0.95	0.98	0.98	0.81	0.98	0.99	0.93	0.98	0.99	0.88	0.98	0.99	0.99	0.99	0.97	0.99

	SHOULDER AB/ADDUCTION			SHOULDER FLEXION/ EXTENSION			HIP AB/ADDUCTION			COMBINED HIP/KNEE FLEXION/ EXTENSION			COMBINED ARM AB/ADDUCTION			SAGITTAL/FRONTAL TRUNK LEANS		
JOINT DISP	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML
head	0.96	1.00	1.00	0.98	0.99	0.99	0.90	0.98	0.94	0.91	0.99	0.95	0.75	0.94	0.95	0.77	0.98	0.93
spine top	0.15	0.95	0.60	0.35	0.90	0.12	0.68	0.70	0.97	0.89	0.85	0.96	0.60	0.54	0.32	0.94	0.07	0.87
right shoulder	0.37	0.59	0.33	0.42	0.36	0.34	0.84	0.87	0.94	0.86	0.93	0.94	0.29	0.31	0.36	0.54	0.42	0.84
left shoulder	0.89	1.00	1.00	0.99	0.99	0.99	0.78	0.97	0.98	0.85	0.98	0.99	0.69	0.94	0.96	0.82	0.98	0.96
right elbow	0.16	0.95	0.50	0.52	0.80	0.15	0.50	0.81	0.97	0.85	0.92	0.98	0.67	0.45	0.29	0.94	0.18	0.93
left elbow	0.34	0.53	0.33	0.31	0.24	0.35	0.86	0.91	0.91	0.81	0.98	0.98	0.12	0.20	0.28	0.53	0.34	0.94
right wrist	0.55	0.19	0.30	0.20	0.33	0.45	0.52	0.14	0.94	0.42	0.53	0.98	0.28	0.45	0.35	0.91	0.76	0.69
left wrist	0.92	0.78	0.76	0.96	0.31	0.75	0.91	0.53	0.96	0.95	0.35	0.99	0.89	0.66	0.77	0.98	0.55	0.99
right hand	0.32	0.54	0.69	0.75	0.25	0.41	0.61	0.36	0.97	0.89	0.46	0.99	0.90	0.25	0.66	0.59	0.12	0.26
left hand	0.80	0.97	0.67	0.68	0.91	0.85	0.73	0.83	0.97	0.94	0.94	0.99	0.22	0.35	0.17	0.97	0.90	0.98
spine mid	0.24	0.97	0.20	0.56	0.94	0.60	0.45	0.60	0.97	0.92	0.84	0.99	0.27	0.54	0.48	0.93	0.78	0.98
spine base	0.70	0.99	0.99	0.94	0.98	0.97	0.65	0.98	0.96	0.94	0.98	0.99	0.14	0.91	0.83	0.79	0.98	0.96
right hip	0.73	0.99	0.99	0.98	0.98	0.98	0.65	0.96	0.97	0.95	0.98	0.99	0.39	0.89	0.93	0.89	0.96	0.92
left hip	0.96	0.49	0.84	0.97	0.77	0.87	0.90	0.42	0.96	0.96	0.86	0.99	0.86	0.69	0.82	0.99	0.90	0.99
right knee	0.62	0.80	0.64	0.52	0.33	0.77	0.92	0.86	0.95	0.88	0.93	0.94	0.33	0.42	0.13	0.95	0.09	0.96
left knee	0.46	0.84	0.59	0.66	0.61	0.63	0.94	0.87	0.94	0.91	0.97	0.97	0.43	0.31	0.51	0.93	0.10	0.98
right ankle	0.39	0.34	0.27	0.30	0.24	0.44	0.22	0.87	0.90	0.63	0.94	0.94	0.18	0.25	0.28	0.18	0.26	0.37
left ankle	0.45	0.16	0.24	0.25	0.17	0.23	0.46	0.92	0.90	0.55	0.98	0.97	0.06	0.08	0.16	0.36	0.44	0.41
right foot	0.97	1.00	1.00	0.98	0.99	0.99	0.91	0.98	0.95	0.93	0.99	0.96	0.90	0.95	0.96	0.75	0.97	0.92
left foot	0.96	1.00	1.00	0.99	0.99	0.99	0.79	0.97	0.97	0.91	0.98	0.99	0.74	0.95	0.97	0.90	0.98	0.97

Table 7: Average Cross-Correlation Coefficients of Joint Angles measured across the same trials used in Table 6 and grouped by movement.

	SIT TO STAND			TIMED UP AND GO			ALTERNATING BARBELL LUNGES			OVERHEAD SQUATS			MARCHING IN PLACE			TIME TO STABILIZATION		
JOINT ANGLES	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ
arm flexion r	0.50	0.56	0.50	0.13	0.26	0.19	0.41	0.41	0.87	0.41	0.24	0.58	0.32	0.62	0.40	0.30	0.46	0.61
arm flexion l	0.86	0.44	0.44	0.19	0.65	0.15	0.67	0.49	0.72	0.58	0.43	0.26	0.56	0.70	0.45	0.47	0.48	0.33
arm adduction r	0.62	0.61	0.50	0.22	0.68	0.21	0.93	0.98	0.93	0.75	0.81	0.86	0.65	0.95	0.65	0.47	0.48	0.58
arm adduction l	0.44	0.61	0.60	0.18	0.70	0.24	0.70	0.74	0.85	0.69	0.70	0.89	0.81	0.92	0.87	0.46	0.58	0.61
elbow flexion r	0.64	0.67	0.49	0.82	0.51	0.57	0.82	0.46	0.43	0.83	0.55	0.41	0.56	0.41	0.36	0.90	0.38	0.46
elbow flexion l	0.63	0.65	0.36	0.72	0.23	0.36	0.69	0.32	0.24	0.69	0.36	0.41	0.55	0.37	0.53	0.76	0.55	0.45
hip flexion r	0.98	0.98	1.00	0.91	0.90	0.93	0.62	0.64	1.00	0.94	0.95	1.00	0.92	0.92	1.00	0.82	0.80	0.99
hip flexion l	0.98	0.98	1.00	0.89	0.90	0.94	0.66	0.65	0.99	0.92	0.94	0.99	0.93	0.93	1.00	0.87	0.90	0.97
hip adduction r	0.69	0.42	0.57	0.27	0.35	0.40	0.20	0.27	0.61	0.45	0.13	0.22	0.40	0.54	0.44	0.28	0.54	0.54
hip adduction l	0.40	0.56	0.80	0.34	0.43	0.50	0.16	0.24	0.73	0.35	0.36	0.86	0.51	0.47	0.53	0.39	0.31	0.27
knee flexion r	0.96	0.96	1.00	0.47	0.47	1.00	0.90	0.93	0.96	0.75	0.76	1.00	0.84	0.85	1.00	0.85	0.86	0.99
knee flexion l	0.96	0.96	1.00	0.50	0.54	0.94	0.92	0.93	1.00	0.74	0.74	1.00	0.79	0.80	0.99	0.86	0.92	0.94
knee adduction r	0.87	0.53	0.58	0.56	0.41	0.63	0.43	0.27	0.58	0.77	0.22	0.17	0.64	0.32	0.62	0.50	0.39	0.65
knee adduction l	0.25	0.59	0.25	0.11	0.36	0.31	0.27	0.36	0.53	0.20	0.36	0.39	0.55	0.13	0.30	0.32	0.36	0.20
lumbar extension	0.33	0.56	0.66	0.19	0.11	0.75	0.35	0.21	0.24	0.69	0.65	0.58	0.25	0.15	0.16	0.28	0.46	0.53

	SHOULDER AB/ADDUCTION			SHOULDER FLEXION/ EXTENSION			HIP AB/ADDUCTION			COMBINED HIP/ KNEE FLEXION/ EXTENSION			COMBINED ARM AB/ADDUCTION			SAGITTAL/FRONTAL TRUNK LEANS		
JOINT ANGLES	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ
arm flexion r	0.22	0.30	0.83	0.96	0.90	0.90	0.60	0.81	0.57	0.45	0.53	0.42	0.11	0.23	0.31	0.40	0.49	0.25
arm flexion l	0.17	0.23	0.76	0.91	0.75	0.82	0.33	0.80	0.57	0.20	0.44	0.29	0.11	0.08	0.40	0.05	0.23	0.21
arm adduction r	0.96	0.97	0.96	0.27	0.35	0.98	0.98	0.98	0.99	0.95	0.96	0.97	0.13	0.35	0.62	0.78	0.82	0.79
arm adduction l	0.98	0.89	0.89	0.38	0.43	0.98	0.95	0.97	0.98	0.88	0.97	0.94	0.56	0.90	0.62	0.78	0.87	0.78
elbow flexion r	0.28	0.14	0.09	0.35	0.12	0.22	0.72	0.71	0.74	0.68	0.57	0.52	0.74	0.18	0.23	0.97	0.57	0.57
elbow flexion l	0.21	0.13	0.19	0.51	0.10	0.42	0.79	0.65	0.55	0.55	0.33	0.29	0.69	0.59	0.24	0.84	0.24	0.22
hip flexion r	0.26	0.36	0.98	0.24	0.22	0.96	0.91	0.91	1.00	0.85	0.76	0.97	0.36	0.36	0.98	0.71	0.74	0.98
hip flexion l	0.68	0.67	0.96	0.20	0.39	0.86	0.95	0.95	1.00	0.45	0.83	0.78	0.31	0.40	0.94	0.55	0.86	0.66
hip adduction r	0.13	0.14	0.94	0.03	0.06	0.91	0.68	0.49	0.68	0.17	0.23	0.88	0.25	0.26	0.99	0.09	0.17	0.96
hip adduction l	0.15	0.13	0.60	0.13	0.28	0.73	0.59	0.68	0.78	0.59	0.29	0.85	0.18	0.22	0.73	0.55	0.24	0.32
knee flexion r	0.48	0.46	0.94	0.07	0.08	0.91	0.87	0.88	1.00	0.72	0.83	0.96	0.21	0.23	0.99	0.79	0.89	0.95
knee flexion l	0.52	0.40	0.66	0.15	0.38	0.40	0.90	0.88	1.00	0.84	0.86	0.94	0.36	0.25	0.74	0.72	0.86	0.88
knee adduction r	0.59	0.28	0.69	0.28	0.35	0.97	0.37	0.48	0.64	0.84	0.83	1.00	0.17	0.25	0.91	0.87	0.88	0.99
knee adduction l	0.17	0.21	0.55	0.22	0.26	0.77	0.13	0.47	0.29	0.24	0.49	0.90	0.34	0.28	0.74	0.17	0.92	0.17
lumbar extension	0.46	0.79	0.42	0.22	0.65	0.15	0.40	0.40	0.27	0.37	0.77	0.51	0.75	0.39	0.25	0.26	0.97	0.30

Table 8: Average Errors in Kinect-based MOCAP Joint Center Displacement estimation across two separate trials from four different subjects (each subject repeated the tests twice) and grouped by movement types

	SIT TO STAND			TIMED UP AND GO			ALTERNATING BARBELL LUNGES			OVERHEAD SQUATS			MARCHING IN PLACE			TIME TO STABILIZATION		
JOINT DISP	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML
head	4.8	4.1	8.3	7.2	4.5	8.5	8.4	8.5	1.9	3.2	4.2	1.6	3.3	2.2	3.6	3.5	3.2	6.7
spine top	7.7	3.4	1.3	5.8	3.8	2.8	7.8	3.4	2.0	5.6	7.4	3.7	1.7	0.8	1.7	3.2	1.4	2.0
right shoulder	2.3	2.3	0.8	11.2	2.2	2.6	10.3	3.0	3.4	2.8	4.4	1.1	1.5	4.3	0.8	2.7	2.5	1.1
left shoulder	3.9	3.8	8.3	8.7	5.0	10.6	7.7	8.0	2.7	3.0	4.1	0.9	3.0	2.2	4.2	3.3	3.2	9.5
right elbow	7.1	2.6	1.3	5.3	4.2	3.2	7.4	4.5	2.1	5.7	7.8	3.7	1.5	0.9	1.6	3.5	1.1	2.1
left elbow	1.9	2.4	0.4	11.8	2.4	3.3	10.8	3.1	2.0	2.1	4.7	1.0	1.1	4.1	0.8	8.8	12.5	4.0
right wrist	17.8	8.5	0.8	23.9	6.5	3.7	8.4	6.5	2.4	8.9	9.9	1.2	1.6	0.7	1.5	18.0	1.3	2.5
left wrist	2.8	5.3	0.6	6.1	3.9	2.0	8.3	2.5	1.0	3.0	4.2	0.7	0.6	0.4	1.1	3.2	2.8	2.4
right hand	2.7	5.9	0.7	6.7	4.8	2.4	5.3	2.6	1.1	2.7	4.2	0.9	0.9	0.5	1.2	2.8	2.8	2.1
left hand	2.5	3.7	1.2	7.7	3.7	4.0	10.4	4.6	1.1	3.5	7.2	4.3	1.8	1.3	1.0	3.4	1.8	2.3
spine mid	2.5	4.9	1.5	8.4	4.4	5.4	9.9	2.9	1.5	3.9	3.8	1.3	2.4	1.1	0.9	4.5	2.1	2.2
spine base	5.2	2.4	2.7	6.5	4.3	8.1	10.4	12.0	2.2	8.9	6.8	2.3	4.9	2.0	3.6	3.1	2.3	2.2
right hip	3.4	3.0	5.7	7.3	5.2	8.6	11.4	11.8	1.7	11.3	4.8	2.8	4.6	3.5	3.3	2.1	2.8	5.3
left hip	3.7	16.2	0.5	22.1	9.6	3.5	10.5	6.5	2.2	2.4	16.5	0.6	0.8	0.2	1.8	25.0	3.6	4.2
right knee	3.2	2.8	0.6	7.3	2.7	2.6	10.2	4.5	3.1	3.6	4.6	1.7	2.5	2.4	1.0	2.8	5.2	1.3
left knee	3.8	2.9	0.6	7.4	2.8	3.0	12.0	5.3	1.5	3.6	5.4	2.1	2.4	2.3	0.8	4.9	3.6	2.1
right ankle	2.4	3.7	1.0	15.2	3.5	3.5	12.8	3.9	3.9	2.7	3.7	1.7	2.1	3.3	1.6	4.5	2.8	2.4
left ankle	2.2	3.5	0.6	16.0	2.3	4.2	12.7	3.3	2.9	2.9	4.4	1.3	1.9	3.0	1.5	16.7	8.3	5.7
right foot	3.1	4.5	8.9	8.0	6.2	8.9	8.1	6.1	2.2	2.7	3.2	2.6	2.6	3.0	2.6	3.8	3.5	7.2
left foot	3.0	4.3	6.8	9.8	6.3	10.8	7.0	5.1	2.5	2.9	3.0	1.7	2.4	2.8	2.4	3.0	4.0	9.6

	SHOULDER AB/ADDUCTION			SHOULDER FLEXION/EXTENSION			HIP AB/ADDUCTION			COMBINED HIP/KNEE FLEXION/EXTENSION			COMBINED ARM AB/ADDUCTION			SAGITTAL/FRONTAL TRUNK LEANS		
JOINT DISP	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML
head	1.6	3.2	3.3	5.6	10.5	2.6	1.5	2.2	6.5	1.7	1.8	7.1	1.5	12.6	6.4	3.9	3.0	10.5
spine top	0.8	1.4	0.3	0.7	0.7	1.2	0.7	1.6	2.0	1.0	1.7	2.1	1.3	0.8	0.6	1.9	2.7	1.3
right shoulder	0.5	1.4	0.2	0.3	1.1	0.1	2.2	5.7	0.7	1.1	5.3	6.2	0.2	1.0	0.1	0.6	1.4	0.4
left shoulder	1.6	2.8	2.1	3.9	5.8	2.9	1.0	3.5	5.4	1.6	1.9	4.7	2.0	12.9	5.0	2.9	2.0	7.6
right elbow	0.8	1.8	0.3	0.7	0.9	1.2	0.8	1.1	1.7	1.1	1.7	1.8	1.2	0.9	0.6	1.8	2.6	1.4
left elbow	0.5	1.5	0.1	0.2	1.0	0.1	2.3	4.0	0.7	1.1	2.3	3.8	0.3	0.7	0.2	0.5	1.7	0.3
right wrist	1.2	3.5	0.4	0.8	1.3	0.5	0.8	1.5	2.6	2.0	1.2	1.7	1.7	1.2	0.2	2.5	2.4	4.4
left wrist	1.2	0.5	0.3	0.5	0.6	0.6	0.5	0.4	2.5	2.7	0.8	2.5	0.4	0.3	0.4	1.5	3.9	1.3
right hand	0.8	2.1	0.4	0.7	0.7	0.5	1.0	0.6	2.2	1.0	1.4	1.6	0.8	0.7	0.2	0.9	3.5	1.5
left hand	0.8	1.5	0.6	1.3	2.1	0.9	0.9	1.4	2.9	1.7	1.2	3.2	2.7	1.3	0.6	1.8	3.9	1.1
spine mid	1.6	2.3	0.5	1.1	1.1	0.7	1.4	1.3	2.5	1.1	1.5	2.3	2.6	1.2	0.4	1.8	3.7	1.0
spine base	2.9	2.9	2.1	4.7	3.7	2.0	1.6	2.2	3.0	2.0	2.9	2.5	2.2	6.8	3.6	4.4	5.3	1.4
right hip	1.7	2.1	1.5	2.0	4.8	1.1	1.3	3.6	1.9	2.1	2.6	2.1	2.9	7.2	2.9	2.0	5.0	2.2
left hip	1.0	0.4	0.5	1.0	0.2	0.4	2.2	0.5	4.3	1.7	0.6	8.4	0.5	0.2	0.3	3.7	5.0	6.4
right knee	1.1	2.0	0.4	1.5	3.3	0.3	3.8	3.8	1.0	1.6	2.8	2.1	0.8	0.9	0.2	1.5	3.3	0.7
left knee	1.1	1.7	0.3	0.6	1.2	0.4	2.9	3.6	1.1	1.4	1.9	1.7	0.7	0.7	0.2	1.1	2.7	0.5
right ankle	1.2	2.0	0.5	1.1	1.7	0.5	4.1	5.7	1.0	1.9	5.6	6.6	3.0	3.4	0.2	1.9	2.4	0.5
left ankle	0.8	1.1	0.3	1.0	1.2	0.2	4.6	2.7	1.0	1.9	2.8	4.2	0.9	1.5	0.1	1.7	1.5	0.6
right foot	2.0	2.6	3.2	6.7	12.2	3.5	1.7	2.8	7.5	1.4	2.0	7.5	1.4	13.7	7.9	4.0	4.7	11.0
left foot	1.4	2.9	2.2	4.7	6.1	2.3	1.6	4.1	6.7	1.7	2.7	5.6	2.0	13.8	5.2	2.5	3.4	8.0

Table 9: Average Errors in Joint angles measured across the same trials as Table 8 and grouped by movement.

	SIT TO STAND			TIMED UP AND GO			ALTERNATING BARBELL LUNGES			OVERHEAD SQUATS			MARCHING IN PLACE			TIME TO STABILIZATION		
JOINT ANGLES	KQ vs QS	K3 D vs QS	K3 D vs KQ	KQ vs QS	K3 D vs QS	K3 D vs KQ	KQ vs QS	K3 D vs QS	K3 D vs KQ	KQ vs QS	K3 D vs QS	K3 D vs KQ	KQ vs QS	K3 D vs QS	K3 D vs KQ	KQ vs QS	K3 D vs QS	K3 D vs KQ
arm flexion r	97	52	92	52	41	72	171	152	37	164	102	77	17	20	22	89	10	93
arm flexion l	88	12	77	66	25	58	140	184	91	97	126	83	38	17	30	68	12	61
arm adduction r	85	37	53	33	13	35	89	62	37	98	65	52	10	7	8	42	10	47
arm adduction l	52	6	50	37	8	40	141	108	34	81	61	38	9	6	7	51	6	49
elbow flexion r	37	54	32	31	43	17	61	94	54	31	50	32	16	16	9	37	52	18
elbow flexion l	33	70	37	34	54	24	67	90	71	32	24	37	19	22	10	33	65	34
hip flexion r	4	4	1	7	7	3	33	32	1	9	10	2	13	12	1	27	32	5
hip flexion l	9	8	3	11	11	4	28	34	9	13	12	4	18	17	3	13	8	5
hip adduction r	4	12	8	5	11	7	6	9	5	6	14	10	4	5	2	5	17	14
hip adduction l	5	15	16	7	10	14	8	9	12	9	21	18	6	4	8	5	5	5
knee flexion r	12	13	3	37	38	3	36	21	17	26	26	4	14	14	2	8	7	2
knee flexion l	16	14	5	28	29	6	18	18	9	28	23	9	16	14	4	9	9	3
knee adduction r	7	16	22	26	31	18	11	26	31	8	16	21	9	15	13	7	7	12
knee adduction l	25	13	29	32	32	22	21	18	25	19	14	28	22	13	13	22	5	22
lumbar extension	22	15	14	113	109	11	135	11	138	31	28	20	19	18	13	15	10	11

	SHOULDER AB/ADDUCTION			SHOULDER FLEXION/ EXTENSION			HIP AB/ADDUCTION			COMBINED HIP/KNEE FLEXION/ EXTENSION			COMBINED ARM AB/ADDUCTION			SAGITTAL/ FRONTAL TRUNK LEANS		
JOINT ANGLES	KQ vs QS	K3 D vs QS	K3 D vs KQ	KQ vs QS	K3 D vs QS	K3 D vs KQ	KQ vs QS	K3 D vs QS	K3 D vs KQ	KQ vs QS	K3 D vs QS	K3 D vs KQ	KQ vs QS	K3 D vs QS	K3 D vs KQ	KQ vs QS	K3 D vs QS	K3 D vs KQ
arm flexion r	32	50	30	14	16	22	20	16	30	42	16	49	127	97	52	109	81	56
arm flexion l	29	53	70	46	15	50	34	16	35	30	19	37	51	48	53	39	28	42
arm adduction r	10	10	7	18	19	8	13	8	11	21	10	14	82	33	52	71	60	29
arm adduction l	9	8	6	17	17	7	16	9	11	27	8	19	32	11	24	32	20	31
elbow flexion r	8	11	15	9	10	11	14	20	12	22	28	14	28	70	59	37	48	29
elbow flexion l	8	8	12	7	8	7	17	18	12	23	33	17	26	54	49	33	55	33
hip flexion r	10	11	1	11	12	1	14	15	1	11	12	1	9	10	1	15	16	13
hip flexion l	11	10	1	12	11	1	17	16	2	15	11	5	9	9	2	16	10	10
hip adduction r	5	5	0	5	5	1	4	4	2	5	5	1	5	5	0	6	4	4
hip adduction l	4	1	5	4	1	4	6	3	8	6	4	7	4	3	6	9	5	11
knee flexion r	22	22	1	23	24	2	15	16	2	14	13	2	20	20	1	26	25	9
knee flexion l	21	19	2	23	21	3	19	16	3	16	11	6	21	19	4	24	23	14
knee adduction r	14	6	19	13	5	17	11	5	15	16	7	15	14	6	16	14	10	19
knee adduction l	21	2	20	21	3	19	18	3	16	18	6	15	18	3	17	25	10	24
lumbar extension	13	6	7	18	7	10	10	7	8	14	4	14	23	6	20	44	19	37

Table 10: Normalized summary metric of accuracy in evaluating joint displacement accuracy in a Kinect-based MOCAP in comparison with a gold standard 3D Qualisys system.

	SIT TO STAND			TIMED UP AND GO			ALTERNATING BARBELL LUNGES			OVERHEAD SQUATS			MARCHING IN PLACE			TIME TO STABILIZATION		
JOINT DISP	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML
head	1.0	4.1	0.2	5.7	2.5	0.9	2.7	1.6	0.6	0.6	4.3	0.2	0.2	0.1	0.5	6.3	1.0	1.1
spine top	0.8	1.3	0.2	1.6	1.0	0.5	2.2	0.6	0.3	0.9	1.2	0.2	0.2	0.2	0.3	0.8	0.7	0.7
right shoulder	0.7	1.3	0.7	2.1	1.1	1.7	2.8	0.7	0.4	1.4	1.0	0.6	0.7	0.4	0.2	1.2	0.5	0.7
left shoulder	0.7	0.9	0.9	2.0	0.9	1.7	2.8	1.2	0.3	1.7	2.0	1.9	0.6	0.5	0.3	0.9	0.5	0.6
right elbow	0.9	0.8	1.9	1.9	1.3	2.7	3.0	3.0	0.5	4.0	1.2	1.8	1.2	0.9	0.9	0.5	0.7	1.4
left elbow	1.4	0.6	1.2	1.7	1.1	3.0	2.7	3.1	0.6	2.9	1.8	1.2	1.2	0.5	0.9	0.8	0.6	0.6
right wrist	1.0	1.0	2.2	2.2	1.3	3.3	2.0	2.0	0.7	0.8	1.0	0.3	0.8	0.6	1.1	0.8	0.8	2.4
left wrist	1.2	1.0	2.3	1.8	1.1	2.8	2.1	2.2	0.5	0.8	1.1	0.5	0.9	0.6	0.9	0.9	0.8	1.8
right hand	0.8	1.1	1.8	2.5	1.6	3.3	1.8	1.3	0.7	0.7	0.7	0.5	0.6	0.7	0.6	0.8	1.0	2.4
left hand	0.8	1.1	2.4	2.0	1.6	3.0	2.1	1.6	0.6	0.7	0.8	0.8	0.7	0.8	0.7	1.0	0.9	1.9
spine mid	0.7	1.5	0.2	1.7	1.2	0.7	1.4	0.7	0.3	0.8	1.2	0.3	0.3	0.2	0.3	0.7	0.8	0.6
spine base	9.4	2.2	0.5	6.7	1.7	1.2	2.2	1.6	0.7	2.8	2.9	0.8	0.7	0.7	0.5	5.1	0.4	1.0
right hip	1.8	0.6	0.9	1.3	1.1	0.9	1.9	1.2	0.6	2.1	2.2	1.7	0.5	0.6	0.4	1.0	0.3	0.7
left hip	2.0	0.9	0.9	1.5	1.0	0.8	2.0	0.9	0.6	1.8	2.2	1.9	0.7	0.3	0.5	0.9	0.4	0.7
right knee	1.0	1.2	0.2	1.9	1.9	0.8	3.1	1.5	0.4	1.0	1.6	0.6	0.6	0.6	0.2	1.3	1.0	0.6
left knee	0.8	1.1	0.2	1.9	2.9	0.8	2.6	1.3	0.9	1.0	1.4	0.5	0.6	0.6	0.3	0.8	1.6	0.4
right ankle	0.7	1.5	0.2	3.1	2.0	0.9	3.4	5.6	0.6	0.8	1.9	0.5	0.3	1.1	0.2	2.7	3.6	1.3
left ankle	1.0	1.6	0.3	3.0	3.3	0.8	3.1	4.5	1.3	1.0	2.0	1.1	0.4	1.2	0.2	0.8	0.9	0.3
right foot	1.4	1.7	0.3	4.3	9.2	1.3	4.3	9.5	1.0	2.3	1.8	1.4	0.6	0.9	0.5	5.8	2.8	2.0
left foot	1.1	2.1	0.5	4.2	9.3	1.1	4.7	5.6	1.7	1.6	1.7	1.2	0.7	1.0	0.6	1.3	1.3	0.8
Average per axis	1.5	1.4	0.9	2.6	2.4	1.6	2.6	2.5	0.7	1.5	1.7	0.9	0.6	0.6	0.5	1.7	1.0	1.1
Average	1.2			2.2			1.9			1.4			0.6			1.3		

	SHOULDER AB/ADDUCTION			SHOULDER FLEXION/EXTENSION			HIP AB/ADDUCTION			COMBINED HIP / KNEE FLEXION/EXTENSION			COMBINED ARM AB/ADDUCTION			SAGITTAL / FRONTAL TRUNK LEANS		
JOINT DISP	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML	AP	VT	ML
head	0.3	0.2	0.1	0.3	0.1	0.1	0.6	0.3	1.1	0.4	0.2	2.1	0.1	0.1	0.1	0.9	1.4	1.6
spine top	0.3	0.2	0.1	0.1	0.4	0.2	0.1	0.2	0.6	0.7	0.6	0.6	0.1	0.1	0.1	0.4	1.8	0.3
right shoulder	1.6	0.6	0.6	0.5	0.3	0.3	0.8	0.6	0.6	0.3	0.4	0.6	2.5	0.6	0.2	0.5	1.2	0.2
left shoulder	0.3	0.4	0.2	0.5	0.6	0.3	0.3	0.4	0.7	0.4	0.3	0.8	3.0	0.9	0.8	0.5	1.1	0.3
right elbow	0.6	0.5	0.4	0.5	1.2	0.3	0.5	0.9	0.5	0.6	0.7	0.5	1.9	2.1	0.8	0.5	1.3	0.6
left elbow	1.1	0.7	0.5	1.3	1.0	0.5	0.6	0.6	0.8	0.5	0.8	0.6	6.6	1.9	1.2	1.4	1.3	0.4
right wrist	0.5	0.7	0.5	1.0	1.5	0.7	0.3	0.9	1.4	0.5	0.5	1.2	0.7	3.5	1.3	0.9	0.5	2.0
left wrist	0.4	0.8	0.8	1.4	2.7	0.6	0.4	0.6	1.7	0.5	0.4	1.9	0.5	3.4	1.7	1.3	0.8	2.8
right hand	0.4	0.7	0.5	1.2	1.5	0.6	0.5	1.1	1.7	0.5	0.7	1.4	0.7	3.7	1.3	0.7	0.9	2.1
left hand	0.5	0.6	0.8	1.7	3.1	0.9	0.5	0.7	2.0	0.4	0.5	2.0	0.4	3.7	2.1	1.4	1.2	3.0
spine mid	0.6	1.0	0.2	0.2	0.7	0.3	0.4	0.5	0.6	0.3	0.7	0.4	0.2	0.7	0.1	0.4	7.3	1.4
spine base	0.5	4.5	0.4	1.0	1.0	0.3	0.4	2.7	0.7	1.2	0.6	0.4	1.5	0.7	0.1	0.7	0.8	1.6
right hip	1.6	0.5	0.2	0.3	0.3	2.0	0.4	0.3	0.4	0.3	0.5	0.5	0.5	0.5	0.5	0.5	3.7	0.4
left hip	1.4	0.4	0.1	0.5	0.2	3.7	0.3	0.6	0.5	0.3	0.5	0.5	0.5	0.4	0.5	0.5	10.4	0.4
right knee	0.6	0.5	0.1	0.2	0.5	0.2	0.8	1.0	0.3	0.4	0.5	0.4	0.4	0.8	0.1	0.3	8.0	0.1
left knee	0.5	0.6	0.2	0.7	2.6	0.1	1.0	1.1	0.3	0.4	0.7	0.6	0.6	0.6	0.4	0.4	9.7	0.2
right ankle	0.4	0.7	0.1	0.2	1.8	0.1	0.7	1.1	0.2	0.3	0.6	1.0	0.7	1.8	0.2	0.3	1.3	0.1
left ankle	0.3	0.6	0.2	0.2	0.8	0.2	0.7	1.6	0.2	0.3	1.4	1.6	0.1	0.8	0.1	0.3	0.8	0.1
right foot	0.5	2.8	0.3	0.9	4.5	0.3	2.6	0.7	0.3	0.9	0.7	1.1	3.3	4.4	0.3	1.2	1.0	0.4
left foot	0.8	1.6	0.4	1.3	1.6	0.3	4.8	1.6	0.3	0.8	1.5	1.8	5.7	3.8	0.1	4.1	2.4	0.4
Average per axis	0.7	0.9	0.3	0.7	1.3	0.6	0.8	0.9	0.7	0.5	0.6	1.0	1.5	1.7	0.6	0.9	2.8	0.9
Average	0.6			0.9			0.8			0.7			1.3			1.5		

Table 11: Normalized summary metric of accuracy in evaluating joint angle accuracy in a Kinect-based MOCAP in comparison with a gold standard 3D Qualisys system.

	SIT TO STAND			TIMED UP AND GO			ALTERNATING BARBELL LUNGES			OVERHEAD SQUATS			MARCHING IN PLACE			TIME TO STABILIZATION		
JOINT ANGLES	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ
arm flexion r	9.9	4.8	9.0	17.9	6.8	14.1	20.1	17.9	2.3	16.3	16.4	7.4	3.2	1.8	3.3	13.7	1.5	7.1
arm flexion l	5.3	1.9	9.1	18.8	1.7	19.8	10.5	18.6	6.7	8.7	14.7	19.6	3.7	1.3	3.8	6.9	1.6	10.1
arm adduction r	6.9	3.0	5.8	6.7	1.0	7.0	4.8	3.2	2.0	5.6	3.2	3.7	1.0	0.4	0.7	4.3	1.3	3.8
arm adduction l	5.9	0.6	4.4	9.1	0.8	7.7	9.7	7.1	2.1	6.1	4.0	2.9	0.8	0.3	0.6	5.4	0.7	3.9
elbow flexion r	3.1	4.2	3.8	1.4	3.0	1.4	3.5	11.8	6.9	1.7	3.2	3.2	1.4	2.3	2.0	2.0	6.7	2.1
elbow flexion l	2.9	5.4	5.4	1.8	8.6	3.0	4.6	15.2	13.4	1.8	2.2	3.3	1.9	3.3	1.2	2.2	6.0	3.9
hip flexion r	0.6	0.6	0.1	2.4	2.6	0.2	2.6	1.6	1.0	1.4	1.4	0.2	0.7	0.7	0.1	1.1	1.2	0.2
hip flexion l	0.7	0.7	0.3	2.2	2.0	0.5	1.3	1.4	0.4	1.5	1.4	0.4	1.0	0.8	0.2	0.7	0.5	0.3
hip adduction r	0.4	2.5	1.9	1.6	3.4	1.6	1.4	5.8	2.8	0.7	9.3	7.6	0.5	1.5	0.6	1.0	1.4	1.1
hip adduction l	3.0	1.8	4.2	11.3	3.0	2.6	3.6	2.8	2.7	3.8	4.1	3.1	1.4	2.3	1.5	2.2	0.8	3.0
knee flexion r	0.6	0.6	0.2	2.4	2.6	0.2	2.1	2.0	0.1	1.4	1.4	0.2	0.8	0.8	0.1	1.1	1.2	0.2
knee flexion l	0.7	0.7	0.2	2.2	2.0	0.4	1.4	1.8	0.5	1.5	1.4	0.4	1.0	0.9	0.1	0.7	0.5	0.2
knee adduction r	0.4	2.3	2.0	1.4	2.7	1.9	1.6	1.7	0.9	0.6	7.3	7.9	0.6	1.2	0.8	0.7	1.2	1.2
knee adduction l	3.2	1.8	5.2	4.9	2.8	3.9	3.0	1.8	1.1	3.0	3.6	3.7	1.1	3.1	1.6	2.0	1.0	4.1
lumbar extension	3.9	1.6	1.1	23.0	35.6	0.7	18.4	2.5	26.8	2.7	2.1	1.7	3.8	5.7	3.9	2.7	0.9	1.1
Average per axis	3.2	2.2	3.5	7.1	5.2	4.3	5.9	6.4	4.6	3.8	5.0	4.3	1.5	1.8	1.4	3.1	1.8	2.8
Average	2.9			5.6			5.6			4.4			1.5			2.6		

	SHOULDER AB/ADDUCTION			SHOULDER FLEXION/ EXTENSION			HIP AB/ADDUCTION			COMBINED HIP / KNEE FLEXION/ EXTENSION			COMBINED ARM AB/ADDUCTION			SAGITTAL / FRONTAL TRUNK LEANS		
JOINT ANGLES	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ	KQ vs QS	K3D vs QS	K3D vs KQ
arm flexion r	12.2	13.7	2.3	1.0	1.1	1.5	3.1	1.0	4.0	4.6	2.1	6.2	46.9	27.0	5.7	8.8	5.1	15.3
arm flexion l	11.9	21.3	8.2	3.6	1.4	4.6	6.0	1.5	3.9	8.8	3.3	8.9	29.2	39.0	8.4	57.8	6.7	14.5
arm adduction r	0.8	0.7	0.5	6.4	4.5	0.4	1.0	0.5	0.6	1.1	0.5	0.7	22.9	6.4	1.9	1.9	1.0	1.5
arm adduction l	0.7	0.7	0.7	3.5	3.1	0.4	1.3	0.5	0.9	1.5	0.5	1.0	2.9	0.8	1.7	2.5	1.0	2.2
elbow flexion r	1.6	14.6	85.0	1.9	7.8	3.6	1.6	2.5	1.5	1.6	2.8	1.8	2.1	18.5	10.6	1.9	4.3	2.8
elbow flexion l	3.3	3.4	3.2	0.8	5.5	1.0	1.4	2.2	1.6	2.0	4.9	3.5	2.0	2.8	9.8	2.4	12.1	6.7
hip flexion r	1.7	1.7	0.1	14.4	7.3	0.1	0.9	0.9	0.1	0.9	0.8	0.1	2.7	2.6	0.1	1.1	1.1	0.2
hip flexion l	1.4	1.7	0.2	6.5	2.6	0.4	0.9	0.7	0.2	1.5	0.7	0.4	1.6	1.8	0.2	1.5	1.0	0.5
hip adduction r	3.6	2.4	0.5	4.6	3.2	0.4	0.9	0.6	0.5	1.2	0.9	0.4	2.6	0.9	0.4	1.4	0.8	0.5
hip adduction l	11.7	0.4	1.2	3.8	0.4	0.8	4.2	0.2	1.8	2.4	0.5	0.8	2.1	0.4	0.9	3.7	0.7	9.2
knee flexion r	3.5	2.7	0.0	8.5	11.3	0.1	1.0	1.0	0.1	0.8	0.8	0.1	4.6	4.2	0.0	1.2	1.2	0.4
knee flexion l	1.8	1.8	0.1	5.2	2.1	0.2	0.8	0.7	0.1	1.4	0.7	0.3	4.7	3.7	0.1	1.4	1.0	0.6
knee adduction r	1.0	0.5	0.7	7.3	2.2	0.5	0.8	0.5	1.0	1.1	0.5	0.5	2.0	0.7	0.5	3.3	1.1	0.4
knee adduction l	2.4	0.7	1.1	3.0	0.5	0.8	3.6	0.2	1.5	3.3	0.8	0.8	2.6	0.4	0.8	3.0	0.6	3.3
lumbar extension	1.6	0.5	1.0	4.6	0.7	4.1	1.2	0.8	1.4	2.6	0.2	1.4	2.3	1.0	6.6	6.4	0.9	4.5
Average per axis	4.0	4.5	7.0	5.0	3.6	1.3	1.9	0.9	1.3	2.3	1.3	1.8	8.7	7.3	3.2	6.6	2.6	4.2
Average	5.1			3.3			1.4			1.8			6.4			4.4		

Effects of Concussion History on Center of Mass Motion During Modified Balance Error Scoring System (BESS) Testing in Women

Stephen M. Glass, Alessandro Napoli, Iyad Obeid, and Carole A. Tucker
Temple University, Philadelphia, PA, USA
e-mail: stephen.glass@temple.edu

Introduction

Concussion and mTBI continues to account for a substantial proportion of the military healthcare burden.[1] Concussion is commonly associated with balance deficits, which are likely related to impaired processing of sensory information.[2] These deficits, in turn, have the potential to impact activities of daily life, job performance, and risk of re-injury as balance is considered a foundational component of nearly all motor behaviors. Impaired balance may present clinically as increased postural sway, particularly in the absence of posture-relevant sensory information.

It has been reported in athletics and in the armed services that women are concussed at comparable rates,[3] experience more severe concussion-related symptoms and limitations, and have longer recovery times when compared with men.[4] Despite these discrepancies in epidemiology, data concerning female-specific neuromotor effects of concussion/mTBI are lacking.[5,6] This poses unnecessary additional risk to brain-injured women as clinical assessment and decision-making may disproportionately rely on knowledge that was developed through the observation of male research subjects.

Previous work has demonstrated sex differences in movement behaviors[7] as well as the relevance of these differences to injury and injury recovery. It is reasonable therefore to suspect that the postural control effects of concussion/mTBI in females are distinct and should be considered separately. The purpose of this research was to identify neuromotor deficits (specifically, balance) between service-age healthy women (CTRL) and women with a history of concussion/mTBI (mTBI). We hypothesized that that history of concussion would be associated with increased postural sway motion and velocity.

Methods

Thirty-one healthy women and 24 women with a history of concussion/mTBI were performed 3 20-second balance trials. Procedures for balance testing were based on the modified Balance Error Scoring System (BESS[8]) protocol and featured 1 trial each of Double Leg (DS), Single Leg (SS), and Tandem (TS) stance. Each testing condition required subjects to stand barefoot with eyes closed and hands-on-hips. DS was performed in bilateral stance, SS was performed standing on the non-dominant limb with a slight bend in the hip and knee of the non-standing leg, and TS was performed with feet inline heel-to-toe (non-dominant limb behind dominant limb). Participants were instructed 1) to remain as motionless as possible throughout a given trial, and 2) to return to the testing pose quickly should the testing position be lost.

Video, infrared, and depth data were acquired using a Microsoft Kinect 2.0™ at a variable frame rate (maximum 30 Hz, not under direct control of the user). These raw data are used to estimate 3D joint center time histories through an on-board classification algorithm. Joint center displacement histories were stored to a local machine running a custom C# software interface. This data was used then used offline to define segment end points, from which the 3D center of mass (COM) displacement time series was estimated using established methods.

Mean velocity (VEL) and standard deviation (SD) of displacement were used to summarize COM motion for each trial in the anteroposterior (AP) and mediolateral (ML) directions. Group performance (CTRL vs. mTBI) was then compared using one-sided Welch's independent samples t-tests for each outcome/stance combination. The *a priori* significance level was $\alpha = 0.05$.

Results

Significant group effects were observed (CTRL < mTBI) in the DS condition for all outcomes (**VEL_{ML}**: CTRL = 0.93 ± 0.72 cm/s, mTBI = 2.83 ± 2.77 cm/s, $t = -3.17_{(25.41)}$, $p < 0.01$; **VEL_{AP}**: CTRL = 1.55 ± 2.58 cm/s, mTBI = 3.21 ± 2.64 cm/s, $t = -2.07_{(40.44)}$, $p < 0.02$; **SD_{ML}**: CTRL = 0.49 ± 0.50 cm, mTBI = 1.79 ± 1.88 cm, $t = -3.18_{(25.50)}$, $p < 0.01$; **SD_{AP}**: CTRL = 0.89 ± 1.72 cm, mTBI = 1.89 ± 1.60 cm, $t = -1.97_{(39.14)}$, $p < 0.03$).

Significant group effects were observed (CTRL < mTBI) in TS condition for **VEL_{ML}** (CTRL = 2.51 ± 0.95 cm/s, mTBI = 4.55 ± 3.71 cm/s, $t = -2.38_{(21.50)}$, $p < 0.01$) and **SD_{ML}** (CTRL = 1.60 ± 0.78 cm, mTBI = 3.53 ± 3.40 cm, $t = -2.47_{(21.01)}$, $p < 0.01$).

No effects were observed in the SS condition ($p > 0.05$).

Conclusion

These data support the conclusion that, relative to healthy female controls, postural control in women with a history of concussion/mTBI is characterized by increased variability and velocity of the COM. The effects we report appear to be specific to the DS and TS conditions. The null finding in the SS condition was unexpected as, among the three, this stance is frequently associated with the poorest balance outcomes.[8] While this pattern of findings may be anomalous, it is also possible that single leg postural control is not effective in discriminating healthy controls from previously concussed individuals in this population. Previous research has demonstrated sex differences in baseline/uninjured balance outcomes wherein females were observed to have better postural control than males using similar experimental tasks.[9] This could suggest that factors related to sex contribute to the presently observed pattern of group effects (healthy vs. concussion/mTBI history), which may be unexpected owing to underrepresentation of females in prior work. If the effects we report are true, post brain-injury balance testing in women may be more appropriately limited to DS or TS conditions.

Further research is warranted to investigate sex-specific effects of concussion/mTBI on balance behaviors prospectively and in direct comparison with comparable male samples. Future work should also consider mechanisms that might account for differential baseline and post-injury behaviors between men and women, such anthropometrics and lower extremity alignment.

Acknowledgements
This work was supported by DoD/USAMRMC Award # W81XWH-15-1-0445.

Disclosure Statement
The authors have no conflicts of interest to disclose.

References

1. Taylor BC, Hagel Campbell E, Nugent S, et al. Three Year Trends in Veterans Health Administration Utilization and Costs after Traumatic Brain Injury Screening among Veterans with Mild Traumatic Brain Injury. *J Neurotrauma.* 2017;34(17):2567-2574.
2. Haran FJ, Slaboda JC, King LA, Wright WG, Houlihan D, Norris JN. Sensitivity of the balance error scoring system and the sensory organization test in the combat environment. *J Neurotrauma.* 2016;33(7):705-711.
3. Laker SR. Epidemiology of Concussion and Mild Traumatic Brain Injury. *PM&R.*3(10):S354-S358.
4. Berz K, Divine J, Foss KB, Heyl R, Ford KR, Myer GD. Sex-specific differences in the severity of symptoms and recovery rate following sports-related concussion in young athletes. *Phys Sportsmed.* 2013;41(2):58-63.
5. Costello JT, Bieuzen F, Bleakley CM. Where are all the female participants in Sports and Exercise Medicine research? *Eur J Sport Sci.* 2014;14(8):847-851.
6. Dye JL, Eskridge SL, Tepe V, Clouser MC, Galarneau M. Characterization and Comparison of Combat-Related Injuries in Women During OIF and OEF. *Military Medicine.* 2016;181(suppl_1):92-98.
7. Hewett TE, Myer GD, Ford KR, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. *The American journal of sports medicine.* 2005;33(4):492-501.
8. Riemann BL, Guskiewicz KM. Effects of mild head injury on postural stability as measured through clinical balance testing. *Journal of athletic training.* 2000;35(1):19.
9. Howell DR, Hanson E, Sugimoto D, Stracciolini A, Meehan III WP. Assessment of the postural stability of female and male athletes. *Clin J Sport Med.* 2017;27(5):444-449.

The Expanded Automatic Assessment of Postural Stability: the xAAPS

Alessandro Napoli, Stephen M. Glass, Carole A. Tucker, and Iyad Obeid
Temple University, Philadelphia, PA, USA
e-mail: a.napoli@temple.edu

INTRODUCTION

Mild Traumatic Brain Injury (mTBI) can lead to temporary or permanent impairment of cognitive, physical and physiological functions and it represents a contributing factor to a substantial number of deaths and permanent disability [1].

mTBI is best detected when the evaluation of possible exposure is carried out in the field, at the earliest opportunity. The Balance Error Scoring System (BESS), which is a brief and easily administered test of static balance, has been devised to detect balance deficits, arising from concussion and musculoskeletal injuries, in the field [2]. The BESS presents four main limitations: 1) it requires the presence of a trained (clinical) observer to score the test; 2) the test-to-test reliability can be biased by the manual scoring system; 3) A visually scored test can result in under-reporting some of the symptoms; 4) The BESS test only measures static posture. To address these limitations, we have developed the Automated Assessment of Postural Stability (AAPS) system, that is an easy to set-up, computerized and quantitative system for automatically administering and scoring the BESS test in a wide variety of non-clinic locations using inexpensive off-the-shelf devices [3], [4].

Furthermore, in order to provide a more comprehensive concussion evaluation tool we are developing the expanded AAPS (xAAPS) to introduce the evaluation of dynamic balance tasks. The xAAPS capability of evaluating coordinated dynamic movements will potentially provide more salient feedback for assessing concussion and suitability for return to duty than using static balance measures alone.

METHODS

The xAAPS system consists of two hardware components: a Windows laptop and a Microsoft Kinect 2.0 device, paired with a custom-developed Windows software application. The xAAPS software has been designed and developed to be user-friendly and to guide the operator through all the necessary steps to correctly administer the testing protocols. At the end of each trial, the xAAPS automatically evaluates, displays and stores the balance scores in under a minute. The xAAPS features a custom developed balance evaluation method based on computer classification algorithms that convert the subject's three-dimensional joint center positions (as derived from the Kinect sensor) into balance metrics. These metrics are equivalent to Functional Movement Screening (FMS) [5] scores assigned by an experienced observer. The FMS consists of seven movement patterns scored on a scale of 0-3 points, where 0 means pain and 3 a perfect execution. The current version of xAAPS focuses on continuous multi-repetition versions of the first three of the seven FMS assessments: Deep Squat (DS), Hurdle Step (HS) and In-line Lunge (ILL).

In order to validate the performance of the xAAPS scoring algorithm, we asked 26 young adults (12 male / 14 female) to perform the three FMS movements while their kinematic data were captured with the xAAPS system. To obtain reference data for comparison, video recordings of the movement tests were scored by an experienced observer. Those scores were then used as labels for the dataset when training the xAAPS classification algorithm.

More specifically, the xAAPS extracts 3D joint coordinates from the Kinect data stream. The Kinect generates these data at a variable sampling frequency, which is then resampled off-line to a constant rate of 30 fps. Subsequently, the resampled joint position time series are low-pass filtered with a fifth order Butterworth with cutoff frequency set at 2Hz to reduce measurement noise. The next step in the signal processing cascade is to extract features that can be successfully used to train a set of classification algorithms. For each trial, a total of 27 kinematic features are extracted and used to evaluate each trial quality. The extracted features range from commonly used kinematic metrics, such as range, mean and standard deviation of velocity, acceleration and jerk of the Center of Mass (COM) trajectory, to more complex features such as spectral power, coefficient of variation and continuous relative phase variability [6] of the COM. Dynamic Time Warping (DTW) distances of COM and Principal Component Analysis (PCA) of the joint 3D displacement time series [7] were also used as features. Next, we trained a set of gold-standard classification algorithms such as, Decision Trees, Support Vector Machine (SVM), k-Nearest Neighbors (k-NN) and Ensemble Bagged Trees. These classifiers' predictive performance was assessed using a 3-fold cross-validation approach. Finally, for each movement type, the optimal combination of features and classification algorithms were identified.

RESULTS

The xAAPS can successfully score the three FMS movements (HS, ILL and DS), with scoring performance well above random classification levels, that given the distribution of our sample population, are 57.1%, 42.3% and 67.7%, respectively. Specifically, the xAAPS displayed the best scoring performance for DS trials, using a SVM classifier with a cross-validated prediction accuracy of 92.3%. The HS assessment accuracy was 84.6% using a Decision Tree algorithm and finally accuracy of 69.2% was measured for ILL using an Ensemble Bagged Tree approach. Furthermore, qualitative analysis of kinematic data time series, indicates that the xAAPS lower performance for ILL trials is due to larger inaccuracies of the Kinect body tracking algorithm when detecting the lower-extremity movements for ILL motion.

CONCLUSION

Our laboratory has recently shown that Kinect 2.0™ data is suitable for instrumenting simple field-expedient clinical static postural stability tests such as the BESS [3]. With the present work, we present the xAAPS, an expanded version of the reliable and quantitative Automated Assessment of Postural Stability (AAPS). The statistical performance of the innovative xAAPS algorithms in predicting the human-assigned FMS scores for three movements, namely HS, ILL and DS, as performed by 26 subjects, shows that the xAAPS can be a valuable in-field expedient to evaluate dynamic balance, without the need of human scorers.

Furthermore, despite the current version of the system being optimized for three specific movements, the feature extraction and classification algorithms have been designed to be flexible, easily adjustable and re-trainable for the evaluation of further motion types and different clinical testing protocols.

REFERENCES

[1] P. Leo and M. McCrea, "Epidemiology," in *Translational Research in Traumatic Brain Injury*, D. Laskowitz and G. Grant, Eds. CRC Press/Taylor and Francis Group, 2016.

[2] D. R. Bell, K. M. Guskiewicz, M. A. Clark, and D. A. Padua, "Systematic review of the

balance error scoring system," *Sports Health*, vol. 3, no. 3. SAGE Publications, pp. 287–295, May-2011.
[3] A. Napoli, S. M. Glass, C. Tucker, and I. Obeid, "The Automated Assessment of Postural Stability: Balance Detection Algorithm," *Ann. Biomed. Eng.*, vol. 45, no. 12, pp. 2784–2793, 2017.
[4] A. Napoli, C. R. Ward, S. M. Glass, C. Tucker, and I. Obeid, "Automated Assessment of Postural Stability System," in *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, 2016, pp. 6090–6093.
[5] G. Cook, L. Burton, and B. Hoogenboom, "Pre-Participation Screening: The Use of Fundamental Movements As An Assessment of Function–Part 1," *North Am. J. Sport. Phys. Ther. NAJSPT*, vol. 1, no. 2, p. 62, 2006.
[6] R. H. Miller, S. A. Meardon, T. R. Derrick, and J. C. Gillette, "Continuous relative phase variability during an exhaustive run in runners with a history of iliotibial band syndrome," *J. Appl. Biomech.*, vol. 24, no. 3, pp. 262–270, 2008.
[7] P. A. Federolf, "A novel approach to study human posture control: 'Principal movements' obtained from a principal component analysis of kinematic marker data," *J. Biomech.*, vol. 49, no. 3, pp. 364–370, 2016.

ACKNOWLEDGMENTS
This work was supported by DoD/USAMRMC Award # W81XWH-15-1-0445.

DISCLOSURE STATEMENT
The authors have no conflicts of interest to disclose.

Journal of Applied Biomechanics, 2019, 35, 32-36
https://doi.org/10.1123/jab.2018-0056

Human Kinetics
ORIGINAL RESEARCH

Validity of an Automated Balance Error Scoring System

Stephen M. Glass,[1] Alessandro Napoli,[2] Elizabeth D. Thompson,[3] Iyad Obeid,[2] and Carole A. Tucker[2,3]
[1]Ohio State University Wexner Medical Center; [2]Temple University College of Engineering;
[3]Temple University College of Public Health

The balance error scoring system (BESS) is a human-scored, field-based balance test used in cases of suspected concussion. Recently developed instrumented alternatives to human scoring carry substantial advantages over traditional testing, but thus far report relatively abstract outcomes that may not be useful to clinicians or coaches. In contrast, the automated assessment of postural stability (AAPS) is a computerized system that tabulates error events in accordance with the original description of the BESS. This study compared AAPS and human-based BESS scores. A total of 25 healthy adults performed the modified BESS. Tests were scored twice each by 3 human raters and the computerized system. Interrater (between human) and intermethod (AAPS vs human) agreement (interclass correlation coefficient$_{2,1}$) were calculated alongside Bland–Altman limits of agreement. Interrater analyses were significant ($P < .01$) and demonstrated good to excellent agreement. Intermethod agreement analyses were significant ($P < .01$), with agreement ranging from poor to excellent. Computerized scores were equivalent across rating occasions. Limits of agreement ranges for AAPS versus the human average exceeded the average limits of agreement ranges between human raters. Coaches and clinicians may consider a system such as AAPS to automate balance testing while maintaining the familiarity of human-based scoring, although scores should not yet be considered interchangeable with those of a human rater.

Keywords: kinematics, clinical biomechanics, motion analysis, motor behavior, sports medicine

Static postural control assessed during quiet standing is commonly used as an indicator of injury and recovery status in cases of suspected concussion or mild traumatic brain injury.[1,2] Quantitative assessment of postural control is perhaps best achieved through laboratory-grade instrumentation and testing protocols,[3,4] but such methods are seldom used for wide-scale or field-based testing due to their prohibitive cost and lack of accessibility. Where laboratory methods are impractical, field-based tests may be used in their place.

The balance error scoring system (BESS)[5] and its variants (eg, modified BESS[6] [mBESS]) are among the most familiar field-based tests of balance. Although these methods address the feasibility limitations associated with laboratory-based measurement, the quantity and quality of data that can be captured by humans are limited and inherently subjective. Such limitations pose a logistical barrier to the effective use of BESS tests as screening tools, particularly in high-volume settings where human resources are limited.[7,8] An automated scoring system relying on portable, low-cost instrumentation could overcome these limitations while maintaining applicability for field and clinical use.

To date, instrumented alternatives to standard BESS or mBESS[9–14] testing have produced novel, signal-based outcomes representing scales, which are (1) unfamiliar to most end users and (2) not directly comparable to the more common BESS scores. To address these issues, we developed the automated assessment of postural stability (AAPS). This system automates BESS administration and scoring using a low-cost, mass-produced sensor to quantify 3-dimensional kinematic motion.[15] Importantly, the outcomes generated by our system are the same as those made familiar by nearly 20 years of clinical and scientific use of the BESS[5]—specifically, error counts per testing condition.

The purpose of this research was to study the agreement of the AAPS system with standard clinician-based scoring of the BESS. To characterize AAPS performance in reference to the criterion of scoring by trained observers, we examined interrater (human vs human) and intersystem (AAPS vs human) agreement. We hypothesized that (1) interrater agreement of (human) BESS scoring would range from good (interclass correlation coefficient [ICC] = .6 to .75, the upper portion of Fleiss' "fair to good" range) to excellent (ICC > .75) as reported in previous work[16] and (2) intermethod agreement (AAPS vs human BESS scoring) would similarly range from good to excellent. Finally, we additionally sought to determine whether AAPS could be considered interchangeable with humans as a method for scoring BESS tests.

Methods

Subjects

A total of 25 healthy participants (13 females: 25.57 [3.13] y, 167.64 [6.72] cm, 67.13 [17.21] kg; 12 males: 24.77 [3.81] y, 180.34 [6.22] cm, 84.87 [14.89] kg) were recruited to participate in this study. The protocol was approved by the Temple University institutional review board. All participants provided written, informed consent prior to participating.

Procedures

The mBESS—excluding the double-leg stance trials—was administered and scored in accordance with previously published procedures.[6,17] This mBESS was specifically chosen considering the lack of information gained from the double-leg standing conditions with this population.[6,18] Testing was performed in a quiet room

Glass is with the Department of Otolaryngology, The Ohio State University, Columbus, OH. Napoli, Obeid, and Tucker are with the Department of Electrical and Computer Engineering, Temple University College of Engineering, Philadelphia, PA. Thompson and Tucker are with the Department of Physical Therapy, Temple University College of Public Health, Philadelphia, PA. Glass (stephen.glass@osumc.edu) is corresponding author.

with minimal distractions. Participants completed each of the conditions in the following order: (1) single-leg stance/stable surface (Single-Stable), (2) tandem stance/stable surface (Tandem-Stable), (3) single-leg stance/foam surface (Single-Foam), and (4) tandem stance/foam surface (Tandem-Foam). All trials were 20 seconds in duration and were performed with hands on hips and eyes closed facing the sensor. Single-leg trials were performed on the nondominant limb, while the dominant limb was held in 20° of flexion at both the hip and knee. Tandem trials were performed with feet positioned in-line (heel-to-toe) with the nondominant limb placed in the back.

Participants were provided standard instructions to maintain each of the testing positions to the best of their ability for each trial and to reassume the testing position as quickly as possible should they commit an error. The standard following errors were used in scoring: (1) removing the hands from the hips, (2) opening the eyes, (3) step, stumble, or fall, (4) flexion/abduction of the hip in excess of 30°, (5) lifting/moving of the feet, and (6) remaining out of the testing position for more than 5 seconds.[5]

Instrumentation

All testing was conducted in view of a Microsoft Kinect 2.0™ (Microsoft Corp., Redmond, WA) gaming device secured to a tripod and placed 1.37 m above the ground and facing the participant at a distance of 3.00 m. The Kinect 2.0™ integrates depth, color, and infrared data streams to render 3-dimensional joint positions using Microsoft's proprietary algorithm at a maximum frame rate of 30 Hz. The exact sampling frequency cannot be controlled directly; however, data collection for a given trial was terminated if the frame rate dropped below 12 Hz.

The Kinect™ sensor was interfaced with a Dell Latitude PC (Dell, Inc. Round Rock, TX) (Windows 10 64-bit, 2.6-GHz Core i7 processor, 8-GB RAM, and 500-GB solid-state drive) through a customized C# application developed using the Microsoft Kinect SDK 2.0 libraries.[15] Each raw BESS trial video was 26 seconds in duration. A 20-second interval for testing was identified beginning with the time stamp of the first frame in which the participant met all the conditions of the required stance position (eg, eyes closed, hands on hips, vertical alignment, and proper foot placement). This frame, and the corresponding time, was identified manually for each video by S.M.G. Once these frames were identified, a truncated video and .csv time series file were generated for human and computer scoring, respectively.

Outcomes

The Kinect™ video data were evaluated twice by each of 3 human raters who were experienced with BESS test administration: 2 physical therapists (each with >10 y of experience, including BESS experience deriving from research, instruction, and musculoskeletal screening) and 1 personal trainer (>10 y of experience, with >5 y of experience scoring and analyzing the BESS for research). Prior to the initial video viewing, the 3 raters conferred to review the BESS scoring criteria. Rating sessions were separated by a minimum of 2 weeks and raters were blinded to (1) each other's scores and (2) their own scores from the first rating sessions.

All trials were also scored twice by the AAPS system. The AAPS-based automated error detection used Kinect® joint position and eye data, exported as .csv files, to determine whether movements related to any error categories exceeded a baseline threshold established over a 1-second pretest observation window. Thresholds were defined by 3-dimensional distances between the left and right centers of the wrist, elbow, ankle, and knee joints; frontal/sagittal trunk and hip angles; and left and right forefoot segments. (For further technical details regarding automated error tabulation, please see our previous work.[15])

Statistics

We calculated interrater (between human) and intermethod (human vs AAPS) correlation coefficients ($ICC_{2,1}$ for absolute agreement)[19] using the psych[20] package in R-3.4.1 (The R Foundation, Vienna, Austria). For intermethod ICC, the average score among the 3 raters was used for the human component of the analysis. ICCs were interpreted using the following guidelines modified from Fleiss[16]: .00 to .40 (poor), .40 to .59 (fair), .60 to .74 (good), and .75 to 1.00 (excellent). Finally, we calculated Bland–Altman 95% limits of agreement[21] (LOA) for AAPS versus the human average and compared it to the average of 3 between-human LOA ranges (1 vs 2, 2 vs 3, and 1 vs 3) to assess whether AAPS can be used interchangeably with human BESS rating. For the latter, human scores were averaged across day prior to determining LOA.

Results

With 2 exceptions, BESS score agreement between human raters was excellent (ICC > .75). The 2 exceptions were Single-Foam on day 1 and Tandem-Foam on day 2, for which agreement was good in both cases. All interrater agreement analyses reached the threshold for statistical significance (Table 1).

Table 1 Interrater and Intermethod ICCs

Condition	ICC (CI_{95})	F_{df}	P value	LOA
Between human day 1				
SS	.86 (.76 to .93)	$20.49_{24, 48}$	<.001*	−2.16 to 1.73
TS	.77 (.56 to .89)	$14.38_{23, 46}$	<.001*	−1.24 to 0.86
SF	.78 (.61 to .89)	$14.01_{23, 46}$	<.001*	−3.19 to 2.13
TF	.65 (.44 to .81)	$7.12_{24, 48}$	<.001*	−3.75 to 2.79
Total	.83 (.68 to .92)	$19.79_{22, 44}$	<.001*	−6.91 to 4.12
Between human day 2				
SS	.96 (.92 to .98)	$68.80_{24, 48}$	<.001*	−1.14 to 0.87
TS	.86 (.74 to .93)	$19.11_{23, 46}$	<.001*	−0.94 to 0.77
SF	.73 (.54 to .86)	$8.89_{23, 46}$	<.001*	−3.10 to 2.77
TF	.79 (.64 to .90)	$12.46_{23, 46}$	<.001*	−2.58 to 2.15
Total	.86 (.75 to .94)	$20.48_{21, 42}$	<.001*	−5.24 to 3.97
AAPS vs human average				
SS	.81 (.63 to .91)	$9.98_{24, 24}$	<.001*	−2.54 to 1.96
TS	.44 (.06 to .71)	$2.52_{24, 24}$	<.01*	−1.89 to 1.67
SF	.38 (−.02 to .68)	$2.21_{23, 23}$	<.03*	−4.21 to 4.94
TF	.72 (.47 to .87)	$6.11_{24, 24}$	<.001*	−3.10 to 2.74
Total	.74 (.47 to .88)	$6.48_{21, 21}$	<.001*	−7.23 to 6.77

Abbreviations: CI, confidence interval; ICC, interclass correlation coefficient; LOA, limits of agreement; SF, Single-Foam; SS, Single-Solid; TF, Tandem-Foam; TS, Tandem-Solid. Note: BESS score agreement between human raters for day 1 (top) and day 2 (middle). Agreement for AAPS versus human raters is shown on the bottom, where human rater scores are collapsed across day and rater.
*Indicates statistical significance.

Intermethod agreement was excellent for Single-Solid, good for Tandem-Foam and the total BESS score, fair for Tandem-Stable, and poor for Single-Foam. All intermethod agreement analyses reached the threshold for statistical significance (Table 1).

Finally, LOA ranges for AAPS versus the human average exceeded the average LOA ranges between human raters. Disparities in LOA ranges were greatest in the Single-Foam condition and least in the Tandem-Stable condition (Table 1). For visualization, we present Bland–Altman plots for AAPS versus the human average alongside the 2 most consistent human raters in our data set for the stable conditions (Figure 1) and foam conditions (Figure 2).

Discussion

The purpose of this research was to study the agreement of an automated BESS scoring system (AAPS) with the existing standard of human observation. We found similar interrater agreement for BESS item and total scores as has been reported previously.[5,22] By comparison, our intermethod analyses suggest that agreement between human-derived and AAPS-derived BESS scores was lower than that between human raters, with 2 of the conditions (Single-Foam and Tandem-Stable) failing to reach the hypothesized "good" level of AAPS-Human agreement (ICC). Finally, although we are encouraged by these initial findings and the qualitative similarity of the aggregate data (Supplementary Data [available online]), 95% LOA ranges suggest further work is required before our system can be considered interchangeable with a human rater for purposes of registering BESS error behaviors in clinical settings.

Efficient collection of high-quality data has been identified as a major concern for organizations conducting large-scale medical or preparticipation exams.[7,8,23] Although not a replacement for a trained human BESS scorer, AAPS is 100% consistent in arriving at error scores and catalogs behavior in far greater detail than is possible through human observation. Thus, a system such as ours could be used to automate balance testing procedures in large-scale organizations, potentially increasing testing efficiency, while at the same time addressing limitations related to interrater or intrarater measurement consistency.[24]

In addition to automating the BESS, the present system does so with less abstraction than other methods in which kinematic features are used as a correlate of BESS scores.[10,13] That is, AAPS quantifies balance performance in the original BESS unit of error counts as opposed to a novel metric. The intention with this strategy was to present the clinician or researcher with information that is not only intuitively meaningful but which is readily compared with previously published normative data.[25] Creating such a system, however, requires finding a balance between the relative strengths of human and machine observation.

The BESS was originally designed to provide a tool that could be used by humans to assess balance in nonlaboratory settings.[5] As a human-friendly tool, there are some areas in which the scoring criteria favor human intuition. For example, humans may be able to separate a single prolonged error from multiple distinct errors using intuition that cannot currently be coded. This bias in favor of human intuition comes at the cost of limitations on human focus and multitasking. The BESS scoring criteria do not require the

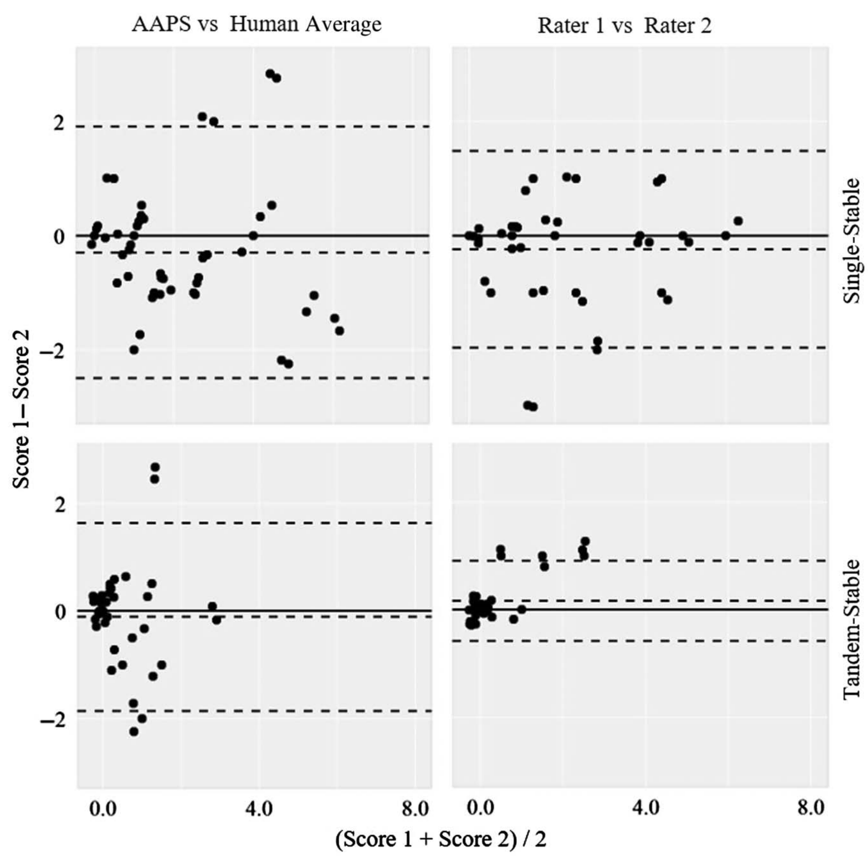


Figure 1 — Bland–Altman limits of agreement plots in stable stance conditions for (1) AAPS versus the human average collapsed across day and rater (left), and (2) for rater 1 versus rater 2 (right), which represents the highest level of between-human agreement observed in our data set. Here, rater 1 and rater 2 scores are averaged across day. AAPS indicates automated assessment of postural stability.

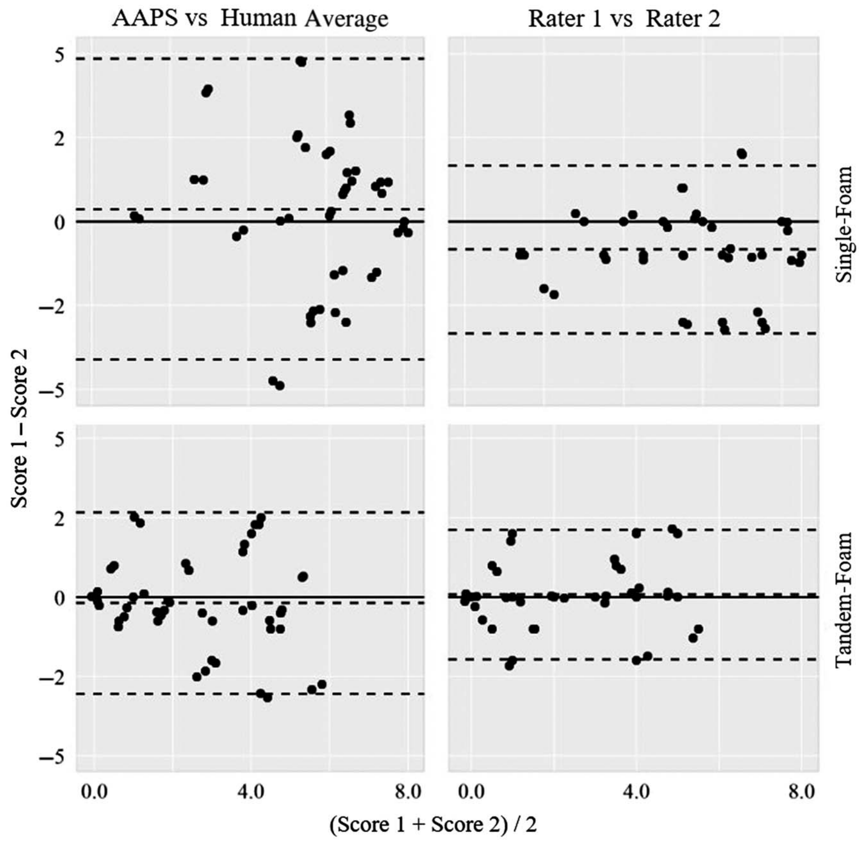


Figure 2 — Bland–Altman limits of agreement plots in Foam stance conditions for (1) AAPS versus the human average collapsed across day and rater (left), and (2) for rater 1 versus rater 2 (right), which represents the highest level of between-human agreement observed in our data set. Here, rater 1 and rater 2 scores are averaged across day. AAPS indicates automated assessment of postural stability.

observer to log simultaneous error events, error event types, or times of error occurrence. Arguably, this is not because such information is unimportant, but rather because it is not feasibly monitored by a human. So, although there may be limits to how precisely an automated system can distinguish errors from non-errors, the user has access to information that is more consistent and more descriptive (eg, error type, time of occurrence, and magnitude) than human judgment.

The results of this study are limited to the sample of balance participants used for our analyses. Further work will be required to cross-validate the error detection algorithm with different samples. Another limitation concerns the content of the errors detected. Because BESS trial scores from human raters present only the total error count, we cannot determine the extent to which our algorithm registers the same error events observed by a human rater. This limitation is currently being addressed with the development of a real-time annotation program, which will aid human raters in producing labeled trials.

In conclusion, the AAPS system provides a low-cost method of quantifying balance performance, requiring only a laptop and a Kinect™ for administration. (Future releases will support use with other RGB-D sensors.) It is currently available for research use. AAPS can be applied to automate balance test administration in high-volume screening settings or others in which human rating is inefficient and/or cost-prohibitive, although error scores from the current algorithm are not yet interchangeable with BESS scores from a trained human rater. Forthcoming work from our laboratory will describe the system's additional features, as well as robustness to sources of noise that might be encountered during field use.

Acknowledgment

The authors have no conflicts of interest to disclose.

References

1. Guskiewicz KM. Postural stability assessment following concussion: one piece of the puzzle. *Clin J Sport Med*. 2001;11(3):182–189. PubMed ID: 11495323 doi:10.1097/00042752-200107000-00009
2. Guskiewicz KM, Ross SE, Marshall SW. Postural stability and neuropsychological deficits after concussion in collegiate athletes. *J Athl Train*. 2001;36(3):263–273. PubMed ID: 12937495
3. Peterka RJ, Black FO. Age-related changes in human posture control: sensory organization tests. *J Vestib Res*. 1989;1(1):73–85.
4. Nashner LM, Black FO, Wall C 3rd. Adaptation to altered support and visual conditions during stance: patients with vestibular deficits. *J Neurosci*. 1982;2(5):536–544. PubMed ID: 6978930 doi:10.1523/JNEUROSCI.02-05-00536.1982
5. Riemann BL, Guskiewicz KM, Shields EW. Relationship between clinical and forceplate measures of postural stability. *J Sport Rehabil*. 1999;8(2):71–82. doi:10.1123/jsr.8.2.71
6. Hunt TN, Ferrara MS, Bornstein RA, Baumgartner TA. The reliability of the modified balance error scoring system. *Clin J Sport*

Med. 2009;19(6):471–475. PubMed ID: 19898074 doi:10.1097/JSM.0b013e3181c12c7b

7. Teyhen D, Bergeron MF, Deuster P, et al. Consortium for health and military performance and American College of Sports Medicine Summit: utility of functional movement assessment in identifying musculoskeletal injury risk. *Curr Sports Med Rep*. 2014;13(1):52–63. PubMed ID: 24412892 doi:10.1249/JSR.0000000000000023
8. Teyhen DS, Shaffer SW, Umlauf JA, et al. Automation to improve efficiency of field expedient injury prediction screening. *J Strength Cond Res*. 2012;26(suppl 2):S61–S72. doi:10.1519/JSC.0b013e31825d80e6
9. Goble DJ, Manyak KA, Abdenour TE, Rauh MJ, Baweja HS. An initial evaluation of the btracks balance plate and sports balance software for concussion diagnosis. *Int J Sports Phys Ther*. 2016;11(2):149–155. PubMed ID: 27104048
10. Alberts JL, Thota A, Hirsch J, et al. Quantification of the balance error scoring system with mobile technology. *Med Sci Sports Exerc*. 2015;47(10):2233–2240. PubMed ID: 26378948 doi:10.1249/MSS.0000000000000656
11. Alsalaheen BA, Haines J, Yorke A, Stockdale K, Broglio SP. Reliability and concurrent validity of instrumented balance error scoring system using a portable force plate system. *Phys Sportsmed*. 2015;43(3):221–226. PubMed ID: 26109242 doi:10.1080/00913847.2015.1040717
12. Amick RZ, Jansen SD, Chaparro A, Hakansson NA, Patterson JA, Jorgensen MJ. Comparison of the SWAY balance mobile application to the abbreviated balance error scoring system. *Athl Train Sports Health Care*. 2015;7(3):89–96. doi:10.3928/19425864-20150422-04
13. Brown HJ, Siegmund GP, Guskiewicz KM, Van Den Doel K, Cretu E, Blouin JS. Development and validation of an objective balance error scoring system. *Med Sci Sports Exerc*. 2014;46(8):1610–1616. PubMed ID: 24500539 doi:10.1249/MSS.0000000000000263
14. Alberts JL, Hirsch JR, Koop MM, et al. Using accelerometer and gyroscopic measures to quantify postural stability. *J Athl Train*. 2015;50(6):578–588. PubMed ID: 25844853 doi:10.4085/1062-6050-50.2.01
15. Napoli A, Glass SM, Tucker C, Obeid I. The automated assessment of postural stability: balance detection algorithm. *Ann Biomed Eng*. 2017;45(12):2784–2793. PubMed ID: 28856486 doi:10.1007/s10439-017-1911-8
16. Fleiss JL. *The Design and Analysis of Clinical Experiments*. New York, NY: Wiley; 1999.
17. Riemann BL, Guskiewicz KM. Effects of mild head injury on postural stability as measured through clinical balance testing. *J Athl Train*. 2000;35(1):19–25. PubMed ID: 16558603
18. Valovich McLeod TC, Perrin DH, Guskiewicz KM, Shultz SJ, Diamond R, Gansneder BM. Serial administration of clinical concussion assessments and learning effects in healthy young athletes. *Clin J Sport Med*. 2004;14(5):287–295. PubMed ID: 15377968 doi:10.1097/00042752-200409000-00007
19. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420–428. PubMed ID: 18839484 doi:10.1037/0033-2909.86.2.420
20. Revelle W. *Procedures for Personality and Psychological Research*. Evanston, IL: Northwestern University; 2015.
21. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8(2):135–160. PubMed ID: 10501650 doi:10.1177/096228029900800204
22. Valovich McLeod TC, Barr WB, McCrea M, Guskiewicz KM. Psychometric and measurement properties of concussion assessment tools in youth sports. *J Athl Train*. 2006;41(4):399–408. PubMed ID: 17273465
23. Teyhen DS, Shaffer SW, Butler RJ, et al. Application of athletic movement tests that predict injury risk in a military population: development of normative data. *Mil Med*. 2016;181(10):1324–1334. PubMed ID: 27753571 doi:10.7205/MILMED-D-15-00297
24. Finnoff JT, Peterson VJ, Hollman JH, Smith J. Intrarater and interrater reliability of the Balance Error Scoring System (BESS). *PM R*. 2009;1(1):50–54. doi:10.1016/j.pmrj.2008.06.002
25. Iverson GL, Koehle MS. Normative data for the balance error scoring system in adults. *Rehabil Res Pract*. 2013;2013:846418. PubMed ID: 23577257