

ORIGINAL ARTICLE

Reliability and concurrent validity of TRAZER compared to 3-dimensional motion capture

Jennifer A. Hogg^{1*}, Lynette M. Carlson¹, Abigail J. Rogers², Mason W. Briles³, Shellie N. Acocello¹, Gary B. Wilkerson¹

1. Department of Health and Human Performance, The University of Tennessee Chattanooga, Chattanooga, TN, USA

2. Sports Medicine and Performance Department, Michigan State University, East Lansing, MI, USA

3. Emory Sports Medicine Center, Emory Healthcare, Atlanta, GA, USA

*Corresponding author

Jennifer A. Hogg

Department of Health and Human Performance, The University of Tennessee Chattanooga, Chattanooga, TN, USA

Email: jennifer-hogg@utc.edu

Article information:

Received: September 9, 2020

Revised: November 13, 2020

Accepted: December 13, 2020

Abstract

Background. Efficient neural processing of visuospatial and proprioceptive input appears to be crucial for avoidance of sport injury. As such, clinically-feasible tests are needed to identify deficiencies found by advanced neuroimaging and electrophysiological tests. Three-dimensional motion capture in a laboratory setting is currently the gold standard for measurement of human movement parameters but is costly and requires extensive training. Non-immersive virtual reality systems with body motion tracking, such as TRAZER, may provide a clinically-feasible and portable means of acquiring similar variables. Test-retest reliability and concurrent validity of these systems is currently lacking.

Aim. To assess the concurrent validity of the TRAZER single-camera system with 3D motion capture system and to assess the test-retest reliability of TRAZER's whole-body reactive agility metrics.

Methods. Participants: For validity, 13 healthy individuals (24.8 ±3.1 yrs, 170.0 ±7.7 cm, 70.0 ±14.2 kg); for reliability, 18 healthy individuals (23.3 ±2.5 yrs, 168.2 ±11.2 cm, 78.2 ±17.8 kg). Design: Validity was a single-session cross-sectional study. Reliability was a three consecutive day test-retest study.

Setting: Controlled laboratory study. Intervention: Assessments utilized randomized movements in eight directions for forty total repetitions as designated by the TRAZER system. TRAZER protocol was simultaneously tracked by Vicon Motion Capture and the TRAZER system. Reliability data were captured on three consecutive days by the TRAZER system. Main Outcome Measures: Maximum acceleration, maximum velocity, and total distance were recorded for validation. In addition to these measures, maximum deceleration, average velocity, average acceleration, average deceleration, and average reaction time were collected for reliability.

Results. Overall, a lack of agreement exists between maximum outputs for TRAZER and 3D motion capture (velocity $r = .808$, acceleration $r = -.090$) but total distance correlation was high ($r = .961$). ICC values between days 1-2-3 for average measures were high (average velocity =.847, average acceleration=.919, average deceleration =.948) with the exception of average reaction time being fair (ICC =.536). ICCs for maximum measures showed a much smaller correlation between days (velocity =.654, acceleration =.171, deceleration =.416).

Conclusions. Even though there is a lack of strong concurrent validity between measures obtained from TRAZER and 3D motion capture systems, there is strong test-retest reliability of the TRAZER system. The applicability of these findings makes TRAZER clinically relevant in scenarios requiring pre and post-testing for return to play decisions, or monitoring of a training regimen where demonstration of validation to a gold standard measurement is not relevant.

Relevance for patients. When test-retest capability is desired, such as in return-to-play protocols following an injury, Trazer is a reliable option.

Key Words: Kinect, virtual reality, functional performance, reactive agility

1. INTRODUCTION

Functional movement assessments are used by sports medicine clinicians to identify movement dysfunction for purposes of injury prevention, evaluation, rehabilitation, and return to activity decisions. Movement screenings, along with proper corrective interventions, are essential components of injury prevention¹. Prevention strategies are critical to enhance physical performance, optimize health, minimize health care expenses, and avoid chronic dysfunction and disability². Portable and convenient acquisition of accurate movement assessment data is necessary for effective injury risk identification and ultimate prevention of sport-related injury.

Lab-based three-dimensional (3D) motion capture systems are the gold standard in functional movement analysis, with reported excellent reliability ($ICC_{3,k} > .93$)³ and validity (± 0.198 mm) (Vicon.com). However, they have limited clinical application due to financial, spatial, and temporal costs⁴ and expertise needed to collect and interpret data. Other means of assessing quality of human movement have been developed and studied extensively, including the Functional Movement Screen⁵, Star Excursion Balance Test⁶, and the Landing Error Scoring System⁷. These tools, however, lack ecological validity, as their required movements are anticipated, whereas responding to unanticipated events and simultaneous performance of cognitive and motor tasks are typically required during athletic activities. Unanticipated events can lead to sensory prediction errors and improper muscle co-contractions, potentially resulting in musculoskeletal injury. Identifying deficits in the simultaneous processing of environmental stimuli and task constraints and in the ability to preplan correct motor sequences (feed-forward) is important for injury prevention⁸⁻¹⁰. Virtual reality systems that track body movements in response to visual stimuli may be valuable for assessing neuromechanical responsiveness, or the ability to optimally integrate neurocognitive and neuromuscular processes¹¹, and integrated perception-motor neural processes¹², which appear to be crucial for preventing athletic injuries¹¹.

Virtual reality is an emerging multidisciplinary tool¹³ and is typically categorized as either immersive or non-immersive^{14,15}. An immersive virtual reality system is one in which the user dons a head-mounted display to interact with their environment, while non-immersive virtual reality is interacted with through a television or computer monitor. TRAZER (Traq Global Ltd, Westlake, OH) is a commercially available non-immersive virtual reality system that utilizes a camera and software gesture recognition. It is marketed as a tool for prevention, rehabilitation, balance training, fall prevention, and concussion management. Its novelty lay in its output metrics. Namely, the measures of *total distance*, *maximum velocity*, *maximum acceleration*, and *reaction time* are indicative of overall functional performance and as such, may be useful for clinicians and coaches aiming to determine one's ability to return to sport or their improvement in overall function^{11,16}.

Briefly, the TRAZER system is a non-immersive virtual reality system that utilizes a Kinect camera. The system employs an infrared camera to create a two-dimensional (2D) representation on a video monitor whereby a participant responds to visual stimuli and is recorded by a Kinect camera. All data captured by the TRAZER system are processed by embedded proprietary algorithms, making it unique to a stand-alone Kinect camera. The Kinect camera has been evaluated for reliability and validity compared to 3D motion capture systems during postural control and balance tasks (Clark et al., 2012, 2015; Eltoukhy, Kuenze, Oh, & Signorile, 2018), dynamic side-cut maneuvers (Eltoukhy et al., 2019), squatting (Schmitz et al., 2015), and single leg squatting (Ressman, Rasmussen-Barr, Grooten, 2020). Studies report good to excellent concurrent validity for kinematic angles during a postural balance task (Pearson's $r > 0.90$)¹⁸, (ICC > 0.75) (Eltoukhy, Kuenze, Oh, & Signorile, 2018), side-cutting maneuvers (absolute agreement ICC range = .77-.99) (Eltoukhy et al., 2019), and squatting (Pearson's $r > 0.55$)²¹. Test-retest reliability has variable results, from excellent reliability (ICC $> .90$) reported by Schmitz et al., 2015 to modest reliability (ICC ≥ 0.70) reported by Clark et al., 2015. Although Kinect systems have been assessed for accuracy, the accuracy of proprietary algorithms overlaid on Kinect, such as those provided by TRAZER, has yet to be determined.

One study (Nyman, 2017; white paper) compared TRAZER outputs with 3D Vicon motion capture system during balance and lateral shuffling tests. Three separate and successive trials of four movement patterns were completed by one female and one male participant during simultaneous TRAZER and Vicon captures. The two systems yielded highly correlated results for joint coordinate positions (ICC range = .75-.99), joint marker distances (ICC range = .93-1.0), marker velocities (ICC range = .93-1.0), joint marker accelerations (ICC range = .83-.96) and joint marker decelerations (ICC range = .91-.97) (Nyman, 2017). However, due to the small sample size (n=2) and single day data collection, there is a need for further investigation.

Therefore, the present study assessed the test-retest reliability of the TRAZER system and its concurrent validity against a 3D Vicon motion capture system. We hypothesized that TRAZER would display moderate test-retest reliability and good to excellent concurrent validity. This study represents a logical step towards the ability to make an informed decision about the fidelity of TRAZER's measurement of functional movement and its use in the clinic and the laboratory. If TRAZER demonstrates agreement against the gold standard 3D motion capture, the utilization of a more clinically-feasible alternative may expand clinicians' resources. The ability to accurately measure movement via a commercially available system has the potential for widespread adoption. Reliability is important in the clinic as well as in the laboratory, as reproducing consistent outcomes allow for accurate comparisons of patient progress.

2. METHODS

2.1 Participants

The data acquired to assess reliability and concurrent validity were obtained from separate cohorts. The reliability cohort consisted of a convenience sample of 18 healthy individuals (11 female, 23.3 \pm 2.5 yrs, 168.2 \pm 11.2 cm, 78.2 \pm 17.8 kg). The validity cohort consisted of a convenience sample of 13 healthy individuals (8 female, 24.8 \pm 3.1 yrs, 170.0 \pm 7.7 cm, 70.0 \pm 14.2 kg). Participants were included

if they were between the ages of 18 and 35. Participants were excluded if they had sustained a lower extremity injury in the previous six months or if they had a history of vestibular disorders, balance disorders, or cardiac conditions or limitations. History of concussion was not considered exclusionary. The reliability and validity components of this study were approved by the university's Institutional Review Board.

2.2 Instrumentation

TRAZER uses a depth-sensing Microsoft Kinect camera to create a three dimensional map of a 1.75 x 1.75 meter capture area. Once within the field of view, a simulated person (avatar) appears on the monitor and mirrors the participant as they respond to visual targets randomly appearing on the perimeter of the capture area. Anatomical landmarks (e.g., joint centers) are determined with a randomized decision forest algorithm with a reported one millisecond latency²²⁻²⁴. Prior to the protocol, each participant stands in the center of the capture area facing the TRAZER television screen for a brief (~5 second) calibration, during which the Kinect camera recognizes and identifies the participant. Following calibration, a visual target randomly appears at one of eight possible locations on the perimeter of the capture area (forward, backward, left, right, forward left diagonal, forward right diagonal, backward left diagonal, or backward right diagonal) (Figure 1). Once the indicator appears, the participant moves as quickly as possible to the location. Once TRAZER detects the participant in the correct location, the indicator disappears, and the participant returns to the start position to prepare for the next repetition. The protocol consists of forty repetitions (five at each of the eight possible locations), each repetition entailing seven to eight feet of travel, with a complete trial taking approximately three minutes to complete. TRAZER does not output raw data, but provides pre-defined performance metrics such as *reaction time*, *average/maximum velocity*, *average/maximum acceleration and deceleration*, and *total distance* traveled (see Nyman, 2017a and <https://trazer.com/science> for comprehensive background information on TRAZER).

2.3 Procedures

Following written informed consent, height and weight were obtained from each participant. To measure reliability, operationally defined as test-retest consistency, participants were not digitized with any 3D markers or instrumentation. They attended data collection sessions on three consecutive days at similar times each day, during which each participant performed the TRAZER protocol. Based on pilot testing, three consecutive days were used to account for a possible learning effect on the first day.

For validity, operationally defined as absolute agreement, each participant attended a single session in which a reflective marker placed on the S2 spinous process was digitized with Vicon Nexus (Vicon, Oxford Metrics, London, England) software. A static capture was recorded. Participants then performed the TRAZER protocol while being tracked concurrently with the TRAZER system sampling at 30 Hz and an 8-camera Vicon 3D Motion Capture system sampling at 60 Hz, consistent with previous validity studies that have used higher sampling rates for 3D motion capture^{4,18,22}.

2.4 Data Handling

TRAZER obtains its metrics by tracking the “base of the spine”, which was operationalized as the digitized S2 spinous process marker. To obtain measurements from three-dimensional data, raw coordinates for the S2 marker were exported from Vicon Nexus. Data were processed in R (R Core Team (2013). R: A language and environment for statistical computing) using the packages ‘signal,’²⁵ ‘imputeTS,’²⁶ ‘zoo,’²⁷ and ‘purrr’²⁸. Raw marker coordinate data were interpolated and filtered with a 12Hz low-pass 4th order Butterworth filter. The cutoff of 12 Hz was chosen after a residual analysis indicated 13 Hz to be the optimum cut point to maximize the signal to noise ratio. Total distance, maximum velocity, maximum acceleration, average velocity, average acceleration, average deceleration, and reaction time were the outcomes of interest. However, due to the proprietary nature of TRAZER calculations, only total distance, maximum velocity, and maximum acceleration were able to be confidently computed independently from TRAZER; thus, for validity, only these three variables were used. They were calculated as follows:

$$\text{Total distance (m)} = \sum |p_i - p_{i-1}|, \text{ where } p = \text{position and } i = \text{capture frame}$$

Maximum velocity (m/s) = $\frac{\max(p_i - p_{i-1})}{t}$, where p = position, i = capture frame, and t = time

Maximum acceleration (m/s²) = $\frac{\max(v_i - v_{i-1})}{t}$, where v = velocity, i = capture frame, and t = time

2.5 Statistics

To assess reliability, a repeated-measures MANOVA was used to determine differences between days. Interclass correlation coefficients (ICC_(3,3)) and standard errors of the measurement (SEM) were computed for total distance, maximum velocity, maximum acceleration, maximum deceleration, average velocity, average acceleration, average deceleration, and reaction time. To assess validity, a repeated-measures MANOVA was conducted to compare means between the TRAZER and Vicon systems for total distance, maximum velocity, and maximum acceleration. Paired t-tests were conducted as post-hoc follow ups. Pearson correlations, ICC_(2,2), and SEM were also computed to determine levels of absolute agreement. ICCs were computed according to McGraw & Wong²⁹ and interpreted as follows: less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.9 were considered poor, moderate, good, and excellent, respectively³⁰. Bland-Altman plots were constructed with the R package 'BlandAltmanLeh'³¹ to visually inspect validity and reliability data. Alpha was set at 0.05 for all statistical tests. Analyses were conducted in SPSS (IBM Corp, Armonk, NY).

3. RESULTS

3.1 Reliability

A significant deviation from normality was evident for day one's reaction time (Shapiro-Wilk $p = .002$), but neither logarithmic nor square root transformation provided any substantial improvement in distribution normality. A repeated measures MANOVA including all eight dependent variables demonstrated a significant difference among trials (Wilks' Lambda = 0.36; $F_{15,54} = 2.23$; $p = .015$). With the exception of the three variables of average reaction time, average deceleration, and maximum acceleration, ICC_{S(3,3)} across all three testing sessions were moderate to excellent, ranging from .65 to .95.

Test-retest consistency among trials, Shapiro-Wilk test of normality result for each trial, and univariable repeated measure analysis of variance results for differences among trials are presented in Table 1.

3.2 Validity

Vicon distributions for total distance and maximum velocity deviated significantly from normality, but neither logarithmic nor square root transformation provided any substantial improvements. Thus, the results of non-parametric correlations (Spearman's rho) were also calculated and are presented in Table 2. A repeated measures MANOVA including all three dependent variables demonstrated a significant difference between systems (Wilks' Lambda = 0.03; $F_{3,10} = 94.00$; $p < .001$). Means, standard deviations, and Pearson r values for total distance, maximum velocity, and maximum acceleration obtained from TRAZER and Vicon are presented in Table 2. ICCs (SEM) for total distance, maximum velocity, and maximum acceleration were 0.79 (11.1m), 0.08 (0.5 m/s), and 0.01 (0.5 m/s²), respectively. Bland Altman plots are presented as visual representations of the absolute agreement between TRAZER and Vicon (Figures 2-4).

4. DISCUSSION

We hypothesized that TRAZER would display moderate test-retest reliability and good to excellent concurrent validity. Our hypothesis was partially supported as evidenced by good to excellent reliability, but poor concurrent validity. To our knowledge, such a comparison has not been previously performed. Although we observed a lack of concurrent validity between TRAZER and 3D motion capture, there was strong test-retest reliability of the TRAZER system particularly with regards to average acceleration, average velocity, and average deceleration.

Test-retest reliability of TRAZER was particularly strong with averaged measures such as average acceleration (ICC=0.919), average velocity (ICC=0.847), and average deceleration (ICC=0.948) in contrast to reliability of peak measures of acceleration, velocity, and deceleration (ICC range = .171-.654). This is reasonable, as averaging a variable over the length of the TRAZER protocol is inherently

more stable than extracting the peak from its path. These results align with Lopes's meta-analysis³², reporting high ICCs for 2D intra-rater reliability. Lopes's meta-analysis reported an ICC of 0.99 for intra-rater reliability in studies that observed frontal plane projection angles with 2D video systems. Although this meta-analysis assessed joint angle reliability, while the current study inspected X-Y coordinate data, this lends further confidence to the reliability of 2D video systems such as TRAZER. As such, TRAZER may be appropriate for use in a setting in which serial measures (test-retest) are obtained to monitor progress, such as a physical therapy clinic or strength and conditioning facility. For instance, periodic measurements could be obtained following a lower extremity injury to quantify the extent to which a patient achieves superior functional performance. Furthermore, baseline TRAZER testing could be appropriate as a post-injury comparison to determine the point at which a patient reaches pre-injury level of function.

Observing high absolute agreement would have provided evidence to support the use of the more clinically-feasible TRAZER system as a clinical injury screening tool in comparison to the more burdensome gold standard 3D motion capture. The lack of strong concurrent validity between TRAZER and 3D motion capture is in stark contrast to a previous study (white paper) conducted by Nyman (Nyman, 2017). The findings of Nyman's comprehensive evaluation testing TRAZER's validity against 3D motion capture demonstrated high correlation for aggregate joint marker accelerations, decelerations, velocities, and distances (ICC range = .83-.96). While both studies possessed similarities, the number of participants differed, Nyman had only two participants, whereas 18 participants were tested in the current study. For validation, we analyzed peak velocity and acceleration while Nyman reported average velocity and average acceleration. Average measures will be more stable than isolated peak measures and may explain the differences between the two sets of findings. Without access to TRAZER's proprietary algorithms for calculation of these metrics, we were unable to assess TRAZER's metric calculations and instead computed the metrics using data obtained from 3D motion capture. It should be noted that Nyman's work was a white paper, for which raw coordinate data were available. As such, TRAZER time

series trajectories of the S2 marker were overlaid on Vicon-obtained trajectories and then submitted to a validity analysis, whereas we assessed the validity of TRAZER-reported discrete metrics with metrics calculated with Vicon coordinate data. Although Nyman reported excellent validity ICCs for time series data, coupling this with the current data suggests that gross metrics (e.g., trajectories, averages) are more valid than discrete (e.g., peaks, maximums, minimums) measures.

Additionally, the systematic error as depicted in the Bland Altman plots for TRAZER and 3D motion capture validity showed the lack of between-system agreement. Over the course of the 40 repetition protocol, TRAZER systematically captured 19.13m less in total distance traveled than 3D motion capture, equating to an 18% discrepancy. A possible explanation for lack of agreement could be the difference in capture rate. It should be noted that the Vicon system sampled data at 60 Hz. This, coupled with an array of eight cameras, naturally allows more accuracy in detecting peaks and maximums, compared with the TRAZER system operating with a single camera and a capture rate of 30 Hz. While TRAZER primarily picks up gross movements, Vicon is designed to record more finely tuned peaks that TRAZER may gloss over, thus producing unagreeable measures. This may explain the systematic differences between the two systems. The authors elected for a 60 Hz Vicon sampling rate because this is more representative of a gold-standard motion capture collection. This is consistent with other researchers who have used different sampling frequencies when comparing two-dimensional with three-dimensional motion capture^{4,18,22}. In fact, in the white paper reported by Nyman, three dimensional motion capture was sampled at 120 Hz, while TRAZER was sampled at 30 Hz. Thus, although different sampling rates likely partially explain the systematic differences observed between the TRAZER and Vicon in the current study, it does not fully account for the discrepancy.

4.1 Clinical and Translational Impact

A key advantage of the TRAZER system is the ability to closely replicate sport demands by presenting visual-cognitive virtual reality challenges that elicit quantifiable whole-body movement responses. Confidence in the system's measurements of reactive responses for clinical documentation of

pre- and post-injury performance capabilities and assessment of injury risk is supported by some of the validity and reliability coefficients derived from our testing. Exceptionally close agreement of the total distance measurement derived from TRAZER with that from the Vicon system, along with very good test-retest reliability, clearly support its use as an indicator of whole-body movement precision in the deactivation of virtual reality targets. These data indicate that TRAZER is appropriate for use as a baseline measure, in addition to post-injury quantification of return to pre-injury levels of functional capacity; thus, allowing for more individualized return-to-play protocols.

An important limitation of this study was the different sampling rates of each system; TRAZER captured at 30 Hz, while Vicon captured at 60 Hz. We acknowledge that this difference may partially account for discrepancies in accuracy. However, as gold standard 3-dimensional motion capture typically samples at higher frequencies, this is a more externally valid comparison. Although participants were instructed to conduct similar activities prior to each day's testing and were tested at similar times each day, we cannot discount the fact that outside factors may have influenced the participants' energy levels and reactions each day. History of concussion was not considered exclusionary for participation. While this does not affect test-retest reliability or concurrent validity, the reader should caution against using these data as reflective of the larger population. Lastly, validation efforts were significantly hampered by the inability to access proprietary algorithms utilized by TRAZER to calculate the mean data for validation. Relying on maximum outputs likely accounted for poor correlation across platforms with different capture rates and camera angles. As evidenced by a higher interrater reliability coefficient and r value for total distance, when measures are taken over time, instead of at maximal points in time, the systems display higher congruency.

TRAZER is designed to be clinically feasible, transportable, and user-friendly compared to the expensive, training-intensive, and sedentary 3D motion capture system. Although the poor validity of TRAZER against a gold standard 3D motion capture hinders that ability to directly compare values obtained from each system, lack of validation does not inhibit TRAZER's ability to provide reliable test-

retest measures. Findings for the test-retest reliability of the TRAZER system may prove beneficial in future studies to determine best uses of TRAZER system in a clinical setting. Additional studies focusing on examining the validity of mean values with a larger sample size could provide further evidence of research and clinical usefulness.

Acknowledgements

This work was not funded.

Conflict of interest

GW previously served in a formal advisory role for Traq Global, Ltd., a role that ended in 2017. The remaining authors have no conflicts of interest. The reported data were acquired from equipment loaned to the US Olympic Committee by Traq Global, Ltd.

References

1. Dinc E, Kilinc BE, Bulat M, Erten YT, Bayraktar B. Effects of special exercise programs on functional movement screen scores and injury prevention in preprofessional young football players. *J Exerc Rehabil* 2017;13(5):535-40. doi:10.12965/jer.1735068.534
2. 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).
3. Ford KR, Myer GD, Hewett TE. Reliability of landing 3D motion analysis: Implications for longitudinal analyses. *Med Sci Sports Exerc* 2007;39(11):2021-8. doi:10.1249/mss.0b013e318149332d
4. Maykut JN, Taylor-Haas JA, Paterno M V, DiCesare CA, Ford KR. Concurrent validity and reliability of 2d kinematic analysis of frontal plane motion during running. *Int J Sports Phys Ther* 2015;10(2):136-46.
5. Cook G, Burton L, Hoogenboom BJ. Functional movement screening: The use of fundamental movements as an assessment of function: Part 1 2014;9(3):396-409.
6. Gribble PA, Hertel J, Facsm A, Plisky P. Using the Star Excursion Balance Test to Assess Dynamic Postural-Control Deficits and Outcomes in Lower Extremity Injury: A Literature and Systematic Review. *J Athl Train* 2012;47(3):339-57. doi:10.4085/1062-6050-47.3.08
7. Padua D, Marshall S, Boling M, Thigpen C, Garrett W, Beutler A. The Landing Error Scoring System (LESS) is a valid and reliable clinical assessment tool of jump-landing

- biomechanics: The JUMP-ACL study. *Am J Sports Med* 2009;37(10):1996-2002 7p.
doi:10.1177/0363546509343200
8. Swanik CB, Cavassin T, Stearne D, Schatz P. The Relationship Between Neurocognitive Function and Noncontact Anterior Cruciate Ligament Injuries. *Am J Sports Med* 2007;35(6):943-8.
 9. Howell DR, Lynall RC, Buckley TA, Herman DC. Neuromuscular Control Deficits and the Risk of Subsequent Injury after a Concussion: A Scoping Review. *Sport Med* 2018:1-19. doi:10.1007/s40279-018-0871-y
 10. McPherson AL, Nagai T, Webster KE, Hewett TE. Musculoskeletal Injury Risk After Sport-Related Concussion A Systematic Review and Meta-analysis 2017:1-8.
doi:10.1177/0363546518785901
 11. Wilkerson GB, Nabhan DC, Prusmack CJ, et al. Detection of Persisting Concussion Effects on Neuromechanical Responsiveness. *Med Sci Sport Exerc* 2018;(10):1750-6.
doi:10.1249/MSS.0000000000001647
 12. Teel E, Gay M, Johnson B, Slobounov S. Determining Sensitivity / Specificity of Virtual Reality-Based Neuropsychological Tool for Detecting Residual Abnormalities Following Sport-Related Concussion. *Neuropsychology* 2016;30(4):474-83.
 13. Anderson KR, Otr L, Woodbury ML, Otr L, Phillips K, Gauthier L V. Virtual Reality Video Games to Promote Movement Recovery in Stroke Rehabilitation : A Guide for Clinicians. *Arch Phys Med Rehabil* 2015;96(5):973-6. doi:10.1016/j.apmr.2014.09.008
 14. Hiraga C, Tonello M, Pellegrini A. Contribution of Virtual Reality (Nintendo Wii) for

- Exercise Training and Rehabilitation. In: Locomotion and Posture in Older Adults. Springer; 2017.
15. Santos F V, Yamaguchi F, Buckley TA, Caccese JB. Virtual reality in concussion management : from lab to clinic. *J Clin Transl Res* 2020;5(4):148-54.
 16. Wilkerson GB, Nabhan DC, Crane RT. Concussion History and Neuromechanical Responsiveness Asymmetry. *J Athl Train* 2020;55(6). doi:10.4085/1062-6050-0401.19
 17. Eltoukhy MA, Kuenze C, Oh J, Signorile JF. Validation of Static and Dynamic Balance Assessment Using Microsoft Kinect for Young and Elderly Populations 2018;22(1):147-53.
 18. Clark RA, Pua Y, Fortin K, et al. Validity of the Microsoft Kinect for assessment of postural control. *Gait Posture* 2012;36(3):372-7. doi:10.1016/j.gaitpost.2012.03.033
 19. Clark RA, Pua Y, Oliveira CC, et al. Reliability and concurrent validity of the Microsoft Xbox One Kinect for assessment of standing balance and postural control. *Gait Posture* 2015;42(2):210-3. doi:10.1016/j.gaitpost.2015.03.005
 20. Eltoukhy M, Kuenze C, Oh J, Apanovitch E, Butler L, Signorile JF. Concurrent Validity of Depth-Sensing Cameras for Noncontact ACL Injury Screening During Side-Cut Maneuvers in Adolescent Athletes: A Preliminary Study. *J Appl Biomech* 2019;35(1):2-10. doi:10.1123/jab.2018-0105
 21. Schmitz A, Ye M, Boggess G, Shapiro R, Yang R, Noehren B. The measurement of in vivo joint angles during a squat using a single camera markerless motion capture system as compared to a marker based system. *Gait Posture* 2015;41(2):694-8.

doi:10.1016/j.gaitpost.2015.01.028

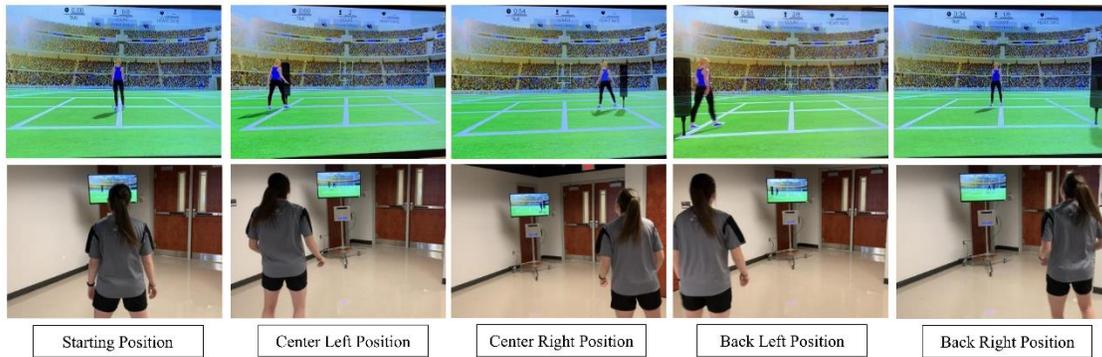
22. Nyman E. A comprehensive evaluation of the TRAZER system: verification. 2017.
23. Menna F, Remondino F, Battisti R, Nocerino E. Geometric investigation of a gaming active device. *Videometrics, Range Imaging, Appl XI*. 2011;8085:80850G.
doi:10.1117/12.890070
24. Shotton J, Sharp T, Fitzgibbon A, et al. Real-Time human pose recognition in parts from single depth images. *Commun ACM* 2013;56(1):116-24. doi:10.1145/2398356.2398381
25. Developers S. signal: Signal processing. 2013. <http://r-forge.r-project.org/projects/signal/>.
26. Moritz S, Bartz-Beielstein T. imputeTS: Time Series Missing Value Imputation in R. *R J*. 2017;9(1):207-8. <https://doi.org/10.32614/RJ-2017-009>.
27. Zeileis A, Grothendieck G. zoo: S3 Infrastructure for Regular and Irregular Time Series. *J Stat Softw*. 2005;14(6):1-27.
28. Henry L, Wickham H. purrr: Functional Programming Tools. 2019. <https://cran.r-project.org/package=purrr>.
29. McGraw KO, Wong SP. Forming Inferences about Some Intraclass Correlation Coefficients. *Psychol Methods* 1996;1(1):30-46. doi:10.1037/1082-989X.1.1.30
30. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* 2016:155-63.
doi:10.1016/j.jcm.2016.02.012
31. Lehnert B. BlandAltmanLeh: Plots (Slightly Extended) Bland-Altman Plots. 2015.

<https://cran.r-project.org/package=BlandAltmanLeh>.

32. Lopes TJA, Ferrari D, Ioannidis J, Simic M, de Azevedo FM, Pappas E. Reliability and Validity of Frontal Plane Kinematics of the Trunk and Lower Extremity Measured With Two-Dimensional Cameras During Athletic Tasks: A Systematic Review With Meta-analysis. *J Orthop Sport Phys Ther* 2018;48(10):1-33. doi:10.2519/jospt.2018.8006
33. Nyman E. A Comprehensive Evaluation of the TRAZER System: Verification. 2017.

Epub ahead of print

Figure 1. Depiction of TRAZER protocol. Top row is the screen as viewed by the participant. Bottom row is the participant reacting to the location of the virtual target.



Prepub ahead of r

Figure 2. Bland-Altman plot depicting limits of agreement between TRAZER and 3 dimensional motion capture for *total distance*. The central dotted line represents systematic error. The outer dotted lines represent ± 2 standard deviations.

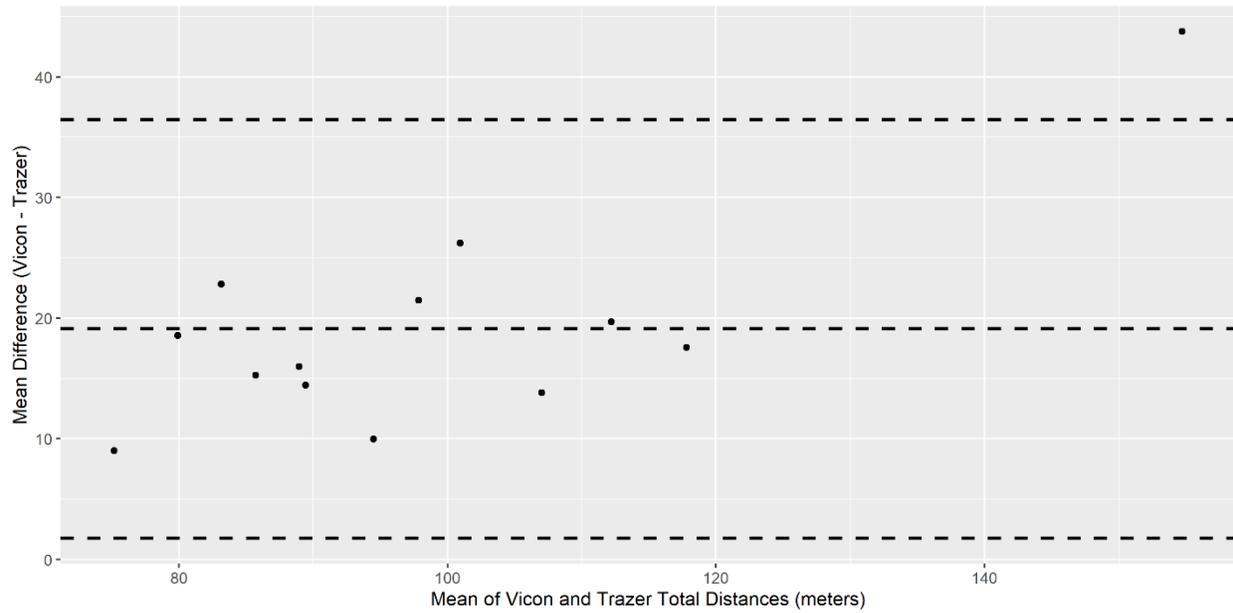


Figure 3. Bland-Altman plot depicting limits of agreement between TRAZER and 3 dimensional motion capture for *maximum velocity*. The central dotted line represents systematic error. The outer dotted lines represent ± 2 standard deviations.

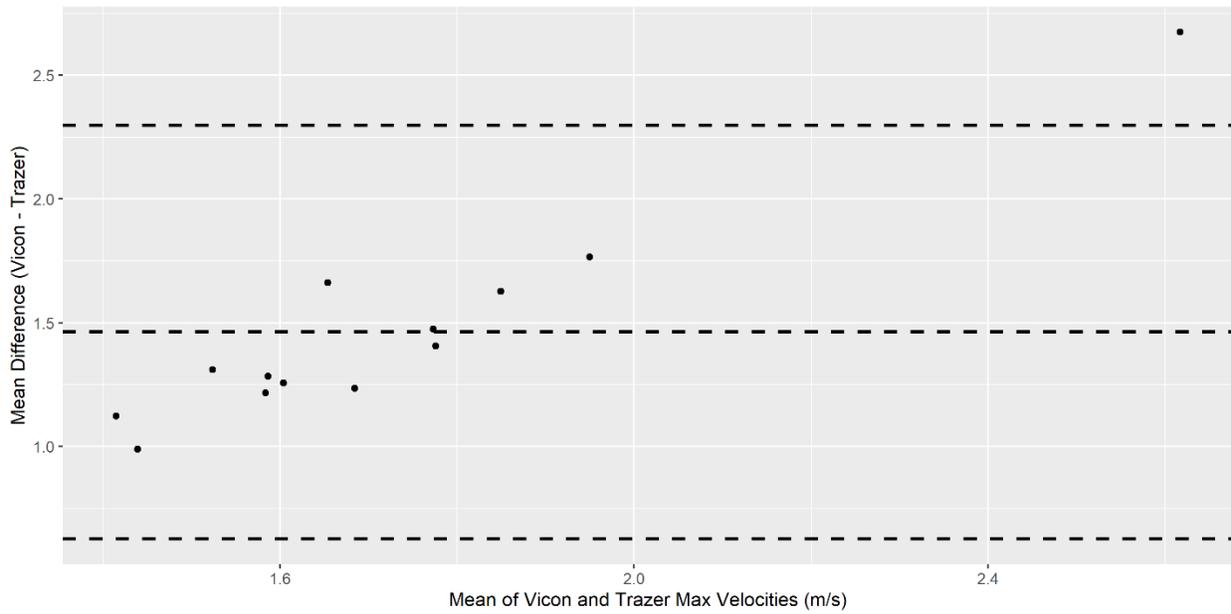


Figure 4. Bland-Altman plot depicting limits of agreement between TRAZER and 3 dimensional motion capture for *maximum acceleration*. The central dotted line represents systematic error. The outer dotted lines represent ± 2 standard deviations.

