INTERNATIONAL JOURNAL OF RACKET SPORTS SCIENCE

VOLUME 4 - ISSUE 2

December, 2022

Index

Articles

Validation of lower limb muscle activation estimated using musculoskeletal modeling against electromyography in the table tennis topspin forehand and backhand
Yoichi Iino, Shinsuke Yoshioka & Senshi Fukashiro 1-10

Reliability and validity of motion sensor and radar for measuring shuttlecock velocity in badminton. Reliability and validity to measure velocity in badminton.
Michael Phomsoupha, Florie Bonneau & Guillaume Laffaye 11-19

The Oldenburg observation sheet for Table Tennis Technique (O3T) as a tool for talent identification and development: a reliability, validity and feasibility study
Till Koopmann, Irene R. Faber, Dirk Büsch & Jörg Schorer 20-31

Accuracy of subjective stats of key performance indicators in tennis
Takashi Jindo Daisuke Mitsuhashi & Tatsumasa Kubota 32-39

Performance analysis in tennis since 2000: A systematic review focused on the methods of data collection
Hiroo Takahash, Shuhei Okamura & Shunsuke Murakami 40-55

Validation of wearables for technical analysis of tennis players
Gabriel Delgado García, Jesús Siquier Coll, Santiago Castro Infantes, Emilio J. Ruiz Malagón, Berta Benito Colio & Francisco Tomás González Fernández 56-61

Int. j. racket sports sci. vol. 4(2), 2022, 1-10. eISSN: 2695-4508
Original Investigation
DOI: 10.30827/Digibug.80316

Validation of lower limb muscle activation estimated using musculoskeletal modeling against electromyography in the table tennis topspin forehand and backhand

Validación de la activación muscular de las extremidades inferiores estimada mediante modelado musculoesquelético y electromiografía en el *topspin* de derecha y revés del tenis de mesa

Yoichi Iino [1]*, Shinsuke Yoshioka[1] and Senshi Fukashiro[1, 2]

1 Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan
2 Japan Women's College of Physical Education, Tokyo, Japan

Received: 08-04-2022

Accepted: 19-09-2022

Abstract

This study aimed to validate the lower limb muscle activation, estimated using static optimization against electromyography (EMG), in the topspin forehand and backhand strokes. The secondary purpose was to compare the estimated activations of the major muscles/muscle groups between the forehand and backhand strokes. Eight male college table tennis players hit the cross-court topspin forehands and backhands with maximum effort. Stroke motions and ground reaction forces were measured using a motion capture system and two force plates. The EMG signals of the 16 lower-limb muscles were recorded using a wireless EMG system. The static optimization algorithm of OpenSim was applied to stroke motions to estimate lower limb muscle activation, which was compared to EMG activation. Of the seven muscles that showed maximum activation > 0.3 during the forehand, five showed a Pearson correlation coefficient > 0.3 Of the four muscles that showed maximum activation > 0.3 during the backhand, all four showed a Pearson correlation coefficient >0.3. However, some muscles, such as the bilateral gluteus medius muscles, showed a low correlation between estimated and EMG activation. A possible cause is the co-contraction of the relevant muscles. Concordance correlation coefficients were smaller than their respective Pearson correlation coefficients. This result reflects that EMG envelope (activation) is also an estimate of muscle activation and is subject to noise and confounding factors. Comparisons with additional independent measurements, such as ultrasound muscle images and instrumented joint loading, are necessary for more robust validation of the musculoskeletal modeling and muscle activation. The gluteus maximus and hamstrings on the playing side, and rectus femoris on the non-playing side exhibited higher activation during the forehand than during the backhand. The overall results suggest that the static optimization algorithm can adequately estimate lower-limb muscle activity during the topspin forehand and backhand strokes.

Keywords: *Musculoskeletal modeling, muscle activation, electromyography, validation.*

Resumen

El objetivo de este estudio fue validar la activación muscular de las extremidades inferiores estimada mediante optimización estática y electromiografía (EMG) en el topspin de derecha y revés. El objetivo secundario fue comparar las activaciones estimadas de los principales grupos musculares entre los golpes de derecha y

Corresponding author: Yoichi Iino, iino@idaten.c.u-tokyo.ac.jp

Cite this article as:
Iino, Y., Yoshioka, S., & Fukashiro, S. (2022). Validation of lower limb muscle activation estimated using musculoskeletal modeling against electromyography in the table tennis topspin forehand and backhand. *International Journal of Racket Sports Science*, 4(2), 1-10.

revés. Ocho jugadores hombre universitarios de tenis de mesa realizaron con el máximo esfuerzo los golpes topspin de derecha y revés cruzados en la pista. Los movimientos de los golpes y las fuerzas de reacción del suelo fueron medidos con un sistema de captura del movimiento y dos placas de fuerza. Las señales EMG de los músculos de los 16 miembros inferiores fueron grabadas con un sistema EMG inalámbrico. Se usó el algoritmo de optimización estática OpenSim para estimar la activación muscular de los miembros inferiores durante los golpes, y luego se compararon los resultados con la activación de la EMG. De los siete músculos que mostraron activación máxima > 0,3 en el golpe de derecha, cinco mostraron un coeficiente de correlación de Pearson > 0,3. De los cuatro músculos que mostraron activación máxima > 0,3 durante el golpe de revés, los cuatro mostraron un coeficiente de correlación de Pearson > 0,3. Sin embargo, algunos músculos, como el glúteo medio, mostraron una baja correlación entre la activación estimada y la EMG. Una posible causa es la cocontracción de los músculos involucrados. Los coeficientes de correlación de concordancia fueron menores que sus respectivos coeficientes de correlación de Pearson. Este resultado refleja que la envolvente (activación) de la EMG es también una estimación de la activación muscular y está sujeta a ruido y factores de confusión. Es necesario realizar comparaciones con otras mediciones independientes, como las imágenes musculares por ultrasonido y la carga articular con instrumentos, para lograr una validación más sólida del modelado musculoesquelético y la activación muscular. El glúteo mayor y los isquiotibiales en el lado de juego, y el recto femoral en el lado de no juego, mostraron una mayor activación durante el golpe de derecha que durante el revés. Los resultados generales sugieren que el algoritmo de optimización estática puede estimar adecuadamente la actividad muscular de las extremidades inferiores durante el *topspin* de derecha y revés.

Palabras clave: *Modelado musculoesquelético, activación muscular, electromiografía, validación.*

INTRODUCTION

Topspin forehand and backhand strokes are fundamental techniques used in table tennis, and mastering the effective execution of these strokes is essential for high performance (Seemiller & Holowchak, 1997). A previous study reported that topspin forehand was the most frequently used stroke in elite matches, followed by counter-topspin topspin forehand, with topspin backhand ranking fourth (Malagoli Lanzoni, Di Michele, & Merni, 2014). The same study found that the topspin forehand and counter-topspin forehand were more related to winners than other strokes. Both strokes are performed by utilizing the kinetic chain of the entire body. Previous studies on table tennis strokes have reported that joint angular velocities (Bańkosz & Winiarski, 2018) and hip joint kinetics (Iino, 2018) are associated with racket speed, and have suggested that the lower limbs energetically contribute to the generation of racket speed in the topspin forehand and backhand (Iino & Kojima, 2011; 2016). Kinematic and kinetic analyses have been conducted on the lower limb motions during table tennis topspin forehands and backhands (He et al., 2021; Le Mansec, Dorel, Hug, & Jubeau, 2018; Malagoli Lanzoni, Bartolomei, Di Michele, & Fantozzi, 2018; Qian, Zhang, Baker, & Gu, 2016; Shao et al., 2020; Wang et al., 2018). Two of these studies examined the lower limb muscles using electromyography (EMG). Wang et al. (2018) compared kinematics and EMG data between elite and amateur players during topspin backhands. They found that the hip and knee flexion angles at backswing were larger in elite players than in amateurs and that the maximum activation of the rectus femoris and tibialis anterior was lower in elite players than in amateurs. Le Mansec et al. (2018) reported the EMG of the eight lower limb muscles of the playing side (right side for a right-handed player) in seven typical strokes, including the topspin forehand and backhand. They found that the EMG peak amplitudes of gluteus maximus and biceps femoris were larger than 60% of their maximum voluntary contraction amplitudes in the topspin forehands and forehand smash. These studies provide valuable insights into the unique muscle activation characteristics of elite athletes and different stroke types. However, surface EMG can only be applied to surface muscles, and this alone cannot provide information on muscle forces. Therefore, the function of the muscles in table tennis forehand and backhand is still not fully understood.

Estimation of muscle activation and forces in table tennis strokes can be used to inform performance improvement and injury prevention. As a non-invasive approach, musculoskeletal modeling has been utilized to estimate lower limb muscle activation and forces in human locomotion such as walking and running through predictive and tracking simulations (e.g., Dorn, Schache, & Pandy, 2012; Liu, Anderson, Schwartz, & Delp, 2008; Neptune, Sasaki, & Kautz, 2008) and has revealed the mechanical functions of the lower limb muscles. To our knowledge, there are no studies that have estimated the lower-limb muscle activation in table tennis strokes using musculoskeletal modeling. The estimated lower limb activation has been validated against EMG in locomotion previously (Alexander & Schwameder, 2016; Dupré, Dietzsch, Komnik, Potthast, & David, 2019; Trinler, Leboeuf, Hollands, Jones, & Baker, 2018; Wibawa et al., 2016; Żuk, Syczewska, & Pezowicz, 2018), and these studies reported moderate to good associations between the estimated and EMG activations. Table tennis topspin forehands and backhands require whole-body rotation (Iino & Kojima, 2009; 2016), and different players may have different techniques for these strokes (Bańkosz & Winiarski, 2018). Thus, the extent to which the estimated lower-limb activations in table tennis can be validated against EMG is unclear.

The following two popular algorithms have been used to estimate muscle activations and forces in OpenSim, which is an open-source software tool for musculoskeletal modeling and simulation of movement (Delp et al., 2007): static optimization and computed muscle control. In this study, a static optimization algorithm was used to estimate the muscle activation because the computed muscle control algorithm conducts forward dynamic simulations to track measured kinematics and requires accurate modeling of the upper body, which is difficult for table tennis strokes that involve complex spine and shoulder motions. Additionally, previous studies have shown that computed muscle control is not always more accurate in estimating muscle activation than static optimization (Alvim, Lucareli, & Menegaldo, 2018; Roelker et al., 2020; Trinler et al., 2018).

The purpose of this study was to validate the lower limb muscle activation estimated using static optimization against EMG in table tennis topspin forehand and backhand strokes. The secondary purpose was to compare the estimated activations of the major muscles/muscle groups between the two strokes because no studies have yet made these comparisons. This could provide a scientific basis for developing effective strength-training programs for table tennis.

METHODS

Participants

Eight male college table tennis players participated in this study. All participants were members of a Division I table tennis team in the Kanto Collegiate Table Tennis League in Japan. Their mean ± standard deviation age, height, body mass, and training experience were 20.2 ± 1.5 years, 1.72 ± 0.06 m, 67.5 ± 6.5 kg, and 11.5 ± 2.3 years, respectively. Six players were right-handed and two were left–handed. The dominant hand was judged by the hand holding the racket. All were offensive players. All participants provided written informed consent. The experimental procedures were approved by a local ethics committee.

Experimental procedure

After an individual warm-up session, the participants were asked to hit the topspin cross-court forehands and backhands with maximum effort (Figure 1). They were asked to place their feet on a separate force plate (type 9281; Kistler, Winterthur, Switzerland) during preparation, but they were allowed to move their feet after the beginning of a stroke. At least three successful forehand and backhand strokes were recorded for each participant. Before data collection, the participants were asked to practice the strokes until they became accustomed to the experimental settings. The position of the table tennis table was adjusted for each stroke type for each participant to ensure that the feet were within the boundaries of the force plates at preparation. All participants used the same shake-hands racket (Timo Boll ALC ST, Tamasu Co., Ltd., Japan) with inverted rubber (Tenergy 05, Tamasu Co., Ltd., Japan). A ball machine (Butterfly Amicus 1000, Tamasu Co., Ltd., Japan) was used to feed the players light backspin balls (Nittaku premium three-star, Nippon Takkyu Co., Ltd.). Balls were projected directly to the foreside and backside of participant's court in topspin forehands and backhands, respectively. The spin rate of the backspin balls after the bounce on the table was 8.6 ± 1.3 rps. The ball feeding frequency was about 43 balls/min. The ball machine was set at -1 for SPIN and 7.0 for SPEED. Finally, they were asked to perform a sequence of hip flexion, extension, abduction, and circumduction of each leg to estimate the locations of the hip joint centers.

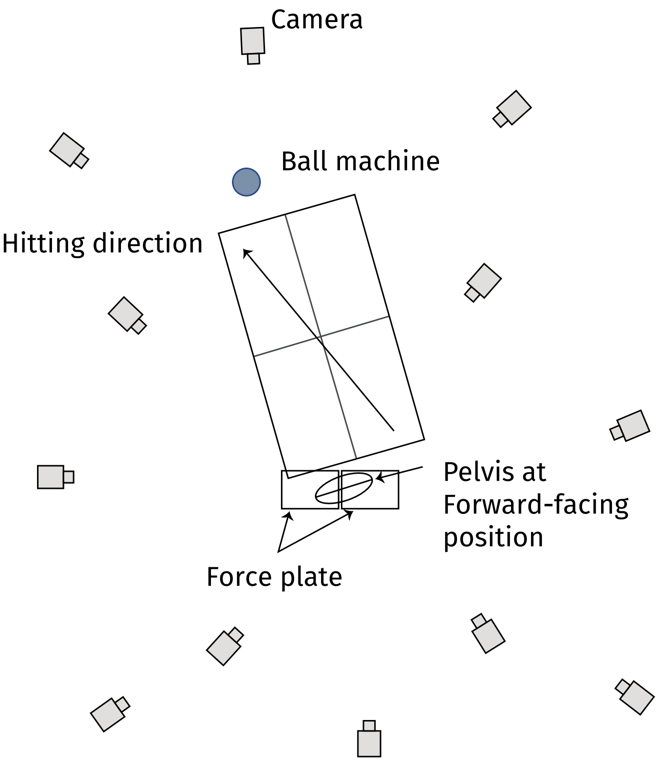


Figure 1. The experimental setup for the topspin forehand. The forward-facing position of the pelvis was defined as the position where the pelvis medio-lateral axis was parallel to the endline of the table tennis table.

Data collection

The participants wore tight-fitting swim pants and table tennis shoes. A total of 51 retro-reflective markers (diameter, 16 m) were attached to landmarks on the whole body. Four markers were attached to the lateral side of the racket face. Three-dimensional marker coordinates were obtained using a 12-camera motion capture system (MAC3D System; Motion Analysis, Santa Rosa, CA, USA) at 200 Hz. The force plate data were recorded at 2,000 Hz. The surface EMG activity was recorded using a wireless EMG system (Trigno Wireless System, DELSYS. Boston, MA, USA). EMG signals were bandpass filtered (20-450 HZ) and sampled at 2,000 Hz. EMG electrodes were placed bilaterally on the gluteus maximus, gluteus medius, biceps femoris, rectus femoris, vastus medialis, tibialis anterior, soleus, and gastrocnemius lateralis muscles. Electrode placement was determined according to SENIAM guidelines (Hermens, Freriks, Disselhorst-Klug, & Rau, 2000). The skin where the electrodes were placed was shaved if necessary and cleaned with alcohol to reduce impedance.

Data processing

The forehand and backhand strokes with the highest racket tip speed for each participant were selected for analysis. Several virtual landmarks were created to scale a generic model using the OpenSim scale tool. The ground reaction force data and kinematics were smoothed using a zero-lag 6 Hz second-order low-pass Butterworth filter. The EMG signals were full-wave rectified and filtered to create a linear envelope using a 6 Hz second-order low-pass Butterworth filter. The positions of the hip joint centers were estimated using a functional method (Gamage & Lasenby, 2002; Halvorsen, 2003).

Musculoskeletal model and estimation of muscle activations

We used a modified version of the OpenSim musculoskeletal model published by Lai et al. (2017). Two additional degrees of freedom of adduction/abduction and internal/external rotation were added to each knee joint. The range of motion was -5° to 5° for adduction (+) and -30° to 30° for internal rotation (+). These values were determined according to Ramsey & Wretenberg, (1999).

Lower limb muscle activation was estimated using a static optimization algorithm with OpenSim 3.3 (Seth, Sherman, Reinbolt, & Delp, 2011). The modified model was scaled for each participant in a static standing position. The maximum isometric force of each muscle actuator was scaled by a factor of 1.8–2.1 to account for possible stronger muscles in younger players. The scale was set so that the estimated activation level would not remain at the maximum (i.e., 1) for more than 15 ms for all muscles because such persistent full activation was not observed in EMG activation. Subsequently, the joint angles during stroke sequences were determined using the Inverse Kinematics tool. The calculated joint angles and ground reaction forces were then used to estimate the lower limb muscle activations through static optimization, which resolves the indeterminacy of muscle forces by minimizing the squared sum of muscle activations.

EMG envelope data were time-shifted by 40 ms to account for the electromechanical delay (Begovic, Zhou, Li, Wang, & Zheng, 2014; Dupré et al., 2019; Zhou, Lawson, Morrison, & Fairweather, 1995). The timings of stroke events were determined in accordance with previous studies (Iino & Kojima, 2011; 2016); However, the origin of the pelvis instead of the shoulder joint was used to define the beginning of a backhand stroke as described below because the present study focused on the lower limb movements. The beginning of a forehand stroke was defined as the time when the pelvis negative (clockwise) axial angular velocity exceeded -0.5 rad/s (for the players who temporarily stopped the pelvis rotation between two strokes) and the time when the pelvis rotated backward beyond the forward-facing position (for the remaining players) (Figure 1). The beginning of a backhand stroke was defined as the time when the origin of the pelvis (midpoint of both hip joint centers) made a preliminary downward movement. The beginning of the forward swing was defined as the time when the pelvis began to rotate forward for the forehand and the time when the pelvis began to move upward for the backhand. Time was normalized to the duration from the beginning of the stroke to the peak racket speed. Muscle activation data were analyzed from 0% (beginning of stroke) to 120% (follow-through) normalized time.

Statistical analysis

Pearson correlation coefficients were determined between the EMG envelope and the estimated muscle activation data for 121 time points. Pearson correlation coefficients were classified as small ($r = 0.1–0.29$), moderate ($r = 0.3–0.49$) or large ($r = 0.5–1$) (Cohen, 2013). In addition, Lin's concordance correlation coefficients (CCC) (Lin, 1989) were determined between the EMG and estimated muscle activations. CCC quantifies the closeness of the two measurements to the 45 degree line that passes through the origin.

The Shapiro-Wilk test was used to evaluate the non-normality of the distribution for the maximum activation of the following lower limb muscles/muscle groups during the forehand and backhand: gluteus maximus, gluteus maximus, adductor magnus, hamstrings, vastus muscles, gastrocnemius, soleus, and anterior tibialis. The test revealed that the distribution for the gluteus maximus on the playing side during forehand significantly departed from normality ($P = 0.0017$); hence, the Wilcoxon signed-rank test was used to compare the maximum activation between the strokes for that muscle. A two-tailed t-test was used to analyze the remaining muscles. Statistical significance was set at $P < 0.01$.

RESULTS

The maximum racket speed was 20.6 ± 1.6 m/s for the forehand topspin and 21.5 ± 1. 6 m/s for the backhand topspin.

In the forehand, seven muscles showed a maximum estimated activation of > 0.3 (Table 1). The rectus femoris and gluteus maximus of the playing side and the rectus femoris of the non-playing side showed a maximum estimated activation of > 0.5. The tibialis anterior and biceps femoris of the playing side and the gastrocnemius lateralis and gluteus maximus of the non-playing side showed maximum estimated activation between 0.3 and 0.5. The Pearson correlation coefficient between EMG and estimated activations was > 0.3 for eight muscles of the forehand (Table 1). Most muscles that showed substantial maximum activation (> 0.3) exhibited a Pearson correlation coefficient higher than 0.3, except the gastrocnemius lateralis and gluteus maximus on the non-playing side, with mean Pearson correlation coefficients of 0.277 and 0.218, respectively (Table 1). Concordance correlation coefficients were smaller than their respective Pearson coefficients and were > 0.3 for five muscles. Peak EMG activation was observed during the forward swing phase for all muscles that showed substantial activation (Table 1, Figure 2). The rectus femoris and gluteus maximus of the playing side and rectus femoris of the non-playing side showed peak EMG activation after the beginning of the forward swing (Figure 2).

For the backhand, four muscles showed a maximum activation > 0.3 (Table 1). Only the rectus femoris on the playing side showed a maximum activation of > 0.5. The soleus of the playing side and rectus femoris and gluteus maximus of the non-playing side showed maximum activations between 0.3 and 0.5. These muscles exhibited a Pearson correlation coefficient of > 0.3 (Table 1). As was in the forehand stroke, concordance correlation coefficients were smaller than Pearson coefficients and were > 0.3 for seven muscles in the backhand (Table 1). In the backhand stroke, peak EMG activation was also observed during the forward swing phase, except for the vastus medialis on the non-playing side, which showed higher activation during the backswing phase (Figure 3).

The gluteus maximus, hamstrings of the playing side, and rectus femoris of the non-playing side showed a statistically higher maximum estimated activation in the forehand than in the backhand (P = 0.0078, P = 0.00055, and P = 0.0002, respectively, Figure 4). Three muscles showed a maximum estimated activation of > 0.5 during the forehand, whereas only the rectus femoris of the playing side showed a maximum activation > 0.5.

Table 1.
Pearson correlation coefficients and concordance correlation coefficients between EMG and estimated activation levels for lower limb muscles during forehand and backhand topspin strokes.

	Forehand				Backhand			
	Pearson Correlation Coefficient, *r*	Concordance Correlation Coefficient		Maximum activation	Pearson Correlation Coefficient, *r*	Concordance Correlation Coefficient		Maximum activation
		r_c	s.e.			r_c	s.e.	
Playing side								
tibialis anterior	**0.300±0.164**	0.204±0.106	0.062±0.014	**0.31±0.17**	0.033±0.266	0.023±0.193	0.055±0.014	0.08±0.08
gastrocnemius lateralis	0.145±0.272	0.087±0.184	0.061±0.010	0.15±0.07	**0.335±0.434**	0.274±0.355	0.057±0.016	0.13±0.07
soleus	**0.528±0.163**	**0.451±0.161**	0.062±0.010	0.24±0.12	**0.650±0.138**	**0.581±0.153**	0.052±0.011	**0.33±0.17**
vastus medialis	**0.349±0.317**	0.282±0.292	0.064±0.022	0.23±0.16	**0.566±0.244**	**0.534±0.247**	0.056±0.024	0.20±0.08
rectus femoris	**0.616±0.201**	**0.572±0.193**	0.055±0.015	**0.58±0.22**	**0.605±0.152**	**0.555±0.149**	0.058±0.014	**0.53±0.16**
biceps femoris	**0.640±0.267**	**0.530±0.257**	0.049±0.017	**0.43±0.23**	0.139±0.293	0.101±0.209	0.050±0.018	0.06±0.05
gluteus medius	0.088±0.230	0.056±0.143	0.052±0.007	0.24±0.17	0.120±0.339	0.059±0.224	0.059±0.018	0.07±0.05
gluteus maximus	**0.883±0.080**	**0.860±0.089**	0.023±0.013	**0.93±0.10**	0.081±0.314	0.063±0.289	0.064±0.013	0.11±0.10
Non-playing side								
tibialis anterior	0.069±0.368	0.066±0.298	0.066±0.012	0.17±0.15	**0.333±0.217**	0.219±0.176	0.055±0.014	0.14±0.06
gastrocnemius lateralis	0.277±0.0457	0.215±0.324	0.058±0.013	**0.34±0.17**	0.245±0.309	0.221±0.288	0.070±0.021	0.25±0.14
soleus	0.169±0.349	0.128±0.281	0.072±0.013	0.26±0.13	0.288±0.335	0.208±0.264	0.066±0.012	0.16±0.11
vastus medialis	**0.343±0.292**	0.251±0.306	0.047±0.028	0.27±0.31	**0.442±0.264**	**0.345±0.246**	0.045±0.025	0.19±0.16
rectus femoris	**0.795±0.099**	**0.746±0.112**	0.038±0.011	**0.85±0.21**	**0.354±0.248**	**0.323±0.257**	0.067±0.022	**0.34±0.23**
biceps femoris	0.234±0.257	0.171±0.190	0.055±0.011	0.13±0.23	**0.666±0.281**	**0.614±0.287**	0.046±0.024	**0.23±0.11**
gluteus medius	0.00±0.237	-0.010±0.177	0.076±0.010	0.19±0.10	0.282±0.319	0.223±0.310	0.060±0.023	**0.13±0.11**
gluteus maximus	0.218±0.272	0.147±0.219	0.065±0.020	**0.35±0.33**	**0.667±0.223**	**0.604±0.235**	0.049±0.023	0.40±0.14

Correlation coefficients > 0.3 (moderate or large in accordance with Cohen (2913)) are shown in bold for clarity.

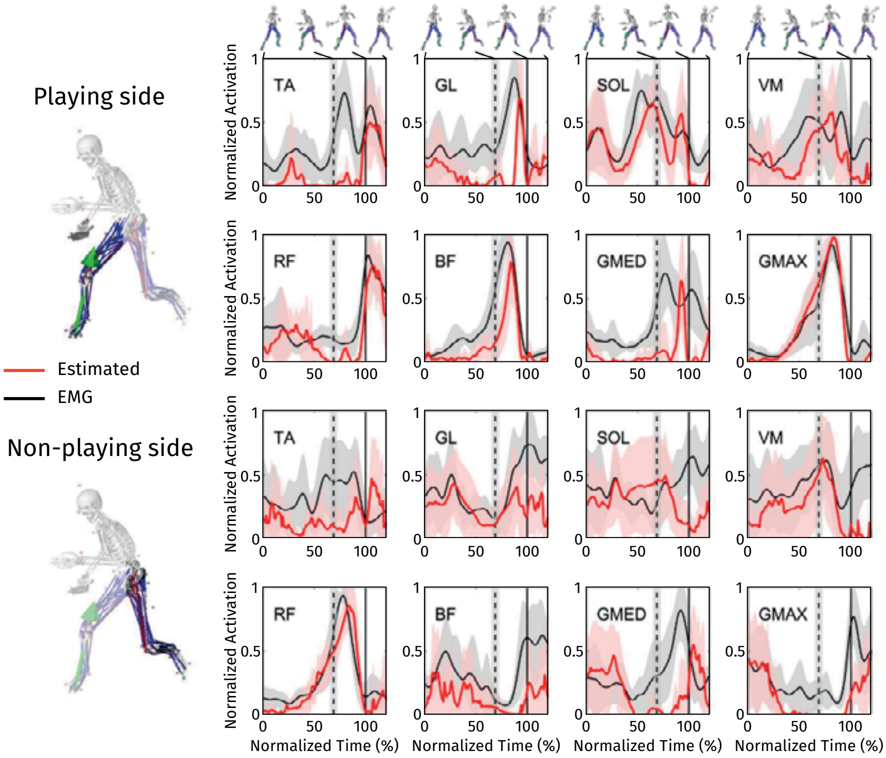


Figure 2. Normalized estimated (red) and EMG (black) activations of the lower limb muscles during the topspin forehand. Vertical lines represent the completion of backswing (dashed) and the occurrence of maximum racket speed (solid). Shared areas show standard deviations for the participants. TA; tibialis anterior, GL; gastrocnemius lateralis, SOL; soleus, VM; vastus medialis, RF; rectus femoris, BF; biceps femoris, GMED; gluteus medius, GMAX; gluteus maximus.

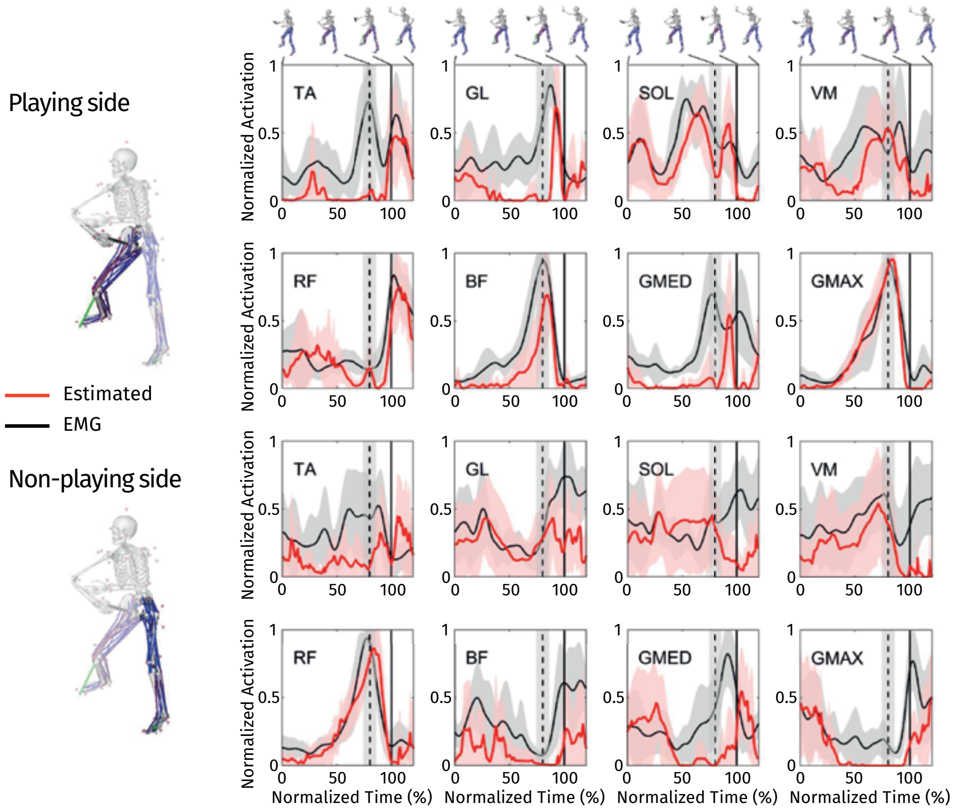


Figure 3. Normalized estimated (red) and EMG (black) activations of the lower limb muscles during the topspin backhand. Vertical lines represent the completion of backswing (dashed) and the occurrence of maximum racket speed (solid). Shared areas show standard deviations for the participants. TA; tibialis anterior, GL; gastrocnemius lateralis, SOL; soleus, VM; vastus medialis, RF; rectus femoris, BF; biceps femoris, GMED; gluteus medius, GMAX; gluteus maximus.

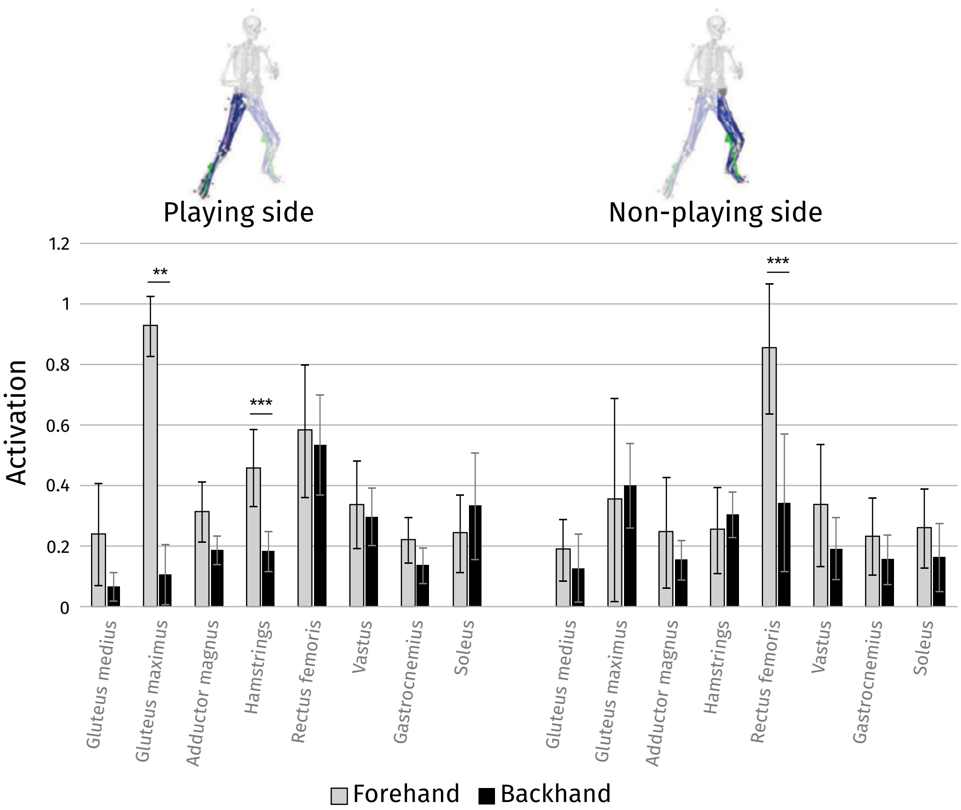


Figure 4. Maximum estimated activation of the lower limb muscles/muscle groups during topspin forehand and backhand. ***P<0.001, **P<0.01.

DISCUSSION

This study aimed to validate the estimation of lower-limb muscle activation during table tennis forehand and backhand through comparison with EMG measurements. We also aimed to compare the estimated activation between forehand and backhand strokes. The maximum racket resultant velocities (20.6 ± 1.6 m/s for the forehand and 21.5 ± 1.3 m/s for the backhand) were similar to or higher than those in previous studies (Bańkosz & Winiarski, 2018; Iino & Kojima, 2009; 2016).

The comparison between the estimated and EMG activation suggests that the static optimization algorithm can adequately estimate lower limb muscle activity during table tennis topspin forehand and backhand. For the four muscles that showed maximum activation of > 0.5, the Pearson correlation coefficients were > 0.5 (Table 1). Of the seven muscles that showed maximum activation between 0.3 and 0.5, five showed a Pearson correlation coefficient > 0.3 (Table 1). However, the Pearson correlation coefficients for the gastrocnemius and gluteus maximus of the non-playing side during forehand were lower than 0.3 (although their maximum activations were higher than 0.3; Table 1).

Co-contraction is a possible cause for the lower correlations observed in the gluteus medius on the playing side and the gluteus maximus and gluteus medius on the non-playing side. EMG recordings showed that the gluteus maximus on the non-playing side was active for 60–80% of the swing phase during the forehand (Figure 2), but the estimated activation was small because the hip joint on the non-playing side exerted flexion torque during that phase (not shown in the Results section). EMG recordings also showed that the bilateral gluteus medius muscles were activated during the follow-through phase of the forehand. However, six of the eight players demonstrated adduction torque at each hip joint during this phase, which resulted in lower estimated activation of the gluteus medius muscles in these players (Figure 2). These results suggest that the co-contraction of the adductor/abductor and flexor/extensor muscles that occurred at both hip joints in the forehand stroke resulted in lower correlations observed in the relevant muscles. Methods for estimating muscle co-contraction using shift parameters (MacIntosh & Keir, 2017) or contraction entropy (Jiang & Mirka, 2007) have been proposed. Future research is needed to establish when and at which joint co-contraction occurs in table tennis strokes to accurately estimate the activation of the lower limb muscles.

Another reason for the lower correlation for some muscles in some participants may be that the model parameters used were not adjusted for each participant as indicated in previous studies (Dupré et al., 2019; Trinler et al., 2018; Wibawa et al., 2016; Żuk et al., 2018). For example, the Pearson correlations for the gastrocnemius of the non-playing side during the forehand and the gastrocnemius of the playing side during the backhand varied substantially among the players (see standard deviation values in Table 1). Muscle parameters, such as the force-length relationship of the muscle and tendon in the generalized model, might not have been appropriate for some players.

Concordance correlation coefficients were smaller than respective Pearson correlation coefficients. Only the gluteus maximus of the playing side during the forehand exhibited a concordance correlation coefficients > 0.8. These results suggest that there was "scale shift" or "location shift" between EMG and estimated activations (Lin, 1989). It should be noted that this reflects that EMG linear envelope is only an estimate of muscle activation and is subject to noise and confounding factors (Staudenmann, Roeleveld, Stegeman, & van Dieën, 2010). For more robust validation of the musculoskeletal modeling and muscle activation, comparisons with additional independent data, such as ultrasound images of muscles and instrumented joint loading, would be necessary (Hicks, Uchida, Seth, Rajagopal, & Delp, 2015).

It is worth comparing the results of the present study with those obtained for walking and other types of locomotion in previous studies (Alexander & Schwameder, 2016; Dupré et al., 2019; Trinler et al., 2018; Wibawa et al., 2016; Żuk et al., 2018). Although these studies have reported that lower limb muscle activation estimated using musculoskeletal modeling generally showed moderate to good agreement with EMG activation, Trinler et al. (2018) suggested that the consistency of agreement between measured and estimated activation levels at different walking speeds was not high enough to recommend immediate clinical adoption. Many factors, such as the musculoskeletal models used (OpenSim, AnyBody), EMG signal processing methods, statistical methods (Pearson or Spearman), and the phase of analysis (stance phase or complete gait cycle), differed between studies, making it difficult to quantitatively compare the correlation coefficients between these studies. Overall, the present study suggests that the static optimization algorithm can estimate lower-limb muscle activity during table tennis forehand and backhand with a similar degree of validity to that of locomotion.

The results suggest that the gluteus maximus and hamstrings of the playing side and the rectus femoris of the non-playing side exhibit higher activation during the forehand than during the backhand. The gluteus maximus and biceps femoris muscles show high activation. These results were consistent with Le Mansec et al.'s (2018) findings on lower limb EMG and the previous studies (Chen et al., 2022; Qian et al., 2016) that suggested that advanced players would use lower limb drive more effectively than intermediate players in the topspin forehand. Our study also suggests that the rectus femoris on the non-playing side is highly activated during the topspin forehand.

In contrast, the lower limb muscles showed relatively low activation during the topspin backhand. This result is consistent with a previous study that found that the angular velocities of playing and non-playing side hip extension and ankle flexion are positively correlated with racket speed in the topspin forehand whereas the angular velocities of the racket arm are correlated with racket speed in the topspin backhand (Bańkosz & Winiarski, 2018). Previous studies (Iino & Kojima, 2011; 2016) have reported that approximately 80% of the mechanical energy of the racket arm at ball impact was due to the energy transfer from the trunk in both the backhand and forehand strokes. Considering that the maximum racket speeds were similar for both strokes in the present study, the trunk muscles may be more highly activated for mechanical work in the backhand than in the forehand, or mechanical energy may be transferred more efficiently through the trunk in the backhand than in the forehand.

The present study has some limitations. First, EMG data were not recorded for maximum voluntary contraction (MVC). Thus, EMG activation could not be normalized to MVC values. Second, the maximum isometric forces of each muscle actuator did not reflect the maximum isometric joint torque for each player because such kinetics were not measured. We focused instead on the patterns of estimated and EMG activations, which were not affected by these normalizing values. Finally, only the static optimization algorithm using the OpenSim model was assessed. Other algorithms such as computed muscle control and other musculoskeletal models should be investigated in future studies.

CONCLUSIONS

The present study suggests that the static optimization algorithm can adequately estimate lower-limb muscle activity during table tennis topspin forehand and backhand strokes. The gluteus maximus and rectus femoris on the playing side and rectus femoris on the non-playing side showed high activation during the forehand. Only the rectus femoris on the playing side showed high activation in the backhand. For these four muscles, the Pearson correlation coefficients were higher than 0.5. A lower Pearson correlation between the estimated and EMG activation was observed for some muscles, including both gluteus medius muscles, during the forehand. A possible cause is the co-contraction of relevant muscles. All concordance correlation coefficients were smaller than their respective Pearson correlation coefficients. The gluteus maximus and hamstrings on the playing side, and rectus femoris on the non-playing side exhibited higher activation during the forehand than during the backhand.

FUNDING

This study was supported by the Japan Table Tennis Association.

CONFLICT OF INTERESTS

The authors declare that there are no conflicts of interest.

REFERENCES

Alexander, N., & Schwameder, H. (2016). Comparison of estimated and measured muscle activity during inclined walking. *Journal of Applied Biomechanics, 32*(2), 150-159. https://doi.org/10.1123/jab.2015-0021

Alvim, F. C., Lucareli, P. R. G., & Menegaldo, L. L. (2018). Predicting muscle forces during the propulsion phase of single leg triple hop test. *Gait and Posture, 59*, 298-303. https://doi.org/10.1016/j.gaitpost.2017.07.038

Bańkosz, Z., & Winiarski, S. (2018). Correlations between Angular Velocities in Selected Joints and Velocity of Table Tennis Racket during Topspin Forehand and Backhand. *Journal of Sports Science & Medicine, 17*(2), 330-338. https://www.ncbi.nlm.nih.gov/pubmed/29769835

Begovic, H., Zhou, G. Q., Li, T., Wang, Y., & Zheng, Y. P. (2014). Detection of the electromechanical delay and its components during voluntary isometric contraction of the quadriceps femoris muscle. *Frontiers in Physiology, 5*, 1-8. https://doi.org/10.3389/fphys.2014.00494

Chen, M.-Z., Wang, X., Chen, Q., Ma, Y., Malagoli Lanzoni, I., & Lam, W.-K. (2022). An analysis of whole-body kinematics, muscle strength and activity during cross-step topspin among table tennis players. *International Journal of Performance Analysis in Sport, 22*(1), 16-28. https://doi.org/10.1080/24748668.2022.2025712

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences* (2nd editio). Hillside, NJ: Routledge. https://doi.org/10.4324/9780203771587

Delp, S. L., Anderson, F. C., Arnold, A. S., Loan, P., Habib, A., John, C. T., Guendelman, E., & Thelen, D. G. (2007). OpenSim: Open-source software to create and analyze dynamic simulations of movement. *IEEE Transactions on Biomedical Engineering*. https://doi.org/10.1109/TBME.2007.901024

Dorn, T. W., Schache, A. G., & Pandy, M. G. (2012). Muscular strategy shift in human running: dependence of running speed on hip and ankle muscle performance. *Journal of Experimental Biology*. https://doi.org/10.1242/jeb.075051

Dupré, T., Dietzsch, M., Komnik, I., Potthast, W., & David, S. (2019). Agreement of measured and calculated muscle activity during highly dynamic movements modelled with a spherical knee joint. *Journal of Biomechanics, 84*, 73-80. https://doi.org/10.1016/j.jbiomech.2018.12.013

Gamage, S. S. H. U., & Lasenby, J. (2002). New least squares solutions for estimating the average centre of rotation and the axis of rotation. *Journal of Biomechanics, 35*(1), 87-93. https://doi.org/10.1016/S0021-9290(01)00160-9

Halvorsen, K. (2003). Bias compensated least squares estimate of the center of rotation. *Journal of Biomechanics, 36*(7), 999-1008. https://doi.org/10.1016/S0021-9290(03)00070-8

He, Y., Sun, D., Yang, X., Fekete, G., Baker, J. S., & Gu, Y. (2021). Lower limb kinetic comparisons between the chasse step and one step footwork during stroke play in table tennis. *PeerJ, 9*, 1-14. https://doi.org/10.7717/peerj.12481

Hermens, H. J., Freriks, B., Disselhorst-Klug, C., & Rau, G. (2000). Development of recommendations for SEMG sensors and sensor placement procedures. *Journal of Electromyography and Kinesiology, 10*(5), 361-374. https://doi.org/10.1016/S1050-6411(00)00027-4

Hicks, J. L., Uchida, T. K., Seth, A., Rajagopal, A., & Delp, S. L. (2015). Is my model good enough? best practices for verification and validation of musculoskeletal models and simulations of movement. *Journal of Biomechanical Engineering, 137*(2), 020905. https://doi.org/10.1115/1.4029304

Iino, Y. (2018). Hip joint kinetics in the table tennis topspin forehand: relationship to racket velocity. *Journal of Sports Sciences, 36*(7), 834-842. https://doi.org/10.1080/02640414.2017.1344777

Iino, Y., & Kojima, T. (2009). Kinematics of table tennis topspin forehands: effects of performance level and ball spin. *Journal of Sports Sciences, 27*(12), 1311-1321. https://doi.org/10.1080/02640410903264458

Iino, Y., & Kojima, T. (2011). Kinetics of the upper limb during table tennis topspin forehands in advanced and intermediate players. *Sports Biomechanics, 10*(4), 361-377. https://doi.org/10.1080/14763141.2011.629304

Iino, Y., & Kojima, T. (2016). Mechanical energy generation and transfer in the racket arm during table tennis topspin backhands. *Sports Biomechanics, 15*(2), 180-197. https://doi.org/10.1080/14763141.2016.1159722

Jiang, Z., & Mirka, G. A. (2007). Application of an entropy-assisted optimization model in prediction of agonist and antagonist muscle forces. *Proceedings of the Human Factors and Ergonomics Society, 2*, 923-927. https://doi.org/10.1177/154193120705101512

Lai, A., Arnold, A. S., & Wakeling, J. M. (2017). Why are Antagonist Muscles Co-activated in My Simulation? A Musculoskeletal Model for Analysing Human Locomotor Tasks. *Annals of biomedical engineering, 45*(12), 2762–2774. https://doi.org/10.1007/s10439-017-1920-7

Le Mansec, Y., Dorel, S., Hug, F., & Jubeau, M. (2018). Lower limb muscle activity during table tennis strokes. *Sports Biomechanics*. https://doi.org/10.1080/14763141.2017.1354064

Lin, L. I. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics, 45*(1), 255-268. https://doi.org/10.2307/2532051

Liu, M. Q., Anderson, F. C., Schwartz, M. H., & Delp, S. L. (2008). Muscle contributions to support and progression over a range of walking speeds. *Journal of Biomechanics, 41*(15), 3243-3252. https://doi.org/10.1016/j.jbiomech.2008.07.031

MacIntosh, A. R., & Keir, P. J. (2017). An open-source model and solution method to predict co-contraction in the finger. *Computer Methods in Biomechanics and Biomedical Engineering, 20*(13), 1373-1381. https://doi.org/10.1080/10255842.2017.1364732

Malagoli Lanzoni, I., Bartolomei, S., Di Michele, R., & Fantozzi, S. (2018). A kinematic comparison between long-line and cross-court top spin forehand in competitive table tennis players. *Journal of Sports Sciences, 36*(23), 2637-2643. https://doi.org/10.1080/02640414.2018.1456394

Malagoli Lanzoni, I., Di Michele, R., & Merni, F. (2014). A notational analysis of shot characteristics in top-level table tennis players. *European Journal of Sport Science, 14*(4), 309-317. https://doi.org/10.1080/17461391.2013.819382

Neptune, R. R., Sasaki, K., & Kautz, S. A. (2008). The effect of walking speed on muscle function and mechanical energetics. *Gait and Posture, 28*(1), 135-143. https://doi.org/10.1016/j.gaitpost.2007.11.004

Qian, J., Zhang, Y., Baker, J. S., & Gu, Y. (2016). Effects of performance level on lower limb kinematics during table tennis forehand loop. *Acta of Bioengineering and Biomechanics, 18*(3), 149-155. https://doi.org/10.5277/ABB-00492-2015-03

Ramsey, D. K., & Wretenberg, P. F. (1999). Biomechanics of the knee: methodological considerations in the in vivo kinematic analysis of the tibiofemoral and patellofemoral joint. *Clinical Biomechanics, 14*(9), 595-611. https://doi.org/10.1016/S0268-0033(99)00015-7

Roelker, S. A., Caruthers, E. J., Hall, R. K., Pelz, N. C., Chaudhari, A. M. W., & Siston, R. A. (2020). Effects of optimization technique on simulated muscle activations and forces. *Journal of Applied Biomechanics, 36*(4), 259-278. https://doi.org/10.1123/JAB.2019-0021

Seemiller, D., & Holowchak, M. (1997). *Winning table tennis: Skills, drills, and strategies*. Champaign, IL: Human Kinetics.

Seth, A., Sherman, M., Reinbolt, J. a., & Delp, S. L. (2011). OpenSim: a musculoskeletal modeling and simulation framework for in silico investigations and exchange. *Procedia IUTAM, 2*, 212-232. https://doi.org/10.1016/j.piutam.2011.04.021

Shao, S., Yu, C., Song, Y., Baker, J. S., Ugbolue, U. C., Lanzoni, I. M., & Gu, Y. (2020). Mechanical character of lower limb for table tennis cross step maneuver. *International Journal of Sports Science & Coaching, 15*(4), 552-561. https://doi.org/10.1177/1747954120922936

Staudenmann, D., Roeleveld, K., Stegeman, D. F., & van Dieën, J. H. (2010). Methodological aspects of SEMG recordings for force estimation – A tutorial and review. *Journal of Electromyography and Kinesiology, 20*(3), 375-387. https://doi.org/10.1016/j.jelekin.2009.08.005

Trinler, U., Leboeuf, F., Hollands, K., Jones, R., & Baker, R. (2018). Estimation of muscle activation during different walking speeds with two mathematical approaches compared to surface EMG. *Gait and Posture, 64*, 266-273. https://doi.org/10.1016/j.gaitpost.2018.06.115

Wang, M., Fu, L., Gu, Y., Mei, Q., Fu, F., & Fernandez, J. (2018). Comparative Study of Kinematics and Muscle Activity between Elite and Amateur Table Tennis Players during Topspin Loop Against Backspin Movements. *Journal of Human Kinetics, 64*(1), 25-33. https://doi.org/10.1515/hukin-2017-0182

Wibawa, A. D., Verdonschot, N., Halbertsma, J. P. K., Burgerhof, J. G. M., Diercks, R. L., & Verkerke, G. J. (2016). Musculoskeletal modeling of human lower limb during normal walking, one-legged forward hopping and side jumping: Comparison of measured EMG and predicted muscle activity patterns. *Journal of Biomechanics, 49*(15), 3660-3666. https://doi.org/10.1016/j.jbiomech.2016.09.041

Zhou, S., Lawson, D. L., Morrison, W. E., & Fairweather, I. (1995). Electromechanical delay in isometric muscle contractions evoked by voluntary, reflex and electrical stimulation. *European Journal of Applied Physiology and Occupational Physiology, 70*(2), 138-145. https://doi.org/10.1007/BF00361541

Żuk, M., Syczewska, M., & Pezowicz, C. (2018). Use of the surface electromyography for a quantitative trend validation of estimated muscle forces. *Biocybernetics and Biomedical Engineering, 38*(2), 243-250. https://doi.org/10.1016/j.bbe.2018.02.001

Int. j. racket sports sci. vol. 4(2), 2022, 11-19. eISSN: 2695-4508
Original Investigation
DOI: 10.30827/Digibug.80314

Reliability and validity of motion sensor and radar for measuring shuttlecock velocity in badminton. *Reliability and validity to measure velocity in badminton*

Fiabilidad y validez del sensor de movimiento y el radar para medir la velocidad del volante en bádminton. *Fiabilidad y validez para medir la velocidad en bádminton*

Michael Phomsoupha [1, 2] *, Florie Bonneau [3, 4] and Guillaume Laffaye [5]

1 M2S (Laboratoire Mouvement, Sport, Santé), Univ. of Rennes, - EA 1274, Rennes, France
2 APCoSS - Institute of Physical Education and Sports Sciences (IFEPSA), Université Catholique de l'Ouest, Angers, France
3 CIAMS, Univ. Paris-Sud, Université Paris-Saclay, 91405 Orsay Cedex, France
4 CIAMS, Université d'Orléans, 45067, Orléans, France
5 Research Center for Sports Science, South Ural State University Chelyabinsk, Russia

Received: 29-11-2021

Accepted: 10-11-2022

Abstract

Radar doppler and inertial measurement unit are often used to analyze the projectile velocity. The aim of the present study was to analyse the reliability and validity of a specifically motion sensor (named: Zepp Tennis) and a radar (Doppler-radar gun) for measuring projectile velocity. Thirty-four (novice, intermediate and expert) stroke badminton smash in a located target. Projectile velocity from five smashes were extracted using Zepp Tennis and Doppler-radar gun data. Between reproducibility of measures was determined by comparing the two sessions. Zepp Tennis and Doppler-radar gun measures were compared with high-frequency video data to establish validity. Both instruments were highly reproducible between trials at different velocity (intra-class correlation coefficient: 0.88-0.94 for radar and 0.78-0.89 for motion sensor). In addition, the positioning of the radar (front of the projectile and angulation) and the placement of the motion sensor and the complexity of the movement (forearm extension and pronation) affect the reproducibility. In terms of validity, radar and motion sensor provides an accurate measure but underestimate projectile velocity (-9.7% and -13.6% respectively).

Keywords: *Shuttle run, performance analysis, ecological validity, lunge.*

Resumen

El radar Doppler y la unidad de medición inercial se utilizan a menudo para analizar la velocidad del proyectil. El objetivo de este estudio fue analizar la fiabilidad y la validez de un sensor de movimiento (denominado Zepp Tennis) y un radar (pistola de radar Doppler) para medir la velocidad del proyectil. Treinta y cuatro jugadores (novatos, intermedios y expertos) realizaron golpes de bádminton en un objetivo localizado. Se extrajo la velocidad del proyectil de cinco golpes utilizando los datos del Zepp Tennis y de la pistola de radar Doppler. La reproducibilidad entre las medidas se determinó comparando las dos sesiones. Las medidas del Zepp Tennis y de la pistola de radar Doppler se compararon con los datos de vídeo de alta frecuencia para establecer su validez. Ambos instrumentos fueron altamente reproducibles entre las pruebas a diferente velocidad (coeficiente de correlación intraclase: 0,88-0,94 para el radar y 0,78-0,89 para el sensor de movimiento). Además, la ubicación del radar (en frente del proyectil y angulación), la ubicación del sensor de movimiento y la complejidad del movimiento (extensión y pronación del antebrazo) afectan a la reproducibilidad. En términos de validez, el radar y el sensor de movimiento proporcionan una medida precisa, pero subestiman la velocidad del proyectil (-9,7% y -13,6% respectivamente).

Palabras clave: *Carrera, análisis del rendimiento, validez ecológica, zancada.*

Corresponding author: Michael Phomsoupha, mphomsou@uco.fr

Cite this article as:
Phomsoupha, M., Bonneau, F., & Laffaye, G. (2022). Reliability and validity of motion sensor and radar for measuring shuttlecock velocity in badminton. *International Journal of Racket Sports Science,* 4(2), 11-19.

INTRODUCTION

The ability to produce a high projectile velocity during a stroke or a shot is one of the main performance factors in sports such as table tennis, tennis, football and baseball. Furthermore, high velocity is correlated with performance in several sports (Laffaye et al., 2012). Projectile can include objects (ball, shuttlecock…) or athletes in flight. In Badminton, the shuttlecock velocity evolved linearly with skill levels (Phomsoupha & Laffaye, 2020).

Especially in badminton, many studies investigated the shuttlecocks velocity to find the biomechanical principles that optimize the motion. In badminton games, the highest velocity is obtained during the smash stroke, which generally allows to finish a rally (Rambely et al., 2005). A recent study showed that during the different Olympic Games finals in men's single, smash is used in about 10 to 14% of the total strokes (Laffaye et al., 2015). The smash can be defined as an aggressive overhead shot with a downward trajectory (Phomsoupha & Laffaye, 2015). At this time, the record of shuttlecock velocity is 493 km/h performed by the Malaysian Tan Boon Heong (Yonex, 2013).

In such a context, assessing the shuttlecock velocity with accuracy is a main stake for athlete monitoring or testing. To record performance, image processing is applied to various fields of sports, such as motion analysis, game analysis (Laffaye et al., 2015), and physical education (Nagasawa et al., 2012). The information including data from the science staff is needed by the player, the supervisor, and the coach to improve the performance (Nagasawa et al., 2012; Takahashi & Kawahara, 2011). Furthermore, this is information feedback of the player's performance.

To record the maximal velocity of the projectile, studies habitually used standard video camera (Hussain & Arshad Bari, 2011; Laffaye et al., 2014). However, this method reveals two main weaknesses. The standard error of measurement depended of the ratio of the launch velocity on the number of frame per second and resolution of the system (Nagasawa et al., 2012). The moment of the peak velocity has to be as close as possible of the moment the shuttlecock quits the racket. The contact between rackets, specifically the string and the ball happens at a very short time (about 5-6 ms) (Miller, 2006). Phomsoupha & Laffaye (2014) showed that the shuttlecock velocity V (in m.s^{-1}) is a logarithmic function of time T (in ms) as follow: V= -9.2 ln(T) +2.4, meaning that the velocity is divided by two times just 0.05 sec after the end of the racket contact. The difference of velocity between elite and high skilled players is about 10 m.s^{-1} (Phomsoupha & Laffaye, 2014), whereas the velocity accuracy of a camera at 50 fps is about 5 m.s^{-1}. This showed that it is impossible to assess a badminton shuttlecock velocity with such devices.

The measurement of the velocity of projectile used in sport games is becoming increasingly common. Such speeds are usually measured using radar guns (Robinson & Robinson, 2016). These devices measure the frequency difference between the reflected signal and the transmitted signal to relate the relative speed of the ball and the radar (Halliday et al., 2011). An increase in frequency of the reflected signal shows the projectile is approaching and a decrease indicates an increase of the distance. Moreover, radar gun devices measure only the radial velocity and will always under-estimate the real velocity (French, 1968; Resnick, 1968).

With advances in microelectronics, wearables recently gained significant attention in sports (Coyle et al., 2009). Specifically, for racket sports, the machine learning methods detect and classify a basic set of shot classes such as forehand and backhand (Petkovic et al., 2001). Moreover, inertial measurement unit (IMU) including accelerometer, gyroscopes, and magnetometer could be used to detect the occurrence of shots (Connaghan et al., 2010). Beyond academic research, few devices have been made and marketed for players and trainers. The most prominent are certainly Babolat Play and Zepp Tennis. An IMU with a wireless transmitting device is attached on the racket handle and data is sent to a smartphone or a tablet for further analysis.

To obtain a better accuracy, some studies used 3D motion analysis system and high speed camera to record the shuttlecock velocity (Domone et al., 2012; Huang et al., 2002; Jaitner & Wolf, 2007; Lee, 1993; Strohmeyer et al., 2009; Tsai et al., 2005, 2006, 2008; Tsai, Chang, et al., 2000; Tsai, Huang, et al., 2000). High speed camera allows that the projectile to evolves in a plan, to avoid parallax error of measurement, whereas 3D motion analysis is free of this kind of error. This allows to obtain an accuracy between 12 and 50 m.s^{-1}, depending on the camera frequency (120 and 500 fps in studies) and it could be considered for the measurement of velocity as the gold standard. However, these devices are expensive for coaches and personal trainers and their use is largely confined to University laboratories and elite sports clubs (Balsalobre-Fernándeza et al., 2015). Furthermore, some of these instruments need specific computer software to analyse the data.

However, to the best of our knowledge, there are no studies validating a motion sensor coupled with a smartphone application or a radar for measuring shuttlecock velocity. The aim of the present study, therefore, was to analyse the validity and reliability of a specifically radar (*Doppler-radar gun*) and a motion sensor or IMU (named: *Zepp Tennis*) for measuring projectile velocity, by comparing with a 'gold standard' measurement system, the Vicon high speed camera system.

MATERIAL AND METHODS

Participants

Thirty-four healthy volunteers (12 novices, 11 intermediates and 11 experts) free of injury (age = 20.1 ± 3.5 years; height = 1.75 ± 10.1 m; body mass = 69.2 ± 13.3 kg; training experienced 8.3 ± 3.1 years) participated in this study (Table 1). Their skills were reflected according to their year of experience and are labelled as followed: novice (lower than 1 year); intermediate (between 3 to 5 years of practice) and experts (more than 5 years of practice). All participants were physically healthy, in good physical condition, and reported no injuries during the time of the study. They were fully informed about the protocol before participating in this study. Informed consent was obtained prior to all testing from all subjects, in accordance with the approval of the local ethical committee and adhered to the latest amendments of the Declaration of Helsinki. The written informed consent was obtained from each participant before experiment. The sample was divided on three groups to obtain different maximal shuttlecock velocity during a smash. The year of practice permit to obtain different velocity.

Table 1.
Age and anthropometric characteristics of the three samples (mean and standard deviation).

Variables	Novice	Intermediate	Expert
Age (years)	24.5 ± 7.6	21.1 ± 4.4	24.4 ± 8.1
Height (cm)	182.3 ± 7.2	179.9 ± 6.3	176.9 ± 9.7
Weight (kg)	76.5 ± 9.8	72.8 ± 9.3	74.3 ± 1.7
Training experienced (years)	0.2 ± 0.7	4.1 ± 1.3	10.6 ± 2.9

Study design

The participant completed a general 10-min warm-up composed of jogging, upper body dynamic stretches and stroke with the racket. Then, each participant performed five badminton smash strokes in a target located in front of him (2m x 2m). During each trial, participants were not informed of their performance. Each smash stroke was separated by 30 sec passive rest period. A shuttlecock was suspended from the ceiling with a string at the player's preferred hitting height. When in contact with the racket, the shuttlecock is pulling away from the celling to produce the trajectory of the smash. No participant expressed residual fatigue from preceding procedure. The experiment took place in two sessions with a minimum of 2 separated days between each one.

Badminton smash strokes

Participants performed badminton smash strokes with the same racket (Wilson Draco Blx; height = 674 mm; weight = 86 g; flexibility = semi-rigid; string tension = 10.5 kg). No instruction was given to the participants on how to proceed during a badminton smash stroke. They were only instructed to stroke as hard as possible.

Equipment

Sports radar. A Doppler-radar gun- Stalker Sport system (Texas, United States) at a frequency of 250 Hz and an accuracy claimed by the constructor of ± 0.027 m/s was used to measure the projectile velocity. The radar permitted to obtain the maximal and the evolution of the velocity during each trial. The experimenter is located 2 meters behind the player in the player-target axis at approximately 2m50 (Chelly & Denis, 2001).

Motion sensor. To record the shuttlecock velocity with *Zepp Tennis*, a mount was attached to the handle of the racket and the sensor was inserted into the mount. The application was designed for analysing the velocity of the racket and the velocity potential of the projectile. Zepp Tennis is available on the Appstore (Apple Inc., USA) and on Google play (Google Inc., USA).

High speed camera reference. The high-speed camera recorded by nine *Vicon V8i* motion capture system at a frequency of 500 Hz (Vicon Peak, Oxford, UK) in order to measure the projectile velocity. A reflective marker of 14 mm diameter was affixed on the front of the shuttlecock. The Vicon system was connected to a PC equipped with the software to analyse and obtain the maximal projectile velocity (Vicon Motion System Ltd., UK). The video-based system is considered as the gold standard reference for establishing concurrent validity of the velocity. The materiel permitted to compare the error of the measure with the other materials.

Statistical analyses

Several analyses were conducted to determine the reliability and validity of badminton smash strokes using the motion sensor and the radar in the present study. To summarise the data from all participants and each trial, descriptive statistics were realised. All data were normally distributed on the basis of Shapiro-Wilk test.

i) Relative reliability is related to the degree to which system maintain their position in a sample with repeated measurements (Atkinson & Nevill, 1998). To analyse the test-retest reliability of both instruments between trials of measurements, intra-class correlation coefficient (ICC) was performed. These coefficients were computed as [ICC = 1- $(SEM/SD)^2$], where SEM is the standard error in measurement and SD is the mean between participant SD of the trial obtained by weighing the variances on the basis of their degrees of freedom (Hopkins, 2000). The SEM was computed as [SEM = SD (between-trial difference in measures) /$\sqrt{2}$]. To analyse the reliability of the motion sensor and the radar when measuring smash stroke of each participant, the coefficient of variation (CV) was used, on the basis of [CV = (SEM/mean)/100%], where the mean takes into consideration all participants and both trials (Atkinson & Nevill, 1998; Hopkins, 2000). To detect systematic bias between trials, Student's t-test was performed (Atkinson & Nevill, 1998).

ii) Data set used concurrent validity has normal distribution. The difference between the materials was examined using a one-way analysis of variance. When a significant F-value was found ($p < 0.05$), the Bonferroni post-hoc was applied (Cohen, 1988). In complement, the bivariate Pearson product moment correlation coefficient (R) was used.

iii) Bland-Altman plots were created, which are known to give a good representation of the agreement between the three instruments (Bland & Altman, 1986). To quantify the statistical dispersion, a White's test was used (White, 1980) to obtain the level of heteroscedasticity.

Concurrent validity was assessed by comparing the mean of trials performed at the maximal performance between two systems. Similar statistical measures to those used to assess reliability were employed for concurrent validity. More precisely, we computed Student's t-test for paired samples (systematic bias), ICC values (relative validity), between-system differences in means (absolute validity in raw units and %) and CVs (absolute validity in %). To obtain a better result, the error size and a maximal error of 5% is considered to be acceptable for a practical application compared to high-speed cameras. On the basis of commonly used thresholds, the relative reliability and validity measures were considered poor, fair and good when the corresponding ICC values were <0.4; 0.4-0.75; >0.75 (Portney & Watkins, 2009). The absolute reliability and validity of measures were considered adequate when the corresponding CV values were equal to or lower than 10% (Stokes, 1985). All calculations were performed using Statistica 10 software (StatSoft Inc., Tulsa, OK), Microsoft Excel 2010 (Microsoft Corp., Redmont, WA, USA) and software R (www.r-project.org).

The intra-session error is free of methodological errors and may be considered as "intrinsic variation" and served as an appropriate baseline for comparisons, remaining independent of other error sources. Intra-session reliability of projectile performance is critically important to ensure that observed differences between testing trials, are not due to systematic bias, such as learning effect, fatigue, or random error due to possible biological or mechanical variation. This variability is usually caused by the emotional state of the participants between the trials and their level of adaptation with the measuring system.

RESULTS

Test-retest reliability

The velocity parameters for each projectile and the mean between-trial difference are reported in Table 2 for the three systems. The between-trial difference in projectile velocity across the level (novice, intermediate and expert) were 5.4 ± 3.7 $m.s^{-1}$, 5.5 ± 2.8 $m.s^{-1}$ and 4.8 ± 3.4 $m.s^{-1}$ for the radar; 6.6 ± 6.4, 4.9 ± 4.2 $m.s^{-1}$ and 5.1 ± 3.4 $m.s^{-1}$ for the motion sensor and 5.2 ± 4.9 $m.s^{-1}$, 4.9 ± 2.9 $m.s^{-1}$ and 5.5 ± 3.0 $m.s^{-1}$ for the high-speed cameras.

The ICC and CV values specific to the reliability at the different test speeds are reported in Table 2 for the three systems. The means of the ICCs was 0.907 ± 0.027 (range 0.88-0.94), 0.840 ± 0.054 (range 0.78-0.89) and 0.940 ± 0.018 (range 0.92-0.96) for the radar, the motion sensor and the high-speed cameras, respectively. Their corresponding mean CV values were 5.8 ± 0.7 (range 5.3-6.6), 7.7 ± 1.8 (range 6.1-9.6) and 4.3 ± 1.0 (range 3.5-5.4). Overall, all three systems demonstrated a good relative and adequate absolute reliability for projectile velocity (Table 3).

Table 2.
Projectile velocity stride parameters calculated using the radar, the motion sensor and the high-speed camera systems.

	Radar			Motion sensor			High speed camera		
Parameter (unit)	Session 1	Session 2	Δ [%]	Session 1	Session 2	Δ (%)	Session 1	Session 2	Δ (%)
				Velocity ($m.s^{-1}$)					
Novice	34.9 ± 8.9	33.3 ± 7.8	4.6	33.6 ± 10.1	33.4 ± 8.4	0.8	37.5 ± 9.1	37.2 ± 7.5	0.9
Intermediate	45.1 ± 6.2	46.2 ± 7.9	-2.6	47.9 ± 4.5	46.6 ± 7.7	2.7	48.1 ± 6.1	49.1 ± 8.9	-1.9
Expert	57.1 ± 9.9	57.9 ± 8.7	-1.5	53.8 ± 8.3	54.4 ± 7.3	-1.3	56.7 ± 9.5	60.4 ± 10.8	-1.2

Note: The mean ± SD for each trial and test velocity, and the difference between trials (Δ, in %) are reported.

Table 3.
The relative (intra-class correlation coefficient, ICC) and absolute (coefficient of variation, CV) reproducibility of projectile velocity stride parameters calculated using the radar, the motion sensor and the high-speed camera system.

	Radar		Motion sensor		High speed camera	
Parameter (unit)	ICC	CV (%)	ICC	CV (%)	ICC	CV (%)
			Velocity ($m.s^{-1}$)			
Novice	0.937	6.6	0.893	9.6	0.940	5.4
Intermediate	0.884	5.5	0.785	7.5	0.921	4.1
Expert	0.901	5.3	0.842	6.1	0.957	3.5

Note: ICC < 0.75 and CV > 10% are italicised and represent fair relative reproducibility and less than adequate absolute reproducibility of measures, respectively.

Also, the Pearson product moment correlation coefficient showed almost perfect correlation between the radar and the high-speed camera measurements for velocity ($r = 0.917$; $p < 0.001$); and good correlation between the motion sensor and the high-speed camera measurements ($r=0.682$; $p < 0.001$) (figure 1 et 2)

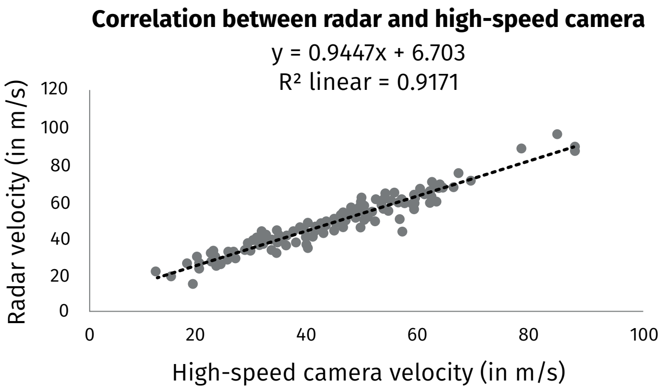


Figure 1. Concurrent validity between radar and high-speed camera.

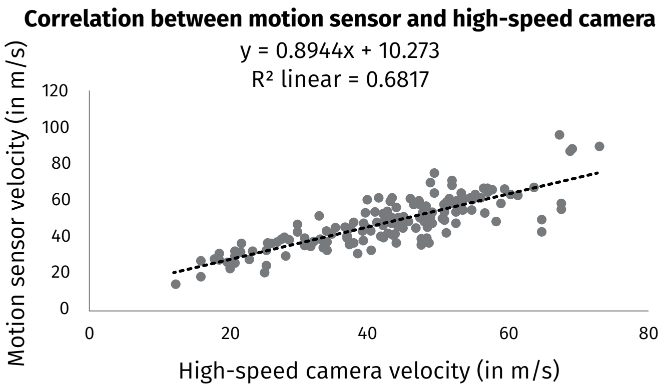


Figure 2. Concurrent validity between motion sensor and high-speed camera.

Validity

The mean of both trials is presented in Table 3 for each system, as are the differences between radar and motion sensor and the reference system (high-speed cameras). Radar and motion sensor recorded significantly shorter velocity compared to the high-speed cameras (Table 4). In contrast, level parameters recorded using the radar and the motion sensor showed no significant differences (all levels, $P > 0.05$).

The ICC and CV values describing the concurrent validity of projectile velocity calculated using the high-speed camera, against the two reference systems are reported in Table 5. The absolute (ICC) concurrent validity of the radar and the motion sensor was overall fair for projectile velocity against the high-speed cameras (0.757 ± 0.207 and 0.866 ± 0.029 respectively). The corresponding relative (CV) concurrent validity measures were higher than adequate (12.5 ± 1.99 and 12.1 ± 0.46%).

Radar and motion sensor values were significantly lower than those obtained with the high-speed camera ($p < 0.05$) (figure 3 and 4).

Table 4.
Projectile velocity stride parameters calculated using the radar, the motion sensor and the high-speed camera systems.

Parameter (units)	Radar	Motion sensor	High speed camera	Radar vs high speed camera	Motion sensor vs high speed camera
	Velocity (m.s^{1})			Δ (%)	
Novice	32.8 ± 8.6	32.2 ± 9.4	37.4 ± 8.3	-14.1*	-16.2**
Intermediate	44.3 ± 7.1	44.1 ± 6.1	48.6 ± 7.1	-9.8	-10.3*
Expert	56.2 ± 9.4	52.8 ± 7.8	60.1 ± 10.1	-6.9	-13.6*

Notes: The mean ± SD for both trials combined at each test speed for each system, and the differences between the high-speed camera and the other two systems (Δ in %) are reported.
*P < 0.05; ** P< 0.01, significant difference between the high-speed camera and the radar or the motion sensor using paired t-tests.

Table 5.
The relative (intra-class correlation coefficient, ICC) and absolute (coefficient of variation, CV) concurrent validity of projectile velocity calculated using the high-speed camera systems against the radar and the motion sensor.

	Radar vs High speed Camera		Motion sensor vs High speed Camera	
Parameter (units)	ICC	CV (%)	ICC	CV (%)
	Velocity ($m.s^{-1}$)			
Novice	0.911	*24.8****	0.627	26.5***
Intermediate	0.841	*16.7****	0.303	14.9***
Expert	0.825	*16.8****	0.295	17.2***

Note: ICC < 0.75 and CV > 10% are italicised and represent fair relative reproducibility and less than adequate absolute reproducibility of measures, respectively.
*** P< 0.001, significant difference between the high-speed camera and the radar or the motion sensor using paired t-tests.

DISCUSSION

The purpose of this study was to analyse the concurrent validity and reliability of a radar (*Doppler-radar gun*) and a motion sensor (*Zepp Tennis*). The radar and motion sensor were reliable, but radar and motion sensor underestimated velocity compared to high-speed cameras. Hence, radar and motion sensor can be considered as a reliable system for computing projectile velocity during a badminton smash stroke ranging from novices to experts. However, motion sensor did not demonstrate good concurrent validity for each level measures and only for novice for radar, warranting caution against the comparisons of results between the high-speed cameras and radar. Intermediate and expert level obtained from the radar proved to be highly reliable and valid compared to our refence systems.

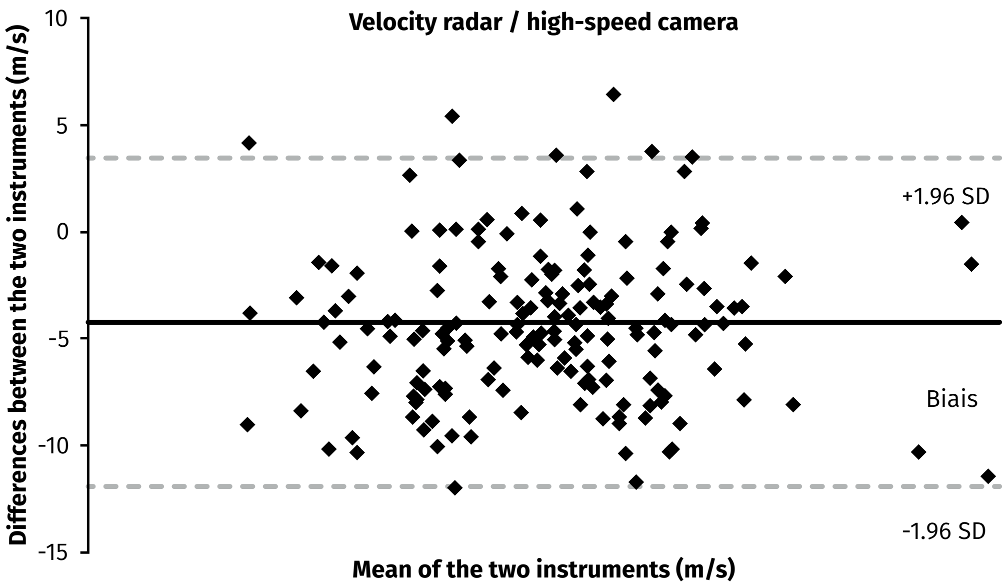


Figure 3. Bland-Altman plots for radar and high-speed camera velocity data. The central line represents the absolute average difference between instruments, while the upper and the lower lines represent ±1.96 SD.

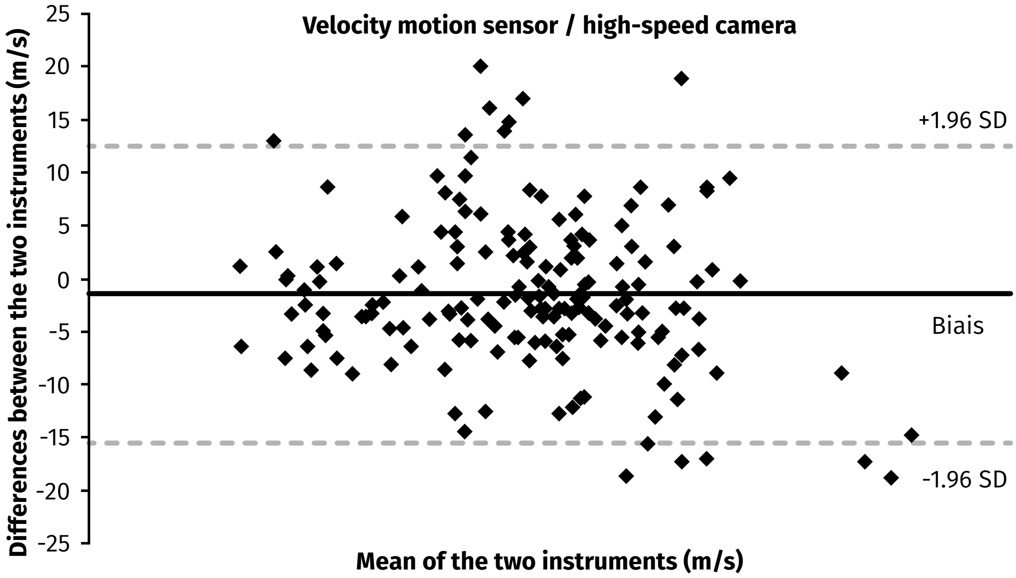


Figure 4. Bland-Altman plots for motion sensor and high-speed camera velocity data. The central line represents the absolute average difference between instruments, while the upper and the lower lines represent ±1.96 SD.

The two systems demonstrated a good homoscedasticity with no significant differences with the White's test on the quantification of the statistical dispersion for radar and for motion sensor (both devices, $p < 0.05$).

The use of several statistical parameters is recommended for quantifying the reliability of measures (Atkinson & Nevill, 1998). In this study, all statistical indicators implied high reliability of velocity derived from the three different systems. The fine distinction in projectile velocity and technique between trials is normal variations expected in any testing situation. Hence the importance of establishing the reproducibility of measures determine which differences exceed typical variations in performance (Gindre et al., 2016). Other than measurement noise, individual variations in stroke biomechanics in the arm movement contributed to the imperfect reliability of measures analysed in the three systems. Kinematics movements of the arm are different and depending of the level that produced different shuttlecock velocity (Phomsoupha & Laffaye, 2020) and contributed to the imperfect reliability of measures in all three systems. The differences in stroke technique between each trial are normal variations expected in any testing situation. The importance of establishing the reproducibility of test measures to determine which differences exceed typical variations in performance (Gindre et al., 2016).

For radar specifically, the between-trial ICCs were all above 0.880; and CVs below 10%. These results tend to highlight that radar recorded the maximal speed during each stroke. Our indicators of reliability of radar device parameters tend to show that the projectile generated higher maximal velocity than racket head (Rambely et al., 2005). In addition, these results tend to highlight that the positioning of the radar (i.e. in front of the projectile and identic angulation) substantially affect the reproducibility of radar parameters (Robinson & Robinson, 2016). For motion sensor specifically, the between-trial ICCs were all above 0.780; and CVs below 10%. These results show that this device has a good reliability. Moreover, neither the positioning of the device (i.e handle or wrist) nor the direction of complexity of the movement (i.e. forearm extension and radio-ulnar pronation) substantially affect the reproducibility of accelerometer derived parameters. The confirmation of the reproducibility of velocity parameters during smash stroke using radar and motion sensor allowed practitioners to be confident in their ability to record these speeds over time using their device.

The concurrent validity of projectile velocity between radar / high-speed cameras and motion sensor / high-speed cameras was fair. On average, radar underestimated projectile velocity by -9.7% and motion sensor by -13.6% between high-speed cameras. These results showed a different way between-system difference in capture velocity and treatment methods. On one hand, the motion sensor consists of an extrapolation of the velocity by the integration formula of the acceleration on three dimensions and the addition of the lever arm. The two major drawbacks of the motion sensor are that there is a possibility that few projectile velocities will not be recorded and the lack of consideration of the racket deflection. The dynamics obtained with a deflexion coupled with a high acceleration of the wrist contribute to racket head velocity (Phomsoupha et al., 2015). Greater flexibility increases the capacity of the racket to store and release more strain energy and to increase the projectile velocity. In addition, the motion sensor seems unable to capture a projectile velocity higher than 325 km/h (≈ 95 m/s). This material is able to measure projectile velocity accurately for all populations, including trained athletes but this is not possible for experts and high speed. There is no requirement and any experience to use and to analyse the data from motion sensor. On the other hand, the radar consists of both a receiver and a transmitter. It sends a radio wave that is reflected of by any object that is in the path. To calculate the speed, the radar gets the echo and uses the principle of Doppler shift. However, the major drawback is the tilt on the sagittal plane which could record the racket instead of the projectile. To ensure better forming results, the radar gun should be positioned near the participant (1 meter at shuttlecock height during the stroke) and the experimenter have to be careful with the recommendation of the manufacturer about the field of angle accuracy. During the experimentation, around 20% of the projectile velocity was not reported. Clinicians and scientist must be aware of these between-system deviations, particularly when comparing results from different studies, laboratory or clinics, and acquiring new equipment for the purpose of quantifying projectile velocity. The data of projectile were homoscedasticity between radar and high-speed cameras and between motion sensor and high-speed cameras.

Considering the low reliability with a low validity of the radar and the motion sensor to measure shuttlecock velocity, correction factors to valid absolute values and to facilitate cross-study comparisons of results may be proposed. Linear regression analyses on our data suggests using the following equation to obtain velocity (xv) from the high-speed camera that are comparable to those from the radar (0.945xv + 6.703) and the motion sensor (0.894xv + 10.273) when individuals velocity ranging from novices to experts (figure 1 and 2).

When analysing the reliability of the motion sensor and radar for measuring the projectile velocity for each participant, the results showed values that were close to the ones obtained with the high-speed camera, despite differences between devices in sampling frequency. Furthermore, the radar and the motion sensor data showed in Bland-Altman plots (Figure 3 and 4) that several of the projectile velocities were close to the mean of the high-speed camera. This is representing a low level of concordance velocities between motion sensor and high-speed cameras (Bland & Altman, 1986). The high ICC showed that motion sensor is no reliability and the results should be tempered and could be increased accuracy with the linear regression.

There are no previous studies that compared different technology for measuring projectile velocity with high-speed camera data. However, some studies used the high-speed camera, which seem to be the best way to record and analyse projectile velocity. This allows to obtain specific values about the highest velocity during the stroke and the time to require it. For the moment, the most accurate systems for measuring projectile velocity are professional and laboratory high speed cameras. This type of camera permits to record at 500 to 1000 Hz compared to 60 Hz for commercial camera. The risk with a standard camera was the maximal velocity could be not recorded during the impact. Thus, experimental data could miss the higher values. Nevertheless, the advancement of new technologies will permit to integrate higher recording frequencies in the future on standard camera (Balsalobre-Fernándeza et al., 2015).

Thus, the orientation has a great impact of the performance data. However, an experience in the use of the radar was required in order to record the correct velocity. This is the first study that validates a motion sensor and radar for measuring the projectile velocity.

CONCLUSION

The ability to evaluate and monitor projectile velocity ability is important in areas of talent identification and sporting performance. The results of the present study showed that projectile velocity can be evaluated using two instruments. Motion sensor could be oriented to the racket sports (tennis, squash…) and golf. Radar was also more efficient for the throwing projectile sports (baseball, football, volleyball…). These findings could help coaches and trainers who wish to monitor the projectile velocity ability of their athletes or clients in a valid and economic way with some ideas of the limitation of each material.

CONFLICT OF INTEREST

All authors have declared there is not any potential conflict of interests concerning this article.

REFERENCES

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine, 26*(4), 217-238.

Balsalobre-Fernándeza, C., Glaisterb, M., & Lockeyb, R. A. (2015). The validity and reliability of an iPhone app for measuring vertical jump performance. *Journal of Sports Science, 33*(15), 1574-1579.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 1*(8476), 307-310.

Chelly, S. M., & Denis, C. (2001). Leg power and hopping stiffness: relationship with sprint running performance. *Medicine and Science in Sports and Exercise, 33*(2), 326-333.

Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (Lawrence E).

Connaghan, D., Conaire, C., Kelly, P., & O'Connor, N. (2010). Recognition of tennis strokes using key postures. *Signals and Systems Conference*, 245-248.

Coyle, S., Morris, D., Lau, K., Diamond, D., & Moyna, N. (2009). Textilebased wearable sensors for assisting sports performance. *Wearable and Implantable Body Sensor Networks*, 307-311.

Domone, S., Wheat, J., Choppin, S., Hamilton, N., & Heller, B. (2012). Wavelet based de-noising of non stationary kinematic signals. *30th Annual Conference of Biomechanics in Sports, 44*, 376-379.

French, A. (1968). *Special Relativity*.

Gindre, C., Lussiana, T., Hebert-Losier, K., & Morin, J. (2016). Reliability and validity of the Myotest® for measuring running stride kinematics. *Journal of Sports Science, 34*(7), 664-670.

Halliday, D., Resnick, R., & Walker, J. (2011). *Fundamentals of Physics* (9th edn).

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine, 30*, 1-15.

Huang, K. S., Shaw-Shiun, C., & Tsai, C. L. (2002). Kinematic analysis of three different badminton backhand overhead strokes. *International Symposium on Biomechanics in Sports*, 200-202.

Hussain, I., & Arshad Bari, M. (2011). Kinematical analysis of forehand and backhand smash in badminton. *Innovative Systems Design and Engineering, 2*(7), 20-26.

Jaitner, T., & Wolf, G. (2007). Analysis of badminton smash with a mobile measure device based on accelerometry. *XXV ISBA Symposium Vol 20*, 282-284.

Laffaye, G., Debanne, T., & Choukou, M. A. (2012). Is the ball velocity dependent on expertise? A multidimensional study in handball. *International Journal of Performance Analysis in Sport, 12*(3), 629-642.

Laffaye, G., Phomsoupha, M., & Dor, F. (2015). Changes in the game characteristics of a badminton match: a longitudinal study through the Olympic Game finals analysis in men's singles. *Journal of Sports Science and Medicine, 14*, 584-590.

Laffaye, G., Wagner, P., & Tombleson, T. (2014). Countermovement jump height: Gender and sport-specific differences in the force-time variables. *The Journal of Strength & Conditioning Research, 28*, 1096-1105.

Lee, B. K. (1993). The effects of the kinematic link principle on performance. *11 International Symposium on Biomechanics in Sports*, 239-242.

Miller, S. (2006). Modern tennis rackets, balls, and surfaces. *British Journal of Sports Medicine, 40*(5), 401-405.

Nagasawa, M., Hatori, Y., Kakuta, M., Hayashi, T., & Sekine, Y. (2012). Smash motion analysis for badminton from image. *Proceedings of the IEEJ Image Electronics and Visual Computing Workshop*, 1-8.

Petkovic, M., Jonker, W., & Zivkovic, Z. (2001). Recognizing Strokes in Tennis Videos Using Hidden Markov Models. *IASTED International Conference on Visualization, Imaging and Image Processing*, 512-516.

Phomsoupha, M., & Laffaye, G. (2014). Shuttlecock velocity during a smash stroke in badminton evolves linearly with skill level. *Computer Methods in Biomechanics and Biomedical Engineering, 17*(Suppl 1), 140-141.

Phomsoupha, M., & Laffaye, G. (2015). The science of badminton: game characteristics, anthropometry, physiology, visual fitness and biomechanics. *Sports Medicine, 45*(4), 473-495.

Phomsoupha, M., & Laffaye, G. (2020). A multiple repeated sprint ability test with four changes of direction for badminton players (part 2): predicting skill level with anthropometry, strength, shuttlecock and displacement velocity. *Journal of Strength and Conditioning Research, 34*(1), 203-211.

Phomsoupha, M., Laffaye, G., Cohen, C., & Clanet, C. (2015). How to use the elasticity of a badminton racket to increase its speed by 80%? *Computer Methods in Biomechanics and Biomedical Engineering, 18*.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: applications to practice* (Upper Sadd). N.J: Pearson/Prentice Hall.

Rambely, A. S., Osman, N. A. A., Usman, J., & Wan Abas, W. A. B. (2005). The contribution of upper limb joints in the development of racket velocity in the badminton smash. *23 International Symposium on Biomechanics in Sports*, 422-426.

Resnick, R. (1968). *Introduction to Special Relativity*.

Robinson, G., & Robinson, I. (2016). Radar speed gun true velocity measurements of sports-balls in flight: application to tennis. *Physica Scripta, 91*(2), 1-19.

Stokes, M. (1985). Reliability and repeatability of methods for measuring muscle in physiotherapy. *Physiotherapy Theory and Practice, 1*(2), 71-76.

Strohmeyer, H. S., Armstrong, C., Litvinsky, Y., Nooney, R., Moore, J., & Smith, K. (2009). Intersegmental coordination differences between beginning performers executing a badminton smash for accuracy of velocity. *27 International Conference on Biomechanics in Sports*.

Takahashi, H., & Kawahara, T. (2011). Conditioning in multi-support project. *The Journal of Japanese Society of Clinical Sports Medicine, 19*(2), 195.

Tsai, C. L., Chang, S. S., & Huang, C. (2000). Biomechanical analysis of differences in the badminton smash and jump smash between taiwan elite and collegiate players. *ISBS'98 - Proceedings II*, 259-262.

Tsai, C. L., Hsueh, Y. C., Pan, K. M., & Chang, S. S. (2008). Biomechanical analysis of different badminton forehand overhead strokes of Taiwan elite femal players. *26 International Symposium on Biomechanics in Sports, Conference, Coaching and Sports Performance*, 719-722.

Tsai, C. L., Huang, C., Lin, D. C., Cheng, C. C., & Lai, C. M. (2000). Biomechanical analysis of the upper extremity in three different badminton overhead strokes. *18th International Symposium on Biomechanics in Sports*, 35-38.

Tsai, C. L., Yang, C. C., Lin, M. S., & Huang, K. S. (2005). The surface EMG activity analysis between badminton smash and jump smash. *International Symposium on Biomechanics in Sports*, 483-486.

Tsai, C. L., Yang, C. C., Lin, M. S., Huang, K. S., & Chang, S. S. (2006). The surface EMG activity of the upper limb muscles of badminton forehand and backhand smashes. *24 International Symposium on Biomechanics in Sports*, 3-6.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817-838.

Yonex. (2013). Nanoray Z speed. https://www.yonex.com/nanoray/#top

DOI: 10.30827/Digibug.80315 Original Investigation

The Oldenburg observation sheet for Table Tennis Technique (O3T) as a tool for talent identification and development: a reliability, validity and feasibility study

Uso de la hoja de verificación de Oldenburg para la técnica en tenis de mesa (O3T) como una herramienta para la identificación y desarrollo del talento: un estudio de fiabilidad, validez y factibilidad

Till Koopmann [1] *, Irene R. Faber [1,2], Dirk Büsch [1] and Jörg Schorer [1]

1Institute of Sport Science, University of Oldenburg, Oldenburg, Germany
2International Table Tennis Federation, Lausanne, Switzerland

Received: 02-03-22

Accepted: 18-11-22

Abstract

Background: The assessment of technical skills as part of a multidimensional approach for talent identification and development in table tennis appears promising. The O3T was developed to assess young table tennis players' technical skills in a highly representative match situation. In this study, two expert coaches (highest coaching license, 25+ years of professional experience) used the O3T to assess the technical skills of 24 young Dutch table tennis players (9 girls, 15 boys; <12 years) based on video recordings. **Results:** Results for variables 'technical quality', 'serve quantity' and 'stroke quantity' were analyzed to assess the O3T's reliability (both inter- and intra-rater), construct validity and feasibility. Bland-Altman-plots and ICCs showed sufficient general reliability with acceptable measurement errors. Variable 'technical quality' showed a moderate relationship (r = .44) with overall table tennis performance at T0 in combination with an increasing trend over time, this way indicating high validity. Finally, the O3T proved to be highly feasible with some possibility to improve based on a feasibility questionnaire. **Conclusions:** Overall, this study presents good prospects for the O3T's measurements properties. In future, the O3T should be used by coaches in various (talent) contexts to further improve its design and to show its added value for talent activities. Furthermore, this approach could be transferred to other performance aspects and sports.

Keywords: *Technique; Talent identification; Talent Development; Children; Racket sports.*

Resumen

Antecedentes: La evaluación de las habilidades técnicas como parte de un enfoque multidimensional para la identificación y el desarrollo de talentos en el tenis de mesa parece prometedora. La O3T se desarrolló para evaluar las habilidades técnicas de jugadores jóvenes de tenis de mesa en una situación de partido altamente representativa. En este estudio, dos entrenadores expertos (licencia de entrenador más alta, más de 25 años de experiencia profesional) utilizaron la O3T para evaluar las habilidades técnicas de 24 jugadores jóvenes holandeses de tenis de mesa (9 niñas, 15 niños; <12 años) basándose en videos. **Resultados:** Se analizaron los resultados de las variables "calidad técnica", "cantidad de saques" y "cantidad de golpes" para evaluar la fiabilidad de la O3T (tanto inter- como intra-evaluador), la validez de constructo y la factibilidad. Los gráficos de Bland-Altman y los CCI mostraron una fiabilidad general suficiente con errores de medición aceptables. La variable "calidad técnica" mostró una relación moderada (r = .44) con el rendimiento general en tenis de mesa en T0 en combinación con una tendencia creciente con el tiempo, indicando así una alta validez. Finalmente, la O3T demostró ser altamente factible con alguna posibilidad de mejora basada en un cuestionario de factibilidad. **Conclusiones:** En general, este estudio presenta buenas perspectivas de las propiedades de medición de la O3T. En el futuro, la O3T debería ser utilizada por entrenadores en diversos contextos (de talento) para mejorar aún más su diseño y demostrar su valor añadido para las actividades de talento. Además, este enfoque podría transferirse a otros aspectos del rendimiento y a otros deportes.

Palabras clave: *Carrera, análisis del rendimiento, validez ecológica, zancada.*

Corresponding author: Till Koopmann, M.Sc., till.koopmann@uol.de

Cite this article as:
Koopmann, T., Faber, I., Büsch, D., & Schorer, J. (2022). Reliability and validity of motion sensor and radar for measuring shuttlecock velocity in badminton. *International Journal of Racket Sports Science,* 4(2), 20-31.

BACKGROUND

Elite sport is embraced by many countries since success in sports, especially at international events, is considered to contribute to a nation's political and economic position, to reinforce the national identity and social cohesion, and to increase positive feelings and well-being (Elling et al., 2014; Grix & Carmichael, 2012; Silva et al., 2020). Talent identification and development (TID) is one of the important pillars of increasing the chances of (future) international sporting success (DeBosscher et al., 2008). Consequently, many countries invest great amounts of resources in talent programs to identify, monitor and develop young and talented athletes (Abernethy, 2008; Vaeyens et al., 2009). National sport associations try to increase these talent programs' effectiveness and efficacy by, among others, improving the talent identification process.

For this purpose, sports associations in cooperation with (embedded) sports scientists attempt to determine performance indicators that help to predict future success (Baker et al., 2011, 2021; Baker et al., 2017; Faber, Bustin, et al., 2016; Johnston et al., 2018). Previous research revealed various indicators that discriminate between playing levels and/or predict future performance in various sports, e.g., anthropometrics, physiological parameters, and sport-specific technical, mental and goal-management skills (Elferink-Gemser et al., 2007; Faber, Bustin, et al., 2016; Huijgen et al., 2012; Johnston et al., 2018). It is suggested that specifically the sport-specific technical skills are of great importance within a multidimensional skill set to be able to reach the elite level in various sports due to their close relation to the highly demanding and specialized proficiencies required for elite performance (Glazier, 2017; Kolman et al., 2019; Koopmann et al., 2020). An optimal technical skill development provides an athlete with the best opportunities to use the full range of technical and tactical solutions (Kannekens et al., 2011). This is the case especially in sports that rely highly on technical proficiency.

Table tennis is a typical example of a technique-based sport. Players aiming for the elite level must develop outstanding technical skills including fast stroke technique adjustments, variable, flexible and fast footwork, a pronounced ability to anticipate and react, proper positioning skills, and balance control (Ak & Koçak, 2010; Akpinar et al., 2012; Faber et al., 2021; Friedrich & Fürste, 2015; Sève et al., 2002). Accordingly, the technical development is emphasized by trainers/coaches from the very beginning of a player's career. Technical skills are considered crucial for early development and the age-span of 8 to 12 years represents an important window of opportunity for high potential youth players to develop their technical skills as a fundament of reaching the elite level (Anderson et al., 2012; Huber et al., 2009; Table Tennis Canada, 2015). Early mistakes hindering a player's technical development should be prevented as much as possible (Friedrich & Fürste, 2015).

Despite the significance of technical skills, instruments that can evaluate the technical skill level for TID purposes appear scarce in table tennis (Faber et al., 2021). Moreover, the existing instruments show shortcomings in three areas: (1) they mostly cover only single technical elements although the interplay of various elements appears crucial, (2) they use outcome-related method types (Koopmann et al., 2020) with a focus on ball speed and ball placement without recognizing the movement patterns, and (3) they focus on adult elite players and not on the identification and development of young talented players (Faber, Bustin, et al., 2016; Kolman et al., 2019). For that reason, the Oldenburg observation sheet for Table Tennis Technique (O3T) was developed as an instrument for the representative assessment of technical skills in youth table tennis players (8-12 years) while addressing the shortcomings mentioned above and allowing for the application in talent contexts (Figure 1).

The O3T was constructed based on both scientific and professional literature as well as expert interviews (Faber et al., 2021). It acknowledges two general elements (i.e., individuality and interconnection bet-ween elements) and includes two quantitative (i.e., serve quantity, stroke quantity) as well as five qualitative elements (i.e., bat grip, ready position, footwork/body positioning, serve, and stroke) of technical skills. In addition, criteria for both flawed and excellent executions in talented young players (≤12 years) were identified and described for each of the five specific qualitative elements to improve objectivity. The O3T is set up to be applied in talent contexts including the observation of competitive table tennis matches. This entails the application in training selections camps to identify high potential youth players and select them for intensive training programs.

The present study was designed to evaluate the O3T's measurement properties to ensure its utility and added value in future practical scenarios (Bowen et al., 2009; Morrow Jr et al., 2015). Here, as the concept of talent in sports is longitudinal and as TID activities are always teamwork and include various coaches and other stakeholders, the O3T must be usable reliably by different people leading to similar rankings of players. Accordingly, reliability (both intra- and inter-rater) must be assessed for all measures. Reliability of quantity measures was hypothesized to be at an acceptable level as the categorizations of different strokes and serves is expected to be straight forward. Reliability for variable 'technical quality' was also hypothesized to be at a sufficient level due to standardization of observations and the clear descriptions for both flawed and excellent executions based on practice-based evidence. Furthermore, validity was investigated by examining the relationship between the O3T outcomes and the players' current and future table tennis performance outcomes. Here, despite technical skills being an important element, similar to previous studies focusing on the predictive value of perceptuo-motor skills tests (Faber, Elferink-Gemser, et al., 2016; Faber et al., 2017; Faber et al., 2015), only moderate associations were expected as technical skills are only one out of multiple elements in a multidimensional profile which explains overall table tennis performance. Finally, *feasibility* was assessed in detail as the O3T will be used by various people in stressful TID situations with tight time schedules. Thus, it must be easy and intuitive to use following a clear structure while delivering comprehensive and meaningful data. Feasibility is hypothesized to be at a high level since the O3T was developed in very close cooperation with the practitioners and project partners of the German Table Tennis Association (Deutscher Tischtennis-Bund e.V., DTTB).

METHODS

The design of the present study was three-fold and followed the COSMIN guidelines (COnsensus-based Standards for the selection of health status Measurement INstruments) (Mokkink et al., 2010). First, a test-retest design with two expert coaches rating the technical skill level of youth players based on video recordings of table tennis matches was applied to examine both intra- and inter-rater *reliability*. Second, associations between the technical skill ratings based on the initial video-observations and players' table tennis performance at three points in time (i.e., at the moment of video recording and one and two years later) were determined for construct validity (Mokkink et al., 2010). Third, *feasibility* was evaluated using a feasibility questionnaire. All procedures were in full compliance with the Declaration of Helsinki and approved by the ethical committee of the Carl von Ossietzky University Oldenburg in Germany (Reference: Drs.EK/2020/040).

Players

A total of 24 young Dutch table tennis players (9 girls and 15 boys; <12 years) playing matches at international tournaments between the years 2013 and 2019 were randomly selected based on the available video archives of the Netherlands Table Tennis Association (Nederlandse Tafeltennisbond, NTTB). These young players were regarded as most talented of their age-group by national coaches and thus selected to represent their country at international tournaments.

Assessing coaches

Two expert coaches were appointed in consultation with the DTTB. Both coaches held at least the highest German coaching certification (A-license), had at least 25 years of professional coaching experience at the highest level in mainly German but also in international table tennis, and had substantial experience with the education and guidance of specifically young players of the highest level (e.g., as national coaches for youth players).

Instruments and variables

The Oldenburg observation sheet for Table Tennis Technique (O3T)

The newly developed O3T was used to rate the technical skills level of youth table tennis players (Faber et al., 2021). The expert coaches were instructed to mark all serve and stroke techniques that were demonstrated by the player during the match to assess the quantity of technique. For the quality, they were instructed to rate the technique level between 1 (lowest/worst value) to 10 (highest/best value) regarding bat grip, ready positioning, foot work/body positioning, serves and stroke while using descriptions of flawed and excellent executions as a guideline to ensure objectivity (see Figure 1). Technical skill variables 'serve quantity' (number of different serve types shown by a player), 'stroke quantity' (number of different stroke types shown by a player) and 'technical quality' (mean of the five quality ratings; Cronbach's alpha .81) were calculated for each video based on the respective O3T's filled out by the coaches.

Competition rating score

Competition rating scores (i.e., ELO-rating) were obtained from the NTTB online archives (https://www.nttb-ranglijsten.nl/) for each player as an indicator of overall table tennis performance. The rating scores indicate the player's individual competition performance at a specific moment in time. The higher the rating score the better is the player's table tennis performance. It allows for comparison of all players (e.g., youth and adult players, male and female players) that participate in any of the Dutch regional and national competition leagues (Faber et al., 2021).

Feasibility questionnaire

Feasibility was assessed using a specifically developed feasibility questionnaire based on other questionnaires and guidelines (Bowen et al., 2009; Robertson et al., 2017) including eleven questions with a focus on the O3T's added value for coaches, its design and structure, its completeness in terms of elements, and the time needed to use it. Questions 4 and 11 used an open format asking for number of minutes and general feedback points, respectively. All other questions were rated on a scale of 1 (lowest/worst value) to 10 (highest/best value) and coaches were able to give additional feedback in an open format.

INSTRUCTION

The observation sheet is designed to evaluate essential technical skills in youth table tennis players between 8-12 years old. It is specifically intended to indicate the level of technical skills at the moment of the observation and not the player's future potential. In addition to this, it is important to mention that by using this observation sheet, it is never intended to follow a 'one-size-fits-all' approach. Technique is considered to be individual and should optimally be a fit to the player's mental and physical characteristics. Moreover, it should be acknowledged that all aspects mentioned in the observation sheet are connected to and influence each other. The evaluation of the connection between technical skills and other performance related aspects like tactics and perception/anticipation is planned to be part of a future project.

QUANTITY OF TECHNICAL SKILLS

Please mark (X) all techniques that were demonstrated by the player during the observed match.

Service

Backhand	contra	backspin	sidespin	topspin	back-sidespin	top-sidespin
Forehand	contra	backspin	sidespin	topspin	back-sidespin	top-sidespin

Stroke

Backhand	contra	push	topspin/drive	block	smash	flick	defense	balloon defense
Forehand	contra	push	topspin/drive	block	smash	flick	defense	balloon defense

QUALITY OF TECHNICAL SKILLS

Please evaluate the player's performance at this moment for each technical aspect by marking (X) the most appropriate score based on the descriptions on the right as a guideline (marked green = excellent performance; marked red = flawed performance). Make sure you always encircle only one number.

Bat grip

1	2	3	4	5	6	7	8	9	10

Ready position

1	2	3	4	5	6	7	8	9	10

Footwork/Body positioning

1	2	3	4	5	6	7	8	9	10

Service

1	2	3	4	5	6	7	8	9	10

Stroke

1	2	3	4	5	6	7	8	9	10

Bat grip

Red	Green
Fixed backhand or forehand grip while NOT using the advantages of that grip.	Mainly neutral grip with small adaptations to use advantages during play.

Ready position

Red	Green
The player shows a highly tensed and fixed position. The center of gravity is not in front of the body; the player leans on whole feet or heels. The player stands straight up with arms hanging aside the body.	The player's body is functionally tensed. The center of gravity is in front of the body while standing on the forefeet. The player has an active position with the whole body with knees and hips bend and holding the bat in front of the body and above the table.

Footwork/Body positioning

Red	Green
Foot position is like 'penguin-feet' or positioned too close/broad and not aligned to the upcoming ball. The body is often unbalanced and in extreme positions. The player shows slow and inefficient footwork leading to delayed and improper body positioning.	Foot position is aligned towards the upcoming ball. The player easily restores balance disturbances and is (almost) never in extreme positions. The player shows fast and efficient footwork creating optimal conditions for the next stroke. The player moves in multiple dimensions: left-right and forward-backward.

Service

Red	Green
The player shows no variation within services (speed, placements, rhythm, rotation).	The player uses different variation within services (speed, placements, rhythm, rotation).

Stroke

Red	Green
The hitting point is mostly aside the body. The kinematic chain: segments are not used/sequenced in the optimal order or contra-act towards the intended stroke. The player does not adapt to the upcoming ball and/or shows similar errors frequently. The player has no variation within strokes (speed, placements, rhythm, rotation). Strokes are not automated and reproducibility is low.	The hitting point is always in front of the body. The kinematic chain: all segments are optimally used (bottom-up). The player adapts to the upcoming ball (e.g., stroke length, bat position, type of stroke). The player uses different variation within strokes (speed, placements, rhythm, rotation) and/or uses different solutions. Strokes are automated and can be produced with high reproducibility.

Figure 1. Oldenburg observation sheet for Table Tennis Technique (O3T).

Data collection

Two videos against two different opponents were available for each youth player included, showing (at least) three games per match (two with a back-view and one with a front-view). Written parental informed consent was obtained prior to using the videos for study purposes. The expert coaches evaluated the players' technical skills level based on video observations. Each coach watched the videos by himself on a personal laptop on two different occasions (7-10 days apart). During the first assessment, both coaches watched the first video of all 24 players in the same order. During the second assessment, both coaches watched the first video of half of the players again and the second video of the other half of the players in random order. Here, it was ensured that videos of the second assessment were not the same for the two coaches and that the coaches were blinded for the technical skill ratings of the first day. Competition rating scores were obtained for each player from the Netherlands Table Tennis Association's open archives from the moment of video recording and one and two years later. The feasibility questionnaire was filled in within four weeks after the second assessment day by the expert coaches.

Statistical analysis

All statistical analyses were conducted using IBM SPSS Statistics 27 (IBM Corp., Armonk, New York, United States of America) and a level of significance of $\alpha = .05$. The normal distribution of the data was checked by comparing (1) the outcomes' means and medians, (2) their spread around the mean (standard deviations) as well as ranges, and (3) the Shapiro-Wilk test for normality. Potential sex effects were checked by comparing female and male players using a Mann Whitney U test for non-normal distributed outcomes and an independent sample t-test for normal distributed outcomes.

Bland-Altman plots were created for all three technical skill variables. For variables 'serve quantity' and 'stroke quantity' medians and 2.5 and 97.5 percentile boundaries were used due their non-normal distribution whereas means and standard deviations were used for variable 'technical quality' as the data were normally distributed. In addition, intraclass correlation coefficients (ICCs) and their 95% confidence intervals could be calculated as reliability outcome for the variable 'technical quality' due its normal data distribution (Weir, 2005). A one-way random model was used to calculate the ICC for the *intra-rater reliability* based on the two ratings by the same coach of the same players based on the same video. A two-way random model (type consistency) was used to calculate the ICC for intra-rater reliability based on the two ratings by the same coach of the same players but different videos. The latter model was also used for the ICCs for the inter-rater reliability and calculated from the ratings of the two coaches based on the same video, and ratings by the two coaches based on two different videos. ICCs were interpreted as acceptable when ≥ .70 (Prinsen et al., 2018). Furthermore, the standard error of measurement (SEm), the smallest detectable difference (SDD) and the coefficient of variation (CV) were also calculated for the normally distributed 'technical quality' data (de Vet et al., 2006; Hopkins, 2000).

To check the O3T's construct validity, associations between the technical skill variables 'serve quantity', 'stroke quantity' and 'technical quality' and the players' competition rating scores at the moment of video recording (T0) and one (T1) and two years (T2) later were analyzed using Spearman and Pearson correlation coefficients for non-normally and normally distributed data, respectively. Here, technical skill outcomes were calculated as the mean of both coaches' ratings for the respective variable based on the first videos. Correlation coefficients (r) were interpreted as small (r = .10 - .29), medium (r = .30 - .49) and large (r > .50) associations (Cohen, 1988).

The feedback of the expert coaches obtained from the feasibility questionnaire was assessed analyzing the range of scores and the mean rating for each question. The open question on general remarks was analyzed based on the qualitative descriptive data. All feasibility feedback is presented descriptively.

RESULTS

Descriptives

Table 1 shows the descriptive statistics for the technical assessment of all players based on the first video in combination with their competition rating scores at the respective time. As the data for variables 'stroke quantity' and 'serve quantity' were not normally distributed, medians and ranges are reported while means and standard deviations are presented for variable 'technical quality'. Comparing female and male players, the data show no statistically significant sex effect (p > .05). Looking at the competition rating scores, all players show a trend of improvement while there are variations in these improvements over time with the lowest change from T0 to T2 being 246 points (Player #14) and the highest being 777 points (Player #2).

Reliability

Figure 2 shows the Bland-Altman plots for the *intra-rater reliability* for all three technical skill variables including also the ICCs for variable 'technical quality'. Regarding variable 'serve quantity' the systematic error (i.e., median difference) was rounded 0 points when observing the same or the other video with a random error between 2 to 3 points (Figure 2A+B). This also was the case for variable 'stroke quantity' when observing the same video twice (Figure 2C). In case of an evaluation by the same coach with two videos of the same players, the systematic error was 2 points with a random error between 4 to 5 points (Figure 2D). The systematic error for variable 'technical quality' (i.e., mean difference) was .1 points with a random error between 1-2 points when the same video was watched twice (Figure 2E). When the same coach watched another video of a player, the systematic error was .4 points and the random error 1 to 2 points (Figure 2F). The ICCs for variable 'technical quality' showed acceptable reliability (> .70) with SDDs of 1.6 and 1.7 with CVs of 10.0 and 10.9 %, respectively (Figure 2F).

The Bland-Altman plots for the inter-rater reliability for all three technical skill variables including ICCs for variable 'technical quality' are presented in Figure 3. The systematic error for variable 'serve quantity' based on the same video was rounded 0 points with a random error between 2 and 3.5 points (Figure 3A). The systematic error for variable 'serve quantity' based on the other video was 1 point with a random error of approximately 2.5 (Figure 3B). For variable 'stroke quantity', systematic errors of rounded 1 and 2 points with random errors of between 3.5 and 4.5 points were found (Figure 3C+D). The systematic error for variable 'technical quality' was 1.3 points with a random error of approximately 1.5 points when the same video was watched twice (Figure 3E). When the same coach rated the 'technical quality' based on another video, the systematic error was 1 point and the random error approximately 1.9 points (Figure 3F). The ICC for variable 'technical quality' based on the observation of one video was on an acceptable level with a SDD of 1.5 and a CV of 9.6 % (Figure 3C). However, the ICC based on two videos dropped below the cut-off value and therefore was considered insufficient with a SDD of 1.9 and CV of 11.8 %.

Table 1. Descriptive statistics for the technical assessment based on the first video observations and the competition rating scores.

Players	Age (years)	Coach A			Coach B			Competition rating score		
		Serve quantity	Stroke quantity	Technical quality	Serve quantity	Stroke quantity	Technical quality	T0	T1#	T2#
Male										
#1	9	2	7	7.4	2	9	8.0	1148	1482	1878
#2	9	4	6	3.8	4	7	5.4	754	1208	1531
#3	9	3	7	4.4	3	6	5.8	777	1052	1164
#4	10	5	7	3.8	3	8	6.2	947	1270	1525
#5	10	1	5	4.0	2	7	5.4	1217	-	-
#6	11	4	8	6.0	3	8	7.2	1082	1400	1663
#7	11	2	7	5.6	4	8	7.4	1046	1209	-
#8	11	3	8	5.4	3	7	5.6	1135	1342	1437
#9	11	4	7	5.8	3	11	6.3	1132	1396	1609
#10	11	3	7	5.8	2	8	6.8	1080	1320	1361
#11	11	2	8	6.2	2	10	7.2	1138	1451	1623
#12	12	2	7	6.2	2	6	6.0	1664	1815	2038
#13	12	3	8	5.2	3	5	6.6	1110	1397	1593
#14	12	8	5	4.6	4	3	6.0	1151	1316	1397
#15	12	4	7	4.2	2	4	5.6	1400	-	-
Female										
#1	10	8	8	6.8	5	8	6.4	1062	1212	1474
#2	10	2	5	5.0	4	8	6.0	644	983	1200
#3	10	2	5	2.8	2	5	4.8	666	1040	-
#4	10	7	7	4.6	4	8	5.6	620	660	-
#5	11	5	6	4.6	5	10	6.4	944	1118	1290
#6	11	1	10	4.8	1	11	6.4	972	1280	1596
#7	11	3	6	4.4	3	7	6.6	642	711	961
#8	11	5	7	5.0	3	6	6.8	1077	1204	1590
#9	11	4	7	3.0	3	6	5.8	674	1052	1126
	Mean (SD)	Median (range)	Median (range)	Mean (SD)	Median (range)	Median (range)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Boys (n = 15)	10.7 (1.1)	3 (1-8)	7 (5-8)	5.2 (1.1)	3 (2-4)	7 (3-11)	6.4 (0.8)	1118 (220)	1358 (180)	1568 (231)
Girls (n = 9)	10.6 (0.5)	4 (1-8)	7 (5-10)	4.6 (1.2)	3 (1-5)	8 (5-11)	6.1 (0.6)	811 (197)	1029 (217)	1320 (243)
Total (n = 24)	10.7 (0.9)	3 (1-8)	7 (5-10)	5.0 (1.1)	3 (1-5)	7.5 (3-11)	6.3 (0.7)	1003 (257)	1224 (253)	1477 (260)

T0 time of video recording; T1 one year after video recording; T2 two years after video recording; SD standard deviation
missing values (-) are due to players dropping out of the sport.

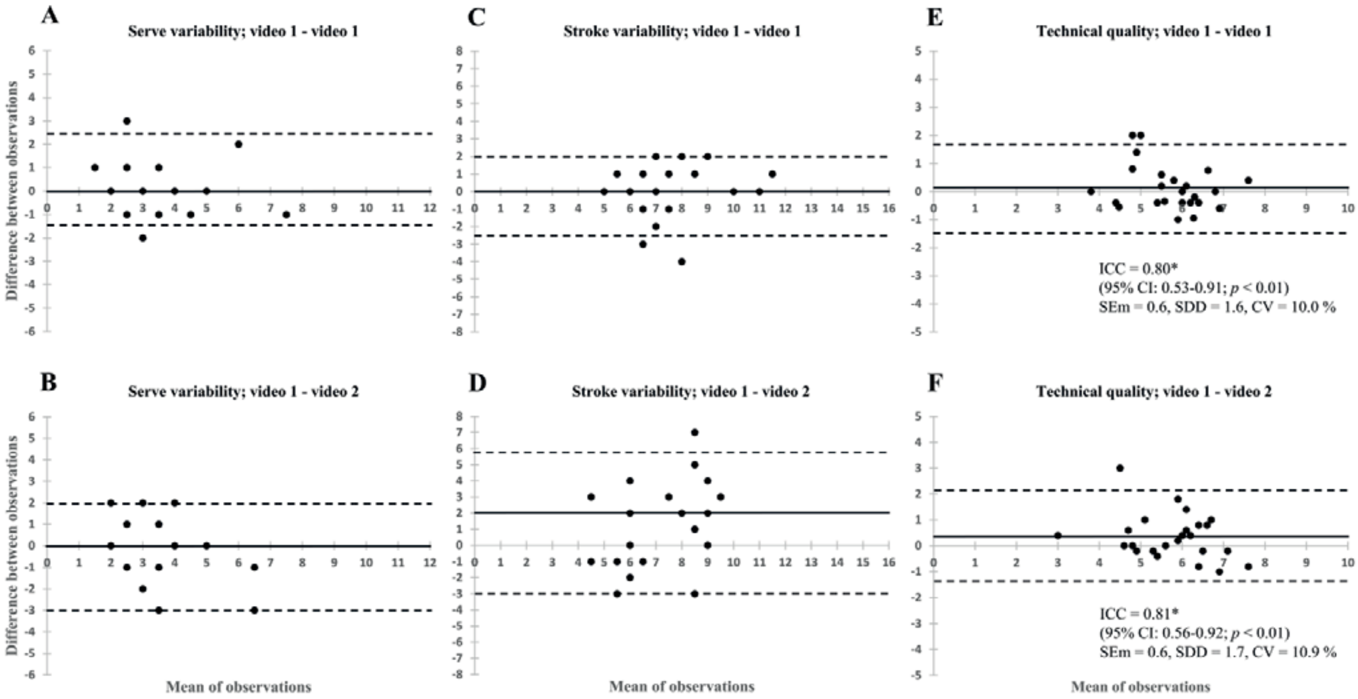


Figure 2. Bland-Altman plots for the intra-rater reliability for all three technical variables.
Legend: Solid line Median value for 'serve quantity' and 'stroke quantity', mean value for 'technical quality'. Dashed line 2.5/97.5 percentiles for 'serve quantity' and 'stroke quantity', 1.96*SD limits for 'technical quality'. CI confidence interval.

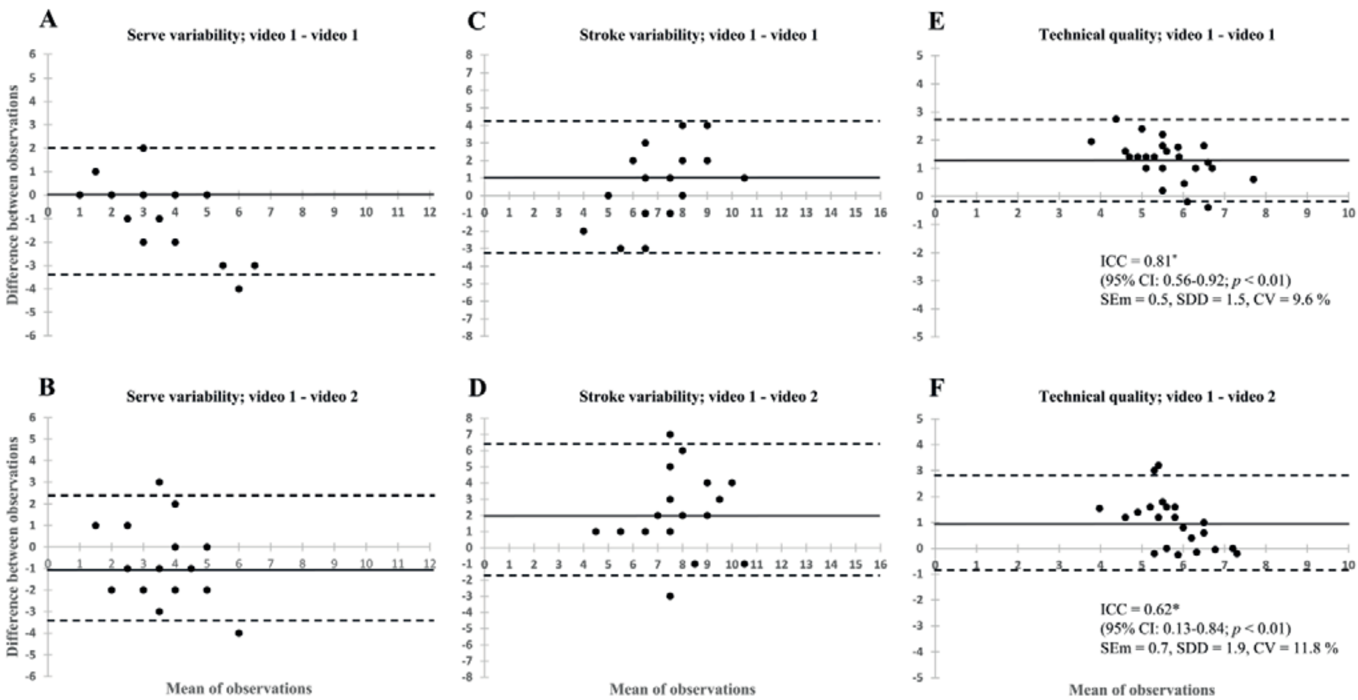


Figure 3. Bland-Altman plots for the inter-rater reliability for all three technical variables.
Legend: Solid line Median value for 'serve quantity' and 'stroke quantity', mean value for 'technical quality'. Dashed line 2.5/97.5 percentiles for 'serve quantity' and 'stroke quantity', 1.96*SD-limits for 'technical quality'. CI confidence interval

Construct Validity

Table 2 shows the results of correlation analysis between the three technical skill variables and competition rating scores at the moment of video recording and one and two years later. The correlation coefficients for variable 'serve quantity' showed weak (T0) and moderate associations (T1, T2) while no (T0) and moderate associations (T1, T2) were found for variable 'stroke quantity'. Correlation coefficients for variable 'technical quality' showed a moderate relationship at T0 that is then by trend increasing and showing large associations at T1 and T2.

Table 2. Relationship between O3T outcomes and competition rating scores (T0, T1 and T2).

	Competition rating score		
	T0 (n = 24)	T1 (n = 22)	T2 (n = 19)
Serve quantity[1]	-.29 (.17)	-.33 (.13)	-.32 (.18)
Stroke quantity[1]	.00 (1.00)	.41 (.06)	.45 (.05)
Technical quality[2]	.44* (.03)	.51* (.02)	.58* (.01)

T0 time of video recording; T1 one year after video recording; T2 two years after video recording.
[1] Spearman rank-order correlation coefficients (p-value).
[2] Pearson correlation coefficients (p-value).
* Statistically significant ($p < 0.05$).

Feasibility

Table 3 shows the results of the feasibility questionnaire for both assessing coaches. Overall, results show high ratings (> 8) for feasibility with slightly lower ratings only for questions 3, 7 and 8. First of all, the feasibility feedback by the two coaches showed that there is a need for a comprehensive tool for the assessment of technical skills in youth players (question 1,). Furthermore, the newly developed O3T is understandable (question 2) and covers all important elements of technical skills well (question 3). However, for the latter question slightly lower scores were rated. Furthermore, coaches needed approximately 15 minutes (question 4) to fill out the O3T and rated this duration as perfectly acceptable (question 5). While both coaches agreed that the O3T could be implemented in the talent selection procedures without problems (question 6), they were slightly divided regarding the question of using the O3T in future camps (ratings of 6 vs. 10; question 9) and recommending its use to other coaches (ratings of 6 vs. 10; question 10).

DISCUSSION

This study presents good prospects for the O3T's measurement properties and its added value in talent contexts. Firstly, reliability appears sufficient with acceptable measurement errors (ICCs > .70, SDD 1.5-1.9, CV 9.6-11.8 %). Only the ICC of .62 for inter-rater reliability for 'technical quality' based on the observation of two different video was below the cut-off value (.70). Secondly, variable 'technical quality' showed a moderate relationship ($r = .44$) with overall table tennis performance at T0 in combination with an increasing trend over time ($r = .51$ and $.58$ for T1 and T2, respectively). Thirdly, the O3T overall appears to be highly feasible with some possibility to improve regarding its structure and its use in real-world talent selection contexts. The O3T's reliability, validity and feasibility appear comparably high compared to similar instruments (Faber et al., 2015; Katsikadelis et al., 2014; Van Biesen et al., 2012) while it includes more detailed items and allows for the assessment of young children's technical skills in a highly representative setting during competition matches. Nevertheless, it is important to consider some important factors for a fair interpretation when using the O3T in practice.

Table 3. Coaches' feasibility feedback.

Question	Coach 1	Coach 2	Average rating
1. Do you see a need for an observation sheet for technical skills for all German table tennis coaches?	8	10	9.0
2. Is the observation sheet understandable and comprehensible?	10	7	8.5
3. Does the observation sheet cover all important elements of technical skills in table tennis?	6	7	7.5
4. How long did it take you to fill out the observation sheet on average? (in minutes)	15	15	15.0
5. Was the time needed to fill out the observation sheet acceptable?	10	10	10.0
6. Would the observation sheet easily be implemented in the talent camp process?	10	8	9.0
7. Could you imagine using observations sheets also for other aspects of table tennis performance (e.g., tactical skills)?	10	5	7.5
8. How much did the observation sheet help you to assess the players' technical skills?	6	7	6.5
9. Would you like to use the observation sheet again at future talent camps?	6	10	8.0
10. Would you recommend the use of the observation sheet to other coaches?	6	10	8.0
11. General feedback:	• Opponent's performance/skill level must be considered. • Match results (win/lose) must be considered. • Training age must be considered. • Quality is more important than quantity. • Rating scales of 1-5 are sufficient. • Elements of each technique should be presented.		

Note All questions besides questions 4 and 11 were rated using a 1-10 rating scale. Questions 4 and 11 followed an open format asking for number of minutes and general feedback points, respectively.

First, using sum scores for both quantity and quality measures presumes the different elements being equally important. For example, a player's technical quality score was calculated as the sum of the respective bat grip, ready position, footwork/body positioning, service and stroke scores. While this procedure may not be without flaws especially for reliability measures, it is common practice in research on sports with compensatory performance profiles (Faber, Elferink-Gemser, et al., 2016; Van Biesen et al., 2012). In future, associations between and weightings of the various elements should be investigated.

Second, although coaches showed a remarkable consensus on the most crucial aspects of technical skills and on their flawed and optimal execution (Faber et al., 2021), it seems that the perception of the demonstrated technical quantity and quality can differ. Here, the questions of where to look and how to look at it should be debated. That is, strategies and approaches to assess and perceive technical skills in practice should be a key theme in coach decision and education. In terms of the statistical analysis used in the present article, calculating the ICCs regarding consistency is fine for solely ranking players. However, when the O3T is used for monitoring players over time, better agreement is needed as may be facilitated with a detailed rubric. Here, specific examples regarding the flawed and excellent executions should be presented (see also the 'Technikleitbild' by the DTTB; https://www.tischtennis.de/technikleitbild.html) and discussed in detail as would be possible with both live and video tutorials that could be implemented in coach education programs. These discussions and potential agreements between coaches may then further increase reliability when using the O3T.

Third, the assessment of technical skills during a match seems inseparable to the evaluation of a player's tactical strategies which has an influence on both the reliability and validity of the assessment (Kannekens et al., 2011). Players may show rather low numbers of different technical variations in case these few variations work successfully and effectively towards winning a match against a respective opponent (see also feasibility feedback, Table 3). That is, assuming a better (i.e., 'more talented') player having the ability to show a higher technique quantity (i.e. variability) is not the same as assuming this player to show this higher technique quantity (i.e. variability) in every match as this is closely connected to the tactical strategies and skills. Accordingly, 'technical quality' may be more important than quantity as measured here in this context (see also feasibility feedback, Table 3) while being able to use various techniques in general still appears beneficial and desirable. Here, technique quantity's 'appropriateness' should be considered when interpreting the O3T results and assessing players.

Finally, the O3T was developed to be part of a multidimensional assessment for talent identification (and development). Accordingly, other crucial aspects of performance and talent must be assessed, and talent selection decisions should be made based on a player's overall profile as weaknesses in one area can be compensated for by strengths in others (Baker et al., 2020; Elferink-Gemser et al., 2011; Faber, 2016). Here, the increasing statistical trend in the longitudinal data may be of high value in the context of talent research as the goal is the identification of factors determining not current but future performance (Johnston et al., 2018). Thus, this trend must be investigated further as it may indicate a high technical skill level in youth to be the base for future high level table tennis performance. If verified, this information would be highly powerful for TID purposes in table tennis. Furthermore, it is important to note that performance measures in children generally vary more compared to those in adults because of differences in growth or training experience (Deutsch & Newell, 2005). This could be an influencing factor especially for the *reliability* measures based on two different videos as players may have shown different levels of technical skills so that lower reliability values would be true and correct and not due to shortcomings in the O3T's design. In addition, these variances in performance measures of children may have also influenced the *validity* assessment. Thus, to adjust for these variations it is recommended to measure children multiple times and, e.g., take the mean of all measurements for the analysis and evaluation.

Three more limitations of the present study need to be acknowledged. First, assessing technical skills only based on video observations lacks representativeness and potentially has introduced some error to our data while it at the same time helped to control other factors, e.g., coaches' visual perspectives and occurrences of match actions. Second, we only used video material from Dutch young players playing at international tournaments. Including talented players from other nationalities would improve generalizability of the findings. Also, video material from talent selection camps including not just matches, but also drills and other activities appears enriching. Third, while including two expert coaches with at least 25 years of experience in the field gave already great insights, data from more coaches would probably expand our findings' generalizability further. Thus, the aforementioned rubrics may help to maintain the high level of reliability. Here, discussions and close exchanges between practitioners, scientists and players are crucial and will be promoted to both enrich the beneficial conversations between coaches and to advance the O3T.

CONCLUSIONS AND PRACTICAL IMPLICATIONS

In conclusion, the results of the present study show that the O3T has good prospects to become a valuable tool to assess technical skills in young table tennis players a reliable and valid manner. However, it can still be improved both in terms of its design (e.g., potentially excluding quantity variables) and its integration into coach education and talent selection processes (e.g., better introduction/tutorial for assessing coaches). It should be used to observe multiple matches of one player against different opponents with two coaches to assure reliable and valid assessments. This way, the O3T can help to assess technical skills as one dimension in a multidimensional approach. Furthermore, future studies should focus on additional data acquisition in practice to improve the O3T and its representative application further. Here, the O3T should be used longitudinally to track player's development over time and to check the increasing trend for the relationship between variable 'technical quality' and table tennis performance as found in this study. Finally, for practice these results emphasize the need for coaches to observe multiple matches against multiple opponents using the O3T to get closer to an objective and realistic rating covering a player's abilities in diverse situations.

DECLARATIONS

Ethics approval and consent to participate

All procedures were in full compliance with the Declaration of Helsinki and approved by the ethical committee of the Carl von Ossietzky University Oldenburg in Germany (Reference: Drs.EK/2020/040). Informed consents were obtained prior to the present study.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

Datasets generated and/or analyzed in the context of the current study cannot be made publicly available for ethical and legal reasons; the public availability would compromise confidentiality and/or participant privacy as the data contain potentially identifying information.

COMPETING INTERESTS

Till Koopmann, Irene Faber, Dirk Büsch and Jörg Schorer declare that they have no competing interests.

FUNDING

This work was supported by the German Federal Institute for Sport Science (Bundesinstitut für Sportwissenschaft; BISp) under Grant 072004/20.

AUTHORS' CONTRIBUTIONS

All authors contributed to the study conception and design. Data collection was performed by TK and IF. Data analyses were performed by TK and IF. The first draft of the manuscript was written by TK and IF and all authors critically revised and commented on previous versions of the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We would like to acknowledge the German Table Tennis Association for their support and cooperation in data-collection, especially Dirk Wagner, Richard Prause and all coaches participating in the present study. Additional thanks go to the Netherlands Table Tennis Association, Titus Damsma, the Dutch youth table tennis players and their parents for sharing their video archive.

REFERENCES

Abernethy, B. (2008). Introduction: Developing expertise in sport-how research can inform practice. In D. Farrow, J. Baker, & C. MacMahon (Eds.), *Developing Sport Expertise: Researchers and Coaches Put Theory into Practice* (pp. 1-14). Routledge.

Ak, E., & Koçak, S. (2010). Coincidence-anticipation timing and reaction time in youth tennis and table tennis players. *Perceptual and Motor Skills, 110*(3), 879-887.

Akpinar, S., Devrilmez, E., & Kirazci, S. (2012). Coincidence-anticipation timing requirements are different in racket sports. *Perceptual and Motor Skills, 115*(2), 581-593.

Anderson, D. I., Magill, R. A., & Thouvarecq, R. (2012). Critical periods, sensitive periods, and readiness for motor skill learning In N. J. Hodges & A. M. Williams (Eds.), *Skill acquisition in sport: research, theory and practice* (pp. 237-254). Routledge.

Baker, J., Cobley, S., & Schorer, J. (2011). *Talent Identification and Development in Sport: International Perspectives.* Routledge.

Baker, J., Cobley, S., & Schorer, J. (2021). *Talent Identification and Development in Sport: International Perspectives* (2nd ed.). Routledge.

Baker, J., Cobley, S., Schorer, J., & Wattie, N. (2017). *Routledge handbook of talent identification and development in sport*. Taylor & Francis.

Baker, J., Wilson, S., Johnston, K., Dehghansai, N., Koenigsberg, A., de Vegt, S., & Wattie, N. (2020). Talent Research in Sport 1990-2018: A Scoping Review. *Frontiers in Psychology, 11*, 607710-607710. https://doi.org/10.3389/fpsyg.2020.607710

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., Bakken, S., Kaplan, C. P., Squiers, L., & Fabrizio, C. (2009). How we design feasibility studies. *American journal of preventive medicine, 36*(5), 452-457.

Cohen, J. (1988). *Statistical power analysis for behavioural sciences*. Lawrence Earlbaum Associates.

de Vet, H. C., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of clinical epidemiology, 59*(10), 1033-1039.

DeBosscher, V., Bingham, J., Shibli, S., Van Bottenburg, M., & DeKnop, P. (2008). Sports policy factors leading to international sporting success: An international comparative study. In: Meyer & Meyer.

Deutsch, K. M., & Newell, K. M. (2005). Noise, variability, and the development of children's perceptual-motor skills. *Developmental Review, 25*(2), 155-180.

Elferink-Gemser, M. T., Jordet, G., Coelho-E-Silva, M. J., & Visscher, C. (2011). The marvels of elite sports: how to get there? *British Journal of Sports Medicine,* 45(9), 683-684. https://doi.org/10.1136/bjsports-2011-090254

Elferink-Gemser, M. T., Visscher, C., Lemmink, K. A., & Mulder, T. (2007). Multidimensional performance characteristics and standard of performance in talented youth field hockey players: a longitudinal study. *Journal of Sports Sciences, 25*(4), 481-489. https://doi.org/10.1080/02640410600719945

Elling, A., Van Hilvoorde, I., & Van Den Dool, R. (2014). Creating or awakening national pride through sporting success: A longitudinal study on macro effects in the Netherlands. *International review for the sociology of sport, 49*(2), 129-151.

Faber, I., Bustin, P. M., Oosterveld, F. G., Elferink-Gemser, M. T., & Nijhuis-Van der Sanden, M. W. (2016). Assessing personal talent determinants in young racquet sport players: a systematic review. *Journal of Sports Sciences, 34*(5), 395-410. https://doi.org/10.1080/02640414.2015.1061201

Faber, I., Elferink-Gemser, M. T., Faber, N. R., Oosterveld, F. G., & Nijhuis-Van der Sanden, M. W. (2016). Can perceptuo-motor skills assessment outcomes in young table tennis players (7–11 years) predict future competition participation and performance? An observational prospective study. *PLOS one, 11*(2), e0149037.

Faber, I. R. (2016). *Diamonds in the rough. Searching for high potential in youth table tennis players* [Sl: sn]].

Faber, I. R., Elferink-Gemser, M. T., Oosterveld, F. G., Twisk, J. W., & Nijhuis-Van der Sanden, M. W. (2017). Can an early perceptuo-motor skills assessment predict future performance in youth table tennis players? An observational study (1998–2013). *Journal of Sports Sciences, 35*(6), 593-601.

Faber, I. R., Koopmann, T., Büsch, D., & Schorer, J. (2021). Developing a tool to assess technical skills in talented youth table tennis players—a multi-method approach combining professional and scientific literature and coaches' perspectives. *Sports Medicine - Open, 7*(1), 42. https://doi.org/10.1186/s40798-021-00327-5

Faber, I. R., Nijhuis-Van Der Sanden, M. W., Elferink-Gemser, M. T., & Oosterveld, F. G. (2015). The Dutch motor skills assessment as tool for talent development in table tennis: a reproducibility and validity study. *Journal of Sports Sciences, 33*(11), 1149-1158.

Friedrich, W., & Fürste, F. (2015). *Tischtennis-verstehen, lernen, spielen: ein Tischtennis Lehrbuch*. Wolfgang Friedrich.

Glazier, P. S. (2017). Could sports biomechanics provide the missing pieces to the talent identification and development puzzle? In J. Baker, S. Cobley, J. Schorer, & N. Wattie (Eds.), *The Routledge handbook of talent identification and development in sport*. Taylor & Francis.

Grix, J., & Carmichael, F. (2012). Why do governments invest in elite sport? A polemic. *International Journal of Sport Policy and Politics,* 4(1), 73-90. https://doi.org/10.1080/19406940.2011.627358

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine, 30*(1), 1-15.

Huber, D., Kohl, M., Weyers, N., & Lüsebrink, I. (2009). *Tischtennis: RTK Tischtennis (Rahmentrainingskonzeption für Kinder und Jugendliche im Leistungssport)*. Limpert.

Huijgen, B. C., Elferink-Gemser, M. T., Lemmink, K. A., & Visscher, C. (2012). Multidimensional performance characteristics in selected and deselected talented soccer players. *European Journal of Sport Science, 14*(1), 2-10. https://doi.org/10.1080/17461391.2012.725102

Johnston, K., Wattie, N., Schorer, J., & Baker, J. (2018). Talent identification in sport: A systematic review. *Sports Medicine, 48*(1), 97-109. https://doi.org/10.1007/s40279-017-0803-2

Kannekens, R., Elferink-Gemser, M. T., & Visscher, C. (2011). Positioning and deciding: Key factors for talent development in soccer. *Scandinavian Journal of Medicine & Science in Sports, 21*(6), 846-852. https://doi.org/10.1111/j.1600-0838.2010.01104.x

Katsikadelis, M., Pilianidis, T., & Mantzouranis, N. (2014). Test-retest reliability of the "table tennis specific battery test" in competitive level young players. *European Psychomotoric Journal, 6*, 3-11.

Kolman, N. S., Kramer, T., Elferink-Gemser, M. T., Huijgen, B. C., & Visscher, C. (2019). Technical and tactical skills related to performance levels in tennis: A systematic review. *Journal of Sports Sciences, 37*(1), 108-121.

Koopmann, T., Faber, I., Baker, J., & Schorer, J. (2020). Assessing Technical Skills in Talented Youth Athletes: A Systematic Review. *Sports Medicine*. https://doi.org/10.1007/s40279-020-01299-4

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology, 63*(7), 737-745.

Morrow Jr, J. R., Mood, D., Disch, J., & Kang, M. (2015). *Measurement and Evaluation in Human Performance, 5E*. Human kinetics.

Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research, 27*(5), 1147-1157.

Robertson, S., Kremer, P., Aisbett, B., Tran, J., & Cerin, E. (2017). Consensus on measurement properties and feasibility of performance tests for the exercise and sport sciences: a Delphi study. *Sports Medicine - Open, 3*(1), 1-10.

Sève, C., Saury, J., Theureau, J., & Durand, M. (2002). Activity organization and knowledge construction during competitive interaction in table tennis. *Cognitive Systems Research, 3*(3), 501-522.

Silva, A., Monteiro, D., & Sobreiro, P. (2020). Effects of sports participation and the perceived value of elite sport on subjective well-being. *Sport in Society, 23*(7), 1202-1216. https://doi.org/10.1080/17430437.2019.1613376

Table Tennis Canada. (2015). *Long Term Athlete Development. Table Tennis 4 Life*. Table Tennis Canada.

Vaeyens, R., Güllich, A., Warr, C. R., & Philippaerts, R. (2009). Talent identification and promotion programmes of Olympic athletes. *Journal of Sports Sciences, 27*(13), 1367-1380.

Van Biesen, D., Mactavish, J., Pattyn, N., & Vanlandewijck, Y. (2012). Technical proficiency among table tennis players with and without intellectual disabilities. *Human Movement Science 31*(6), 1517-1528.

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research, 19*(1), 231-240.

Int. j. racket sports sci. vol. 4(2), 2022, 32-39. eISSN: 2695-4508

DOI: 10.30827/Digibug.80313 Original Investigation

Accuracy of subjective stats of key performance indicators in tennis

Precisión de las estadísticas subjetivas de indicadores clave del rendimiento en tenis

Takashi Jindo [1] *, Daisuke Mitsuhashi [1] and Tatsumasa Kubota [2]

1 Faculty of Health and Sport Sciences, University of Tsukuba
2 School of Pharmaceutical Sciences, University of Shizuoka

Received: 29-11-21

Accepted: 02-12-22

Abstract

The compilation of stats by performance analysis is common in matches with top professional tennis players. However, outside the top level such objectively evaluated stats and feedback for players are rare. With this in mind, an original method was developed that asks players to subjectively evaluate the match stats. This study aimed to investigate the accuracy of subjective stats in tennis. The participants were 30 male collegiate athletes, including some who had participated in national-level competitions. The participants played a 6-game, 1-set practice match, and immediately after the match subjectively evaluated the stats of key performance indicators such as percentages, number of shots, and rally patterns. Objective stats were aggregated using video clips recorded by a digital camera or smartphone. Based on Bland-Altman plots show that subjectively evaluating their own performance indicators helped to confirm the objective stats. Although some variables showed fixed or proportional biases, the mean differences were not significant (percentage of first serve in: 1.733% points; double faults: 0.400 times; net plays: -0.767 times; unforced errors: -2.133 times). These findings support the implementation of a subjective evaluation of key performance indicators in tennis players who might have difficulty incorporating objective evaluations.

Keywords: *Performance analysis, profiling, feedback, tactics.*

Resumen

La recopilación de estadísticas mediante el análisis del rendimiento es común en partidos con jugadores profesionales de élite de tenis. Sin embargo, este tipo de estadísticas y retroalimentación evaluadas objetivamente son poco frecuentes en los niveles de rendimiento inferiores. Teniendo esto en cuenta, se desarrolló un método original que pide a los jugadores que evalúen subjetivamente las estadísticas de juego. El objetivo de este estudio era investigar la precisión de las estadísticas subjetivas en tenis. Los participantes fueron 30 atletas hombres universitarios; algunos de ellos habían participado en competencias nacionales. Los participantes jugaron un partido de práctica a 6 juegos y 1 set, e inmediatamente después evaluaron subjetivamente las estadísticas de indicadores clave del rendimiento tales como porcentajes, número de golpes y patrones de intercambio de golpes. Se añadieron estadísticas objetivas a través de videos grabados con una cámara digital o un teléfono inteligente. Los gráficos de Bland-Altman sugieren que evaluar subjetivamente sus indicadores de rendimiento les ayudó a confirmar las estadísticas objetivas. Aunque algunas variables mostraron sesgos fijos o proporcionales, las diferencias medias no fueron significativas (porcentaje de primeros saques: 1,733 % puntos; dobles faltas: 0,400 veces; jugadas de red: -0,767 veces; errores no forzados: -2,133 veces). Estos hallazgos apoyan la implementación de una evaluación subjetiva de los indicadores clave del rendimiento en jugadores de tenis con dificultades para aplicar evaluaciones objetivas.

Palabras clave: *Análisis del rendimiento, perfilación, retroalimentación, tácticas.*

Corresponding author: Takashi Jindo, Ph.D., jindo.takashi.ge@u.tsukuba.ac.jp

Cite this article as:
Jindo, T., Mitsuhashi, D., & Kubota, T. (2022). Accuracy of subjective stats of key performance indicators in tennis. *International Journal of Racket Sports Science,* 4(2), 32-39.

INTRODUCTION

Various sports have adopted performance profiling to analyze and improve athletic performance (Butterworth et al., 2013). A traditional method introduced by Butler and Hardy (1992) set key performance indicators (KPIs) between a coach and player prior to the match, and then rated the quality of performance on a zero to ten scale. A previous study reported that the majority of coaches praised the usefulness of the performance profiles as part of the wider coaching process (Butterworth et al., 2012). In addition, other studies have reported that athletes believed producing performance profiles in a group setting to be generally very useful (Weston et al., 2011a), thus performance profiling might enhance intrinsic motivation (Weston et al., 2011b).

Meanwhile, even for critical events in a game, coaches' observational accuracy is not perfect (Franks & Miller, 1986; Laird & Waters, 2008), leading to objective performance analysis and performance profiling using video clips or specific software being adopted recently (Butterworth et al., 2013; O'Donoghue, 2005). In these methods, the number of executions, successes, or errors of each performance indicator is objectively aggregated (O'Donoghue, 2005). Other than for world top level athletes, however, in most cases neither coaches nor tournament hosts conduct objective performance analysis. One study reported that the proportion of coaches who engage in performance analysis using video footage decreases corresponding to the lower athletic level of the player (Kraak et al., 2018). In tennis, even at a top level professional tournament, detailed objective stats would not necessarily be provided to the players (Kovalchik, 2021), while lower-level tournaments probably do not record the stats. Therefore, while performance analysis with objective stats confirms the accuracy, barriers for conducting it at various athletic levels may exist.

Compiling the values of performance indicators would be effective for improving performance. On the other hand, athlete-centered or continuous effort has been reported as an important aspect of performance profiling (Weston et al., 2011b). Traditional performance profiling has been used as a subjective tool whereby athletes would assess themselves against one or more KPIs (Butterworth et al., 2013). However, the accuracy of subjective statistics has not been well investigated. Mitsuhashi (2002) reported that subjective stats had underestimated own first service (17.1 %), and overestimated the total number of opponent forehand stroke winners (twice) and backhand stroke errors (4.4 times). However, because the study included only ten participants in five matches, accumulated evidence is needed to understand the accuracy of subjective stats.

A previous study reported that tactical skills were greater in higher level tennis players (Kolman et al., 2019). One likely reason for this is that higher athletic level players make decisions regarding tactics based on accurate subjective evaluation of their performance as well as their opponents' performance during a match. On the other hand, although performance profiling is usually conducted for a specific performance, analyzing that of an opponent is important for tactical planning, especially in ball games. It is anticipated that accurate subjective evaluation might be effective for improving tactics or performance through repetitive feedback and feedforward not only for one match but over a full tournament or season. Based on these assumptions, this study also focused on the difference in the accuracy of subjective stats between athletic levels and accuracy of opponent performance evaluation.

The purpose of this study was to investigate the accuracy of subjective KPI stats in tennis, which will be useful for creating a performance profile with a new method. Moreover, the accuracy of subjective stats for the opponent's performance indicators and the differences between athletic levels were also investigated.

MATERIAL AND METHODS

Participants

This was an observational study targeting singles tennis practice matches. The participants were 30 male collegiate tennis players at a university affiliated with the Kanto Inter-Collegiate Tennis Federation in Japan. They engaged in hitting practice and training for approximately five days a week, three to four hours per session. The participants were recruited using convenience sampling. Data gathering was conducted as part of extracurricular activities. Of the total participants, 19 had taken part in a national or equivalent-level junior or collegiate tournament (higher achievement group; age range: 18–21 years) while the remaining 11 participants had lower achievement (lower achievement group; age range: 18–21 years). The competitive records of the higher achievement group were as follows: two had participated in national-level collegiate tournaments, three had participated in semi-national-level collegiate tournaments, and 14 had participated in national-level junior tournaments. Objective competitive record was not available for the participants of lower achievement group; however, they had participated in the same hitting practice with the higher achievement group and training and considered to be regional competitive level.

The ethical committee of the author's affiliation approved the study protocol (approval number: TAI021-113).

DESIGN AND PROCEDURE

The practice matches were played by players who were considered to be equivalent or at a close athletic level based on each participant's previous achievement. For the participants who had equivalent record or those who did not have available record, practical experience or knowledge were adopted to decide the match combination. Accordingly, nine matches among the higher achievement group, one match between high and lower achievement groups, and five matches among lower achievement groups were played. The scoring format used in this study was a 6 game, 1-set match, with advantages and tie-break scores. Prior to the match, participants were notified that subjective and objective evaluations of their indicators and their opponents' indicators would be conducted. Training or familiarization session for subjective evaluation was not conducted.

The evaluation of subjective stats was conducted by each player immediately after the match using an original score sheet developed by the authors (Figure 1). Using the sheet, an original method was conducted that asked players to subjectively observe and record the values for the KPIs. Specifically, the players were asked players to recall and fill out the percentages of success or error of each shot or rally indicators on a score sheet immediately after a match. If this subjective evaluation has a certain accuracy, a performance profile can be created. This method might easily and continuously be applied by players who have difficulty implementing objective evaluations such as athletes at the lower levels. Objective stats were aggregated using video clips recorded using a digital camera or smartphone.

Performance indicators of subjective and objective stats

This study focused on the indicators of serve, return, and rally, which are commonly used in the performance analysis of tennis (O'Donoghue, 2005). Percentages, numbers of own and opponent shots, and rally indicators, were recorded.

For service, percentages of first service in, percentage of points won when first service in, and total number of double faults were used. For the return, the percentage of return in was used. For rally, the total number of winners, unforced errors, net plays, and percentage of points with net plays were used.

Analyses

The mean and standard deviation (SD) of each indicator were calculated. The Shapiro-Wilk test was used to check the normal distribution of the variables. We created a Bland-Altman plot (Bland & Altman, 1986), and conducted a paired t-test or Wilcoxon signed rank test between subjective and objective stats to investigate the accuracy of the former. The differences in the stats between the two methods were calculated by subtracting the objective values from the subjective values. Accordingly, more positive values indicate subjective overestimation and more negative values indicate subjectively underestimated stats. In a Bland-Altman plot, agreement and normality of distribution (i.e., whether the values lie between ±1.96SD) between the two stats was confirmed (Bland & Altman, 1986). Fixed bias was examined using a one-sample t-test that calculates the mean differences of two ways that significantly differ from zero. Proportional bias was examined using either the Pearson or Spearman correlation coefficients. Bland-Altman plots were shown only for their own performance indicators, which was the main target of this study. In addition, a two-sample t-test was performed to detect differences according to athletic achievement levels.

The analysis was conducted using R version 4.1.1 (The R Foundation). The level of statistical significance was set at $P < 0.05$.

A. Front page

Score sheet

Name

Date

Opponent

Score (win/lose)

What do you think was the reason for wining or losing?

Good points of the game

Bad points of the game

Characteristics of own and opponent performance

B. Back page

[Service]

Percentage own 1st service in	%
Percentage of opponent 1st service in	%
Percentage of point won when own 1st service in	%
Percentage of point won when opponent 1st service in	%
Total number of own double faults	times
Total number of opponent double faults	times

[Return]

Percentage of own returnin	%
Percentage of opponent returnin	%

[Points]

Total number of own winners	times
Total number of opponent winners	times
Total number of own unforced errors	times
Total number of opponent unforced errors	times

[Net play]

Total number of own net play	times
Total number of opponent net play	times
Total number of own point with net play	times
Total number of opponent point with net play	times

Figure 1. Score sheet.

RESULTS

Performance indicators of own shots and rally

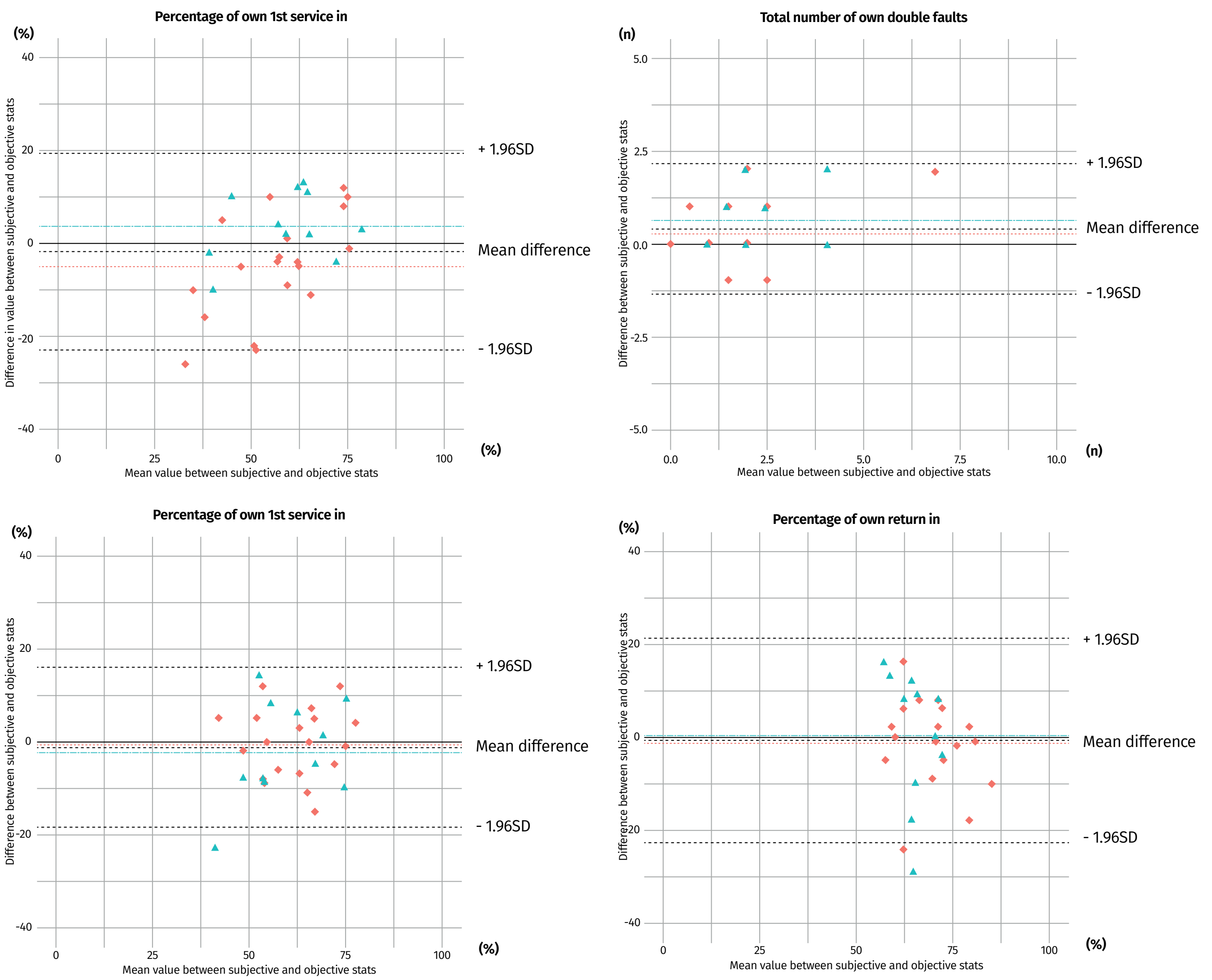
For own shot and rally indicators, significant differences between subjective and objective stats were observed in the total number of unforced errors (P = 0.011) and net plays (P = 0.008) (Table 1). Although most cases lay between ±1.96SD on the Bland-Altman plots, some variables had one or two cases less than -1.96SD (Figure 2). For fixed bias, overestimation was observed in the total number of double faults (P = 0.020), while underestimation was observed in the total number of unforced errors (P = 0.009) and net plays (P = 0.006). The other variables showed no fixed bias. For proportional bias, a tendency for overestimation in the higher percentages in first serve (r = 0.51, P = 0.004) and underestimation in fewer net plays (ρ = -0.39, P = 0.033) were observed. The other variables showed no proportional bias.

Performance indicators of opponent shots and rally

For opponent shot and rally indicators, significant differences between subjective and objective stats were observed in the percentages for first serve (P < 0.001) and points won on first serve (P = 0.024), total number of winners (P = 0.018), unforced errors (P = 0.015), and net plays (P = 0.023). For fixed bias, overestimation was observed in the percentages in first serve (P < 0.001) and points won when the first serve was in (P = 0.026), while underestimation was observed in the total number of winners (P = 0.015) and unforced errors (P = 0.011). The other variables showed no fixed bias. For proportional bias, a tendency of underestimation in the fewer number of winners (ρ = -0.55, P = 0.001) was observed. The other variables showed no proportional bias.

Comparison between athletic achievement level

A significant difference was observed in the percentages in first serve between the low and high achievement groups (P = 0.018). Specifically, the higher achievement group (-4.9±11.2 percentage points) subjectively underestimated the stats compared to the lower achievement group (3.7±7.1 percentage points). The other variables showed no significant differences between groups.



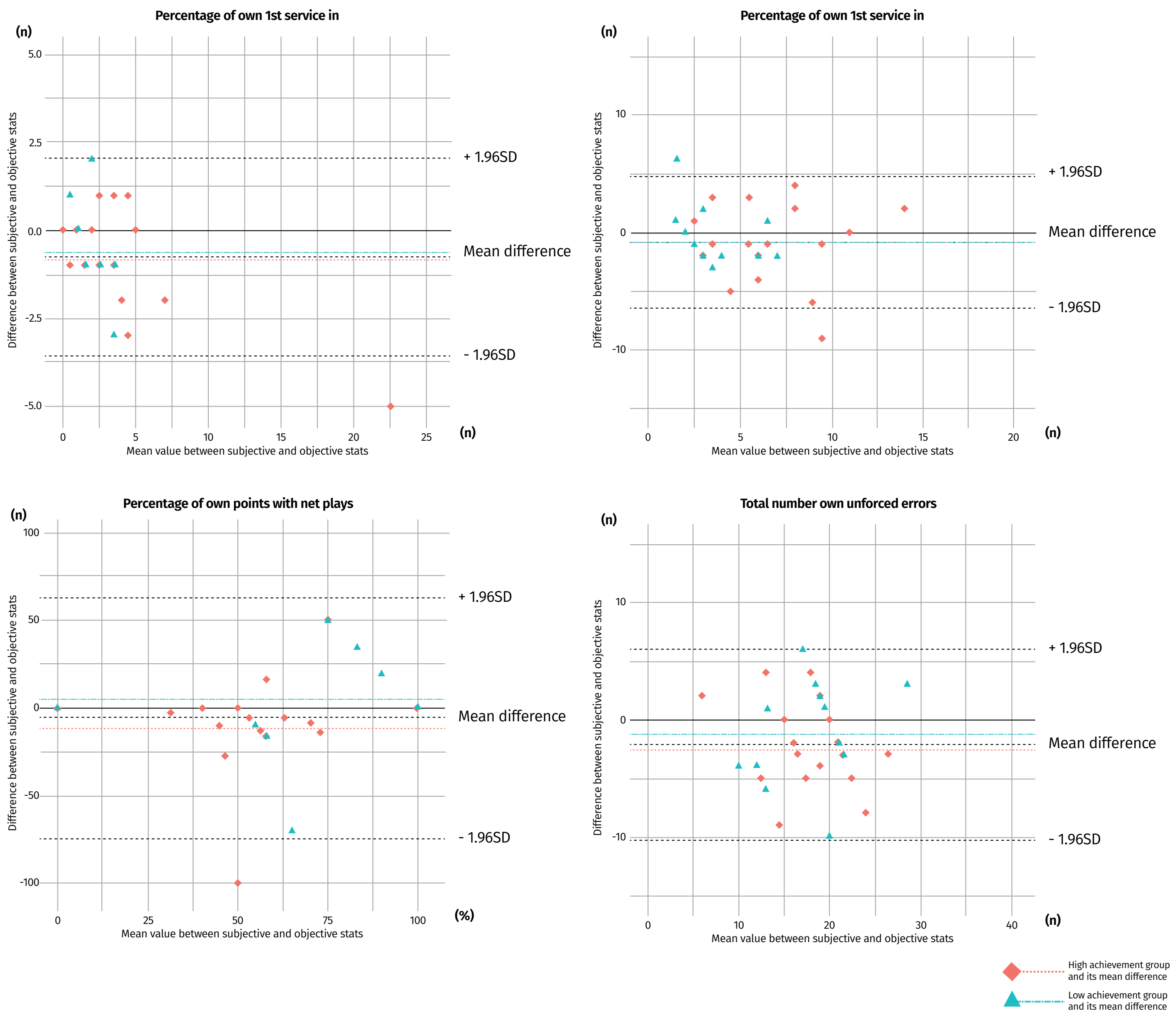


Figure 2. Bland-Altman plot of method comparison between subjective and objective stats.

Table 1. Differences in subjective and objective stats between all participants.

Item	Subjective stats			Objective stats			Mean difference between subjective and objective stats	P-value
	mean	±	SD	mean	±	SD		
Own shot and rally indicators								
Percentages of 1st service in, %	56.5	±	16.3	58.2	±	11.1	-1.733	0.386
Percentage of points won when 1st service in, %	60.7	±	11.7	61.9	±	10.2	-1.200	0.461
Total number of double faults, times †	2.1	±	1.6	1.7	±	1.3	0.400	0.022
Percentage of return in, %	67.9	±	8.4	68.5	±	10.3	-0.600	0.772
Total number of winners, times †	5.2	±	3.1	6.0	±	3.4	-0.800	0.227
Total number of unforced errors, times †	17.1	±	5.1	19.2	±	5.4	-2.133	0.011
Total number of net plays, times †	2.8	±	3.6	3.6	±	4.4	-0.767	0.008
Percentages of points with net play, % †	53.2	±	35.8	58.7	±	30.5	-5.467	0.590
Opponent shot and rally indicators								
Percentages of 1st service in, %	66.9	±	8.9	58.2	±	11.1	8.700	<0.001
Percentage of points won when 1st service in, % †	67.3	±	10.1	61.9	±	10.2	5.467	0.024
Total number of double faults, times †	1.8	±	1.7	1.7	±	1.3	0.167	0.403
Percentage of return in, % †	70.3	±	8.0	68.5	±	10.3	1.800	0.733
Total number of winners, times †	4.8	±	2.1	6.0	±	3.4	-1.133	0.018
Total number of unforced errors, times †	16.3	±	4.1	19.2	±	5.4	-2.867	0.015
Total number of net plays, times †	3.0	±	5.3	3.6	±	4.4	-0.567	0.023
Percentages of points with net play, % †	51.9	±	36.5	58.7	±	30.5	-6.800	0.592

Bold numbers indicate P < 0.05. The dagger (†) indicates that any of the analyzed variable were not normally distributed, and Wilcoxon signed rank test was used.
SD: standard deviation.

Table 2. Comparison of the subjective and objective stats difference between athletic achievement level.

Item	Low achievement group (n=11)			Low achievement group (n=19)			P-value
	mean	±	SD	mean	±	SD	
Own shot and rally indicators							
Percentages of 1st service in, %	3.7	±	7.1	-4.9	±	11.2	0.018
Percentage of points won when 1st service in, %	-2.3	±	10.6	-0.6	±	7.5	0.655
Total number of double faults, times	0.6	±	0.8	0.3	±	0.9	0.262
Percentage of return in, %	0.5	±	14.0	-1.2	±	9.2	0.735
Total number of winners, times	-0.8	±	1.6	-0.8	±	3.4	0.975
Total number of unforced errors, times	-1.3	±	4.6	-2.6	±	3.8	0.428
Total number of net plays, times	-0.6	±	1.3	-0.8	±	1.5	0.698
Percentages of points with net play, %	5.3	±	33.1	-11.7	±	34.5	0.206
Opponent shot and rally indicators							
Percentages of 1st service in, %	8.6	±	12.0	8.7	±	8.7	0.981
Percentage of points won when 1st service in, %	5.4	±	13.0	5.5	±	12.7	0.974
Total number of double faults, times	-0.1	±	0.8	0.3	±	1.1	0.265
Percentage of return in, %	4.8	±	7.2	0.1	±	11.9	0.192
Total number of winners, times	-0.5	±	1.7	-1.5	±	2.6	0.195
Total number of unforced errors, times	-0.7	±	5.2	-4.1	±	5.8	0.120
Total number of net plays, times	-0.8	±	1.4	-0.4	±	1.8	0.508
Percentages of points with net play %	-7.5	±	28.3	-6.4	±	39.8	0.936

Bold numbers indicate P < 0.05.
SD: standard deviation.

DISCUSSION

This study investigated the accuracy of subjective stats of own and opponents' KPIs in targeted tennis practice matches of collegiate male players. The results confirmed the accuracy of subjective stats for their own KPIs. On the other hand, half of the opponent's KPIs showed fixed or proportional biases between subjective and objective stats, which indicates less accuracy in opponent performance evaluation compared to their own. There was no significant difference between athletic achievement levels, except for the percentage of own first serve in, whereby we did not confirm the accuracy difference according to athletic level. These findings may be useful when conducting subjective evaluations and subsequent performance profiling.

In this study, the Bland-Altman plot showed a high degree of agreement, and there were no significant differences or fixed or proportional biases between the subjective and objective stats of most variables of own performance indicators. Even for the variables that showed fixed or proportional biases, the mean differences were not significant (percentage of first serve: -1.733 percentage points; double faults: 0.400 times; net plays: -0.767 times; unforced errors: -2.133 times), so this can be considered an acceptable level in a practical setting. A previous study that examined the accuracy of subjective stats had a small number of participants (Mitsuhashi, 2002) and there is a dearth in knowledge about the accuracy of subjective evaluation of tennis. This study suggests that performance profiling based on subjective statistics could create a profile with a certain accuracy. Such a method can be implemented with an athlete-centered and continuous style, which is expected to have consequences such as increasing the intrinsic motivation of athletes (Weston et al., 2011b). In the case of applying this method, it should be considered that, depending on the items and their levels, biases may occur.

In terms of the percentage of first serve in, where proportional bias was observed, the higher the probability, the more overestimation was observed. It would appear that when the percentage is low, players evaluate their own performance more negatively and, conversely, when the percentage is high, they evaluate it more positively. Such bias may affect the tactical and psychological aspects of a match. The number of net plays was also found to have a proportional bias, but this result may be highly influenced by one player who plays extremely frequent net plays. For a frequent net player, not only the number of net play attempts but also detailed information such as the method of approach and the characteristics of the opponent's pass might be important, and it may be necessary to consider how to utilize the subjective evaluation specifically for such a player.

On the other hand, compared to objective, the subjective stats of the opponent's performance indicators showed significant differences or biases in the percentages and points of first serve in and total numbers of winners, unforced errors, and net plays. Moreover, the tendency of overestimation and underestimation differed for each item. Although not much different from the objective stats, less accuracy was observed compared to the performance indicators. Although the practice matches in this study instructed the player to remember the stats, some information was not accurately recalled. In this context, these biases may be due to the fact that in planning tactics during a match, players must not only evaluate KPIs but also perceive a variety of other information such as the opponent's type of shot, ball speed, and course. Future studies should clarify what aspects of the opponent's performance each player focuses on during a match will lead to proposals for effective subjective evaluation after the match.

With regard to athletic achievement levels, a significant difference was found only in the own first serve probability, where the high achievement group tended to estimate their own performance lower. It should be noted that the percentage of first serve in this study was comparable to the average stats of junior and professional players (around 60%) (Kovalchik & Reid, 2017). In this study, it was hypothesized that the higher the level of athletic achievement, the more accurate the subjective evaluation, due to its importance for tactical decision-making. The hypothesis, however, was not supported, as no differences were found except for the percentage of own first serve in. In other words, the results showed that differences in tactical skills by competition level (Kolman et al., 2019) were not based on an accurate subjective evaluation of KPIs. Further investigation is needed to determine the role of accurate subjective evaluations in matching performance.

As a practical implication, it is possible to create a profile and grasp fluctuations in performance by using the score sheet to continuously record one's own performance stats. In addition, whether repeating such subjective evaluations changes the accuracy of subjective stats and whether it improves performance and affects intrinsic motivation should be examined. Moreover, we focused on the KPIs of own and opponent performance, but items should be selected according to the performance of each player. Since the KPIs might be stable in a certain value (i.e., the percentages for first serve usually fell into around 50 to 60 % (Kovalchik & Reid, 2017)), more detailed subjective evaluation such as service or return stats in deuce- and advantage-side, shots' courses or values that depend on situations (e.g., beginning or later of the match, game, etc.) should be adopted. In addition, not only shots performance but also movement performance such as movement speed or distance (Reid et al., 2016) would be a candidate indicator. Moreover, different indicators for own and opponents' performance for each player might be beneficial because limited resources can be focused on an important aspect.

This study has several limitations. First, with regard to the setting of the game, we informed the participants in advance of the matches that we would conduct a subjective evaluation, which may have impacted the accuracy of the results. In addition, because the matches were played with one set match, which is fewer than in most official matches (three sets), this may have affected the results. Future studies should target actual matches, and consider the length of the matches and differences in the match pattern. At the same time, the effect of familiarization on subjective evaluation, which was not considered in this study, should be accounted for. Second, the insufficient participants' information and the participants' selection might also be limitations. Based on international standards of competition level provided by the International Tennis Federation (ITF), the high group corresponded with intermediate to advanced, and the low group with novice to intermediate. Other characteristics, such as competitive experience (i.e., age, number of participated tournament), styles of play, or training and practice condition might affect the subjective evaluation. Future studies should investigate this information and examine whether the results of this study can be replicated at other levels and participants should be investigated. Finally, the roughness of the evaluated KPIs reveals a limitation for the application of the findings. Specifically, the accuracy of detailed indicators such as service or return stats in deuce- and advantage-sides, and shots' courses or values that depend on situations (e.g., beginning or later of the match, game, etc.) are unknown. Investigating their accuracy could demonstrate whether these detailed stats evaluations would be useful in a practical setting.

CONCLUSIONS

This study confirmed that the subjective stats of own performance indicators in tennis has a certain accuracy. This suggests that conducting performance profiling based on subjective statistics could be useful. On the other hand, the subjective stats of the opponent's performance indicators were not as accurate as that of their own performance indicators. This might be because the players have focused on not only their own KPIs but also on the other opponents' performance indicators during a match. In addition, there was no significant difference in athletic achievement levels. Based on these findings, understanding the perspectives of analysis of one's own and opponents' performance by players at various athletic levels will lead to proposals for effective ways of reflecting on matches.

ACKNOWLEDGMENTS

We would like to express our gratitude to all the participants and Mr. Daiki Hirota for their cooperation in this study.

DECLARATION OF CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

REFERENCES

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8476), 307-310.

Butterworth, A., O'Donoghue, P., & Cropley, B. (2013). Performance profiling in sports coaching: a review. *International Journal of Performance Analysis in Sport, 13*(3), 572-593. https://doi.org/10.1080/24748668.2013.11868672

Butterworth, D. A., Turner, J. D., & Johnstone, A. J. (2012). Coaches' perceptions of the potential use of performance analysis in badminton. *International Journal of Performance Analysis in Sport, 12*(2), 452-467.

Franks, I. M., & Miller, G. (1986). Eyewitness testimony in sport. *Journal of Sport Behavior, 9*(1), 38-45.

Hardy, B. R. J., Butler, R. J., & Hardy. (1992). The performance profile: theory and application. *The Sport psychologist, 6*(3), 253-264. https://doi.org/10.1123/tsp.6.3.253

International Tennis Federation. *ITN Categories*. Retrieved 27 October from https://www.tennisplayandstay.com/itn/itn-categories/itn-categories.aspx

Kolman, N. S., Kramer, T., Elferink-Gemser, M. T., Huijgen, B. C., & Visscher, C. (2019). Technical and tactical skills related to performance levels in tennis: A systematic review. *Journal of Sports Sciences, 37*(1), 108-121.

Kovalchik, S. (2021). Why Tennis Is Still Not Ready to Play Moneyball. Harvard Data Science Review.

Kovalchik, S. A., & Reid, M. (2017). Comparing matchplay characteristics and physical demands of junior and professional tennis athletes in the era of big data. *Journal of sports science & medicine, 16*(4), 489.

Kraak, W., Magwa, Z., & Terblanche, E. (2018). Analysis of South African semi-elite rugby head coaches' engagement with performance analysis. *International Journal of Performance Analysis in Sport, 18*(2), 350-366. https://doi.org/10.1080/24748668.2018.1477026

Laird, P., & Waters, L. (2008). Eyewitness recollection of sport coaches. *International Journal of Performance Analysis in Sport, 8*(1), 76-84.

Mitsuhashi, D. (2002). A difference of the subjective analysis and objective analysis in the technical factor of tennis: An examination of the usefulness of a score sheet. *Bulletin of Tokaigakuen University, 7*, 183-193.

O'Donoghue, P. (2005). Normative Profiles of Sports Performance. *International Journal of Performance Analysis in Sport, 5*(1), 104-119. https://doi.org/10.1080/24748668.2005.11868319

Reid, M., Morgan, S., & Whiteside, D. (2016). Matchplay characteristics of Grand Slam tennis: implications for training and conditioning. *Journal of Sports Sciences, 34*(19), 1791-1798. https://doi.org/10.1080/02640414.2016.1139161

The R Foundation. *The R Project for Statistical Computing*. Retrieved August 18 from http://www.r-project.org

Weston, N. J., Greenlees, I. A., & Thelwell, R. C. (2011a). Athlete perceptions of the impacts of performance profiling. *International Journal of Sport and Exercise Psychology, 9*(2), 173-188.

Weston, N. J., Greenlees, I. A., & Thelwell, R. C. (2011b). The impact of a performance profiling intervention on athletes' intrinsic motivation. *Research Quarterly for Exercise and Sport, 82*(1), 151-155.

DOI: 10.30827/Digibug.80900

Performance analysis in tennis since 2000: A systematic review focused on the methods of data collection

Análisis del rendimiento en tenis desde el año 2000: una revisión sistemática enfocada en los métodos de recolección de datos

Hiroo Takahashi [1] *, Shuhei Okamura [2] and Shunsuke Murakami [3]

1 Professor, National Institute of Fitness and Sports in Kanoya, Kagoshima, Japan.
2 Assistant Professor, Osaka University of Health and Sports Sciences, Osaka, Japan.
3 Senior Assistant Professor, National Institute of Fitness and Sports in Kanoya, Kagoshima, Japan.

Received: 01-04-22

Accepted: 03-03-23

Abstract

In tennis, performance analysis has advanced primarily as notational analysis. And analytical techniques markedly advanced, particularly in the fields of notational analysis and match analysis. In tennis, the Hawk-Eye system was introduced to tour tournaments in 2002. It has recently become used for player tracking and post-match analysis, there are a number of papers using Hawk-Eye data. Along with the development such measuring devices, technologies for analysis of a vast amount of data collected with these devices (big data) have also been developed. In particular, analysis by machine learning using AI was developed in the field of engineering, and it is also increasingly adopted in the field of sports. In the present review, we aimed to clarify the direction of research on performance analysis of tennis by organizing the trend of studies of performance analysis after 2000 with particular attention to the methods of data collection in the hope of furthering the development of this field. As a result of search of reports concerning performance analysis of tennis published after 2000 with particular interest in data collection methods, 90 papers were retrieved. The papers were classified into primary and secondary data collection, and subclassified into six categories, i.e., tracking, video recording, data mining, observations of coaches, websites, and broadcasting. This review of the papers in different categories may aid in developing future directions of research in the field of performance analysis in tennis.

Keywords: *tracking, video recording, data mining, websites, broadcasting.*

Resumen

En tenis, el análisis del desempeño ha evolucionado principalmente como análisis notacional. Y las técnicas analíticas han avanzado de manera notable, especialmente en los campos del análisis notacional y de partidos. En tenis, el sistema Hawk-Eye fue incorporado a los torneos de circuito en 2002. Recientemente se ha usado para el seguimiento de jugadores y el análisis posterior al partido, y existen diversos artículos que usan datos del Hawk-Eye. Junto con el desarrollo de dichos dispositivos de medición, también se ha desarrollado tecnología para el análisis de grandes cantidades de datos recolectados con estos dispositivos (macrodatos). En particular, se desarrolló en el campo de la ingeniería el análisis con aprendizaje automático e IA, y cada vez es más usado en el ámbito deportivo. En esta revisión, el objetivo fue clarificar la dirección de la investigación sobre el análisis del rendimiento en tenis al organizar la tendencia de los estudios de análisis del rendimiento después del año 2000 con particular atención a los métodos de recolección de datos con el fin de continuar con el desarrollo de este campo. Como resultado de la búsqueda de artículos relacionados con el análisis del rendimiento en tenis publicados después del año 2000 enfocada en métodos de recolección de datos, se encontraron 90 artículos. Los documentos se clasificaron en recopilación de datos primarios y secundarios, y se subclasificaron en seis categorías, por ejemplo, seguimiento, grabación de video, minado de datos, observaciones de entrenadores, sitios web y transmisiones. Esta revisión de artículos en diferentes categorías puede ayudar en el desarrollo de otras líneas de investigación futuras en el campo del análisis del rendimiento en tenis.

Palabras clave: *seguimiento, grabación de video, minería de datos, sitios web, transmisión.*

Corresponding author: Hiroo Takahashi, hiroo@nifs-k.ac.jp

Cite this article as:
Takahashi, H., Okamura, S., & Murakami, S. (2022). Performance analysis in tennis since 2000: A systematic review focused on the methods of data collection. *International Journal of Racket Sports Science,* 4(2), 40-55.

INTRODUCTION

Performance analysis is a new concept. Lees (2003) reviewed studies on racket sports by field and selected notational analysis as a category, but did not mention performance analysis.

The paper by Bartlett (2001) is considered to be the first on performance analysis. On the basis of differences between earlier biomechanical studies and studies of notational analysis, Bartlett (2001) defined the value of performance analysis as analysis of good and bad performances of the team and players according to performance indicators used in each genre of studies.

Thereafter, O'Donoghue (2010) defined performance analysis as investigation of sports performance using analytical methods that not only include biomechanics and notational analysis as reported by Bartlett, but also target data collected by physiological and psychological techniques.

In tennis, performance analysis has advanced primarily as notational analysis. As objectives of notational analysis, Hughes (1998) mentioned 1) tactical evaluation; 2) technical evaluation; 3) analysis of movement; 4) development of a database and modelling; and 5) educational use for both coaches and players, and reviews have since been reported according to these 5 goals. O'Donoghue (2004) also wrote reviews using the 5 goals proposed by Hughes (1998), but suggested, as prospects for the future, transformation of match analysis itself with the development of its techniques in addition to the necessity of conducting practical match analysis in the context of coaching.

The development of analytical techniques was previously brought up by Liebermann et al. (2002). Describing analytical methods for sports performance using the latest IT technology at the time, Liebermann et al. (2002) proposed that these technologies should be utilized in everyday coaching.

These reviews generally targeted papers published before 2000. Thereafter, analytical techniques markedly advanced, particularly in the fields of notational analysis and match analysis. In tennis, the Hawk-Eye system was introduced to tour tournaments in 2002 (hawkeyeinnovations.com, online). The initial objective of this system was to assist line judges, but as it has recently become used for player tracking and post-match analysis, there are a number of papers using Hawk-Eye data. In addition, instruments for tracking of the ball and players, such as Trackman (Trackman Inc.) and PlaySight (PlaySight Interactive ltd.), have been developed, and studies using such instruments are being conducted (Edelmann-Nusser et al., 2019; Murata and Takahashi, 2020; Kashiwagi et al., 2021).

Along with the development such measuring devices, technologies for analysis of a vast amount of data collected with these devices (big data) have also been developed. In particular, analysis by machine learning using AI was developed in the field of engineering, and it is also increasingly adopted in the field of sports. In studies concerning tennis, machine learning has been used by researchers including Whiteside and Reid (2017), Ganser et al. (2021), and Fernandes (2017).

In view of the changes in the methods for collection and analysis of data related to performance analysis of tennis, we considered it necessary to evaluate research themes of future performance analysis of tennis based on a review of papers published after 2000, when such changes became apparent.

In the present review, we aimed to clarify the direction of research on performance analysis of tennis by organizing the trend of studies of performance analysis after 2000 with particular attention to the methods of data collection in the hope of furthering the development of this field.

METHODS

This review was conducted according to the procedure of systematic review (Pickering and Byrne, 2014). To retrieve the relevant literature, searches were performed with "tennis", "performance", "analysis", "notation", and "match" as search words with the 'AND' condition, excluding "table" and "paddle" to restrict the search to reports concerning tennis. The databases searched were PubMed, Web of Science, and SPORT DISCUS, which encompass the literature concerning sports science. Searches were performed using the above search words in the default mode of each database, by which papers were retrieved if the search words were included in the title, abstract, or keywords. Two additional conditions, i.e., in English and published after 2000, were used for the search. The last date that we searched was April 23rd, 2021.

By the above method, 1,068 papers were retrieved. Of the retrieved papers, those that fulfilled the following conditions were included in the present review: 1) studies of performance in matches, 2) studies aiming to develop analytical methods, 3) studies analyzing quantitative data, and 4) studies published in the category of "research paper" in each journal. Papers that corresponded to the following were excluded as irrelevant to the objective of the present review: 1) studies focusing on physiological, psychological, and/or biomechanical indices alone as analytical targets, and 2) studies focusing on techniques of tennis and/or their development alone. In the first step, all the papers were screened by title, and all the authors agreed on the 130 papers that were retrieved. These papers were screened by Abstract to identify those that fulfilled the inclusion criteria, and all the authors agreed on the 90 papers that were retrieved.

While classifying the retrieved literature, attention was paid to the methods of data collection employed in each study. After overviewing the 90 retrieved papers, they were classified according to the data collection method from the following viewpoints: 1) primary data collection: match data collected using videos and tracking systems at the sites of actual matches or data collected by the researchers themselves using audio-visual media, and analytical data prepared by the researchers themselves by conducting simulations using data from such sources, and 2) secondary data collection: data collected from broad sources, such as those made public online, or that were broadcast on the television. In addition, reports classified into 1) and 2) were subclassified according to the data collection method, and the characteristics of the subclasses were evaluated. The procedures of present review were showed on Figure 1.

RESULTS

According to the data collection methods, 42 and 48 of the 90 papers were classified as using primary and secondary data collection methods, respectively.

1) Studies using primary data collection

Primary data collection methods were subclassified into automatic data collection using tracking technologies, collecting data from video images, handling data from data mining, and collecting data from the observations of coaches. Tracking technologies were classified into vision-based technologies and inertial measurement unit (IMU)-based technologies. Data from video images were classified into automatic and manual methods.

1-1) Studies using tracking technologies

There were 16 studies using vision-based tracking technologies (Table 1-1). In this category, the study using Hawk-Eye data by Loffing et al. (2010) was published earliest. Of the reports evaluated in this review, all those using tracking technologies were published after 2010. Publication of studies using Hawk-Eye data increased, particularly after about 2016 (Kolbinger and Lames, 2013; Mecheri et al., 2016; Reid et al., 2016; Wei et al., 2016; Kovalchik and Albert, 2017; Kovalchik and Reid, 2018; Cui et al., 2019; Meurs et al., 2021). A characteristic of these studies was a large data size. Among the studies using Hawk-Eye data, the study by Mecheri et al. (2016) collected data from more than 100,000 points, and the study by Kovalchik and Albert (2017) targeted more than 30,000 services. They can be regarded as big data analyses.

There were two studies using IMU-based technologies (Table 1-2). A study by Myers et al. (2019) adopted Sony Smart Sensor, and a study by Edelmann-Nusser et al. (2019) adopted BABOLAT and HEAD sensors. However, as these studies reported that the measurement accuracy of the sensors were inacceptable, the studies that used these sensors were not published.

The reports in the table are arranged in: 1) chronological order, and 2) alphabetical order using the name of the authors. Studies by the same authors and those using the same methods are unified in the same row.

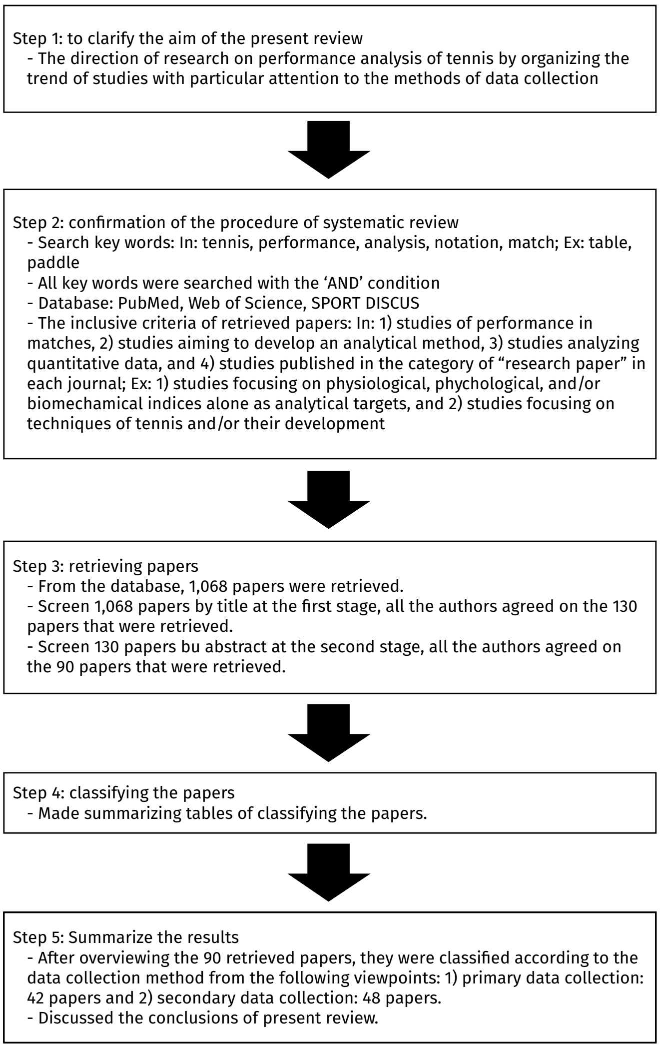


Figure 1. The procedures of present review.

1-2) Studies using video images

There were three studies that collected video images by the automatic method (Table 2-1). These studies used an independently developed system that processed video images automatically.

There were 18 studies that collected video images by the manual method (Table 2-2). These studies had been published since 2000. The methods of data collection in this category consisted of two types: observation of video images (Johnson and McHugh, 2006; Jans, 2007; Mergheş et al., 2014; Schmidhofer et al., 2014; Martin-Lorente et al., 2017), and developing independent systems (Klaassen and Magnus, 2003; Hizan et al., 2010; 2011; 2014; 2015; Klaus et al., 2017; Prieto-Lage et al., 2018). Many of studies targeted singles matches from Grand Slam tournaments.

Table 1-1. Data collection methods of vision-based tracking studies.

Authors	Year	Subject	Methods	Output data
Loffing et al.	2010	8098 rallies from 37 men's and 17 women's matches played at ATP, WTA, and Grand Slam	Hawk-Eye	% of the ball placements on opponent's backhand side
Kolbinger and Lames	2013	10418 serves of 53 right-handed male players from 56 men's singles Grand Slam matches of 2010 and 2011on hard court	Hawk-Eye	the placement of the ball of right-handed men's serves
Martínez-Gallego et al. Martínez-Gallego et al. Martinez-Gallego et al.	2013a 2013b 2019	188 games in 8 matches recorded at the ATP tournament 500 Valencia in 2011 11 professional players (age 24.8 ± 2.9) ranked between 5 and 113 on the ATP ranking	the SAGIT tracking system	(2013a) distance covered, average speed, time spent in the areas (2013b) % of unforced errors, % of winners and forced errors (2019) time, distance covered, speed, winners, errors
Stare et al.	2015	boys U14 (n=11) and girls U14 (n=10) in the national championships in Slovenia ATP tournaments (n=7)	the SAGIT tracking system	the efficiency of the first and second serves the efficiency of the forehand and backhand the efficiency of the forehand and backhand in the return of serve the efficiency of topspin forehand or backhand, the slice of forehand or backhand
Mecheri et al.	2016	professional tennis tournaments (ATP and WTA) including Grand Slam between 2003 and 2008 75587 points for the women 187009 points for the men	Hawk-Eye	the relationships between the various characteristics of the serve (speed, location, spin, etc) and winning-point probabilities
Reid et al.	2016	102 male and 95 female players during the 2012-2014 Australian Open	Hawk-Eye	Serve performance Return of serve performance Groundstroke performance Movement characteristics
Wei et al.	2016	8780 shots of the top 3 players (Djokovic, Nadal, Federer) in the 2012 men's Australian Open	Hawk-Eye	Ground stroke speed ratio Ground stroke depth ratio Ground stroke angle ratio Lateral player movement ratio
Kovalchik and Albert	2017	175 matches from 2016 Australian open 87 matches of men and 88 matches of women	Hawk-Eye	time-to-serve rally length shot importance
Pereira et al.	2017	8 professional players during 4 matches of an international tournament (Futures level) on outdoor clay court in Brazil	Automatic tracking system by Figueroa et al. (2006)	Physical performance Technical performance
Kovalchik and Reid	2018	246 matches and 270,023 shots from men and 257 matches and 178,136 shots from women in 2015-2017 Australian open	Hawk-Eye	shot types (clustered by location, shape and speed) % of point won
Pereira et al.	2018	10 of U18 players from ITF tournament 8 professional players from Futures 10 professional players from ATP250	Automatic tracking system by Figueroa et al. (2006)	Time spent of interpersonal coordination patterns during lateral displacements: Anti-phase, In-phase, Serving player phase and Returning player phase
Cui et al.	2019	1188 of men, 189 individual players, from four Grand Slam's 2015-2017	Hawk-Eye	technical-tactical and physical performance
Floyd et al.	2020	5 matches from 2015 US Open	no show (only showed as 'tennis player-tracking data')	ESV (Expected Shot Value)
Meurs et al.	2021	64 men's matches from 2017 Australian open	Hawk-Eye	PA (Positional Advantage) index by Carvalho et al. (2013)

Table 1-2. Data collection methods of IMU-based tracking studies.

Authors	Year	Subject	Methods	Output data
Myers et al.	2019	14 junior players, 12 males and 2 females	Sony Smart Tennis Sensor	Hitting volume Ball speed
Edelmann-Nusser et al.	2019	4 matches by 8 players (10-18yrs, 4 female, 4 male) 2,098 strokes	BABOLAT PURE DRIVE PLAY BABOLAT POP HEAD Tennis Sensor PlaySight	number of strokes service speed

Table 2-1. Data collection methods of video images by automatic system studies.

Authors	Year	Subject	Methods	Output data
Connaghan et al.	2013	twelve complete matches with players of various skill levels, 825 min in total	Automated tennis event indexing system	accuracy of event detection
		same as above	Match Point: visual coding system	user's evaluation
Polk et al.	2014	two-set match of the best singles players on the coaches' team	TennisVis	the scoreline by Pie Meter View point outcome by Fish Grid View match summary by Filters and Bar Charts
Lara et al.	2018	a simulated match by two players	comparison of the manual versus automatic tracking	player's positioning

1-3) Studies using data mining

There were two studies using data mining theory (Table 3). These studies aimed to predict the results of matches or simulate the progression of matches.

1-4) Studies using data from the observations of coaches

A study that aimed to clarify performance analysis in tennis using data of the observations of coaches (Torres-Luque et al., 2018) was classified into this category (Table 4).

2) Studies of secondary data collection

Secondary data collection methods were subclassified into data collection from official websites and data collection from video images published by television broadcasting and websites.

2-1) Studies using data collected from websites

There were 38 studies using information released on websites (Table 5). Such studies were more common after 2010. Most of the studies targeted men's singles matches and collected data from the official ATP website and official Grand Slam website. Some studies targeted women's, doubles, and junior matches (Brenzik, 2013; Kovalchik et al., 2017; Cui et al., 2018; Sogut, 2018; Fernandez-Garcia et al., 2021; Li et al., 2020; Grambow et al., 2021). Other studies collected data from websites that gathered match data independently (Kovalchik and Reid, 2017; Kovalchik and Ingram, 2018; Fagan et al., 2019; Ingram, 2019; Makino et al., 2020). In addition, there were some studies that had no information about the data source (Pollard et al., 2006; Newton and Aslam, 2009; Tudor et al., 2014; Gu and Saaty, 2019; Stefani, 2020). A characteristic of these studies was their large data size.

2-2) Studies using data collected from broadcasting

There were 10 studies using data collected from broadcasting (Table 6). The studies collecting data from terrestrial and satellite broadcasting were published between 2000 and 2012 (O'Donoghue, 2001; O'Donoghue and Ingram, 2001; Gillet et al., 2009; Yu et al., 2009; Nowak and Panfil, 2012), whereas recent studies collected video images from websites (Carboch et al., 2018a, 2018b, 2019, 2020; Martinez-Gallego et al., 2020). Most of the studies targeted singles matches of Grand Slam tournaments, and one study targeted doubles matches (Martinez-Gallego et al., 2020).

DISCUSSION

1) Studies using primary data collection

Methods using automated vision-based tracking techniques, mainly Hawk-Eye, will continue to be the mainstay of primary data collection. Concerning studies using Hawk-Eye data, a group participated in by Tennis Australia has recently been active in reporting Australian Open matches (Reid et al., 2016; Wei et al., 2016; Kovalchik and Albert, 2017; Kovalchik and Reid, 2018; Meurs et al., 2021). There have also been studies focusing on other tournaments (Loffing et al., 2010; Kolbinger and Lames, 2013; Mecheri et al., 2016; Cui et al., 2019) and expansion of the research field was confirmed. The Hawk-Eye system is routinely employed in major tournaments. Groups conducting these studies reached an agreement with the tournament organizers about the use of the data obtained in the tournaments by the Hawk-Eye system for research. Building such relationships between tournament organizers and Hawk-Eye providers is considered a process indispensable for the development of research in this field.

Table 2-2. Data collection methods of video images by manual system studies.

Authors	Year	Subject	Methods	Output data
Klaassen and Magnus Klaassen and Magnus Klaassen and Magnus	2001 2003 2009	(2001) 481 matches (male: 258, female: 223) at Wimbledon during 1992-1995 57,319 points in male, 28,979 points in female (2003,2009) all singles matches at Wimbledon 1992-1995	(2001,2009) no information for data collection (In each match we know the two players and the complete sequence of points.) (2003) TENNISPROB	(2001) dynamic binary panel data with random effects tests whether points in professional tennis are iid (independent and identically distributed) (2003) forecasting the probability of winning a match (2009) the efficiency of winning a point on serve
Johnson and McHugh	2006	22 players on 3 Grand Slams (8 in RG, 11 in Wimbledon, 9 in US) in 2003	observation from video recording	number of strokes stroke distribution
Jans	2007	3 final matches from 3 Grand Slams (RG, Wimbledon, US) in 2005	observation from video recording	time duration of point time interval of point total time of match time of play
Hizan et al.	2010	tennis coding system	coded the same match on two occasions separated by a 4-week period 5 raters coded 674 shots	intra-rater reliability inter-rater reliability comparison with Hawk-Eye data
Hizan et al. Hizan et al. Hizan et al.	2011 2014 2015	(2011) 28 matches (male:14, female: 14) from 2008 Australian Open, 2666 points (male: 1651, female: 1015) 28 U-16 (male: 14, female: 14) matches and 28 U-12 (male: 14, female: 4) matches from 2008 Australian Boys and Girls championships, 2359 points on U-16 (male: 1239, female: 1120) and 2267 points on U-12 (male: 1175, female: 1092) (2014) 23 matches (male:11, female: 12) from 2008 Australian Open, 1968 successful serves (male:1172, female: 796) 27 U-16 (male: 14, female: 13) matches and 21 U-12 (male: 12, female: 9) matches from 2008 Australian Boys and Girls championships, 2836 succesful serves on U-16 (male: 1439, female: 1397) and 1647 succesful serves on U-12 (male: 916, female: 731) (2015) 23 matches (male:11, female: 12) from 2008 Australian Open, 5221 serves (male: 3272, female: 1949) 27 U-16 (male: 14, female: 13) matches and 21 U-12 (male: 12, female: 9) matches from 2008 Australian Boys and Girls championships, 3391 serves on U-16 (male: 1740, female: 1651) and 1922 serves on U-12 (male: 1050, female: 872)	tennis coding system (by Hizan et al., 2010)	(2011) % 1st in, aces, DF, % 1st won, % 2nd won, % 1st return won, % 2nd return won (2014) serve-return location point winning (2015) serve location point winning
Carvalho et al. Carvalho et al.	2013 2014	(2013) 27 rallies in 3 matches from 2008 Estoril Open (ATP 250) (2014) 28 rallies in 3 matches from 2008 Estoril Open (ATP 250)	recording by DV camera and 2D-DLT	(2013) PA (Positional Advantage) index (2014) GDD (Goal-Directed Displacement) index
Mergheş et al.	2014	9 matches by 3 players (Federer, Nadal, Agassi) in 2 years	observation from video recording	% won on 1st serve % won on 2nd serve % won on return

Table 2-2. Data collection methods of video images by manual system studies (Continuation).

Schmidhofer et al.	2014	12 matches each for 3 groups (U9, U10, U12) from Develop Tournaments for Australian Tennis Association 12 matches of ATP tournaments	observation from video recording	service parameters return parameters ICT (Inter Contact Times) parameters miscellaneous parameters
Fitzpatrick et al.	2017	48 participants MTR: n=18, Age 7.4 ± 0.6, 230 points MTO: n=16, Age 8.5 ± 0.6, 253 points MTG: n=8, Age 9.9 ± 0.4, 280 points FB: n=6, Age 13.7 ± 0.5 247 points	a custom-notational analysis system	service parameters return parameters ICT (Inter Contact Times) parameters miscellaneous parameters
Klaus et al.	2017	8 U-14 national level male players in Australia QF and SF of the Victorian Junior Hardcourt Championships	A developed computerized system Kinovea (version 0.8.15)	type of stroke type of outcome court position
Martin-Lorente et al.	2017	18 matches of Grand Slam and ATP finals between 2011 and 2014 11 men players	observation from video recording	results of inside out and inside in forehand
Prieto-Lage et al.	2018	82 break point events between Nadal and Djokovic on final clay court during 2011 and 2012	observation from video recording with OBSTENNIS	the break points T-Pattern
Martínez-Gallego et al.	2021	2339 points from 19 complete doubles matches of the 2018 ATP World Tour Masters 1000 tournament played in Canada	a data collection system was designed using Microsoft Excel	time characteristics of doubles tennis time characteristics of the points by winning and losing team time characteristics of the points by the type of match

Table 3. Data collection methods of data mining studies.

Authors	Year	Subject	Methods	Output data
O'Donoghue and Simmonds	2019	Traditional tennis games Traditional tiebreaks Fast4 tennis games Tiebreaks in Fast4 tennis Tiebreak Ten	Simulation in various winning point probabilities	The probability of the player who serve first
Li et al.	2021	no information	data mining technology	serve points won and lost

Table 4. Data collection methods of studies using the observations of coaches.

Authors	Year	Subject	Methods	Output data
Torres-Luque et al.	2018	observational instrument	video observation by one observer questionnaire to 10 experts	the list of variables and categories related with the result of the match the list of variables and categories related with the development of the game

Table 5. Data collection methods of public data on websites.

Authors	Year	Subject	Source	Output data
Pollard et al.	2006	4883 matches data from 1995-2004 All Grand Slam tournaments of men's singles	no show	the probability of winning a set in a match iid (independent and identically distributed) in a set
Djurovic et al.	2009	128 matches data from 2007 and 2008 Grand Slam hard court tournaments from IBM	IBM DB2 applicaction	the latent (factor) area of a tennis match
Newton and Aslam	2009	330 players over 59 ATP tournaments in 2007 four Grand Slams four top players (Federer, Nadal, Roddick, Blake)	no show	percentage of points won on serve percentage of points won on receiving serve Monte Carlo simulations
Reid et al.	2010	2007 Matchfact information of the top 100 male professional players	ATP website	correlation coefficients between the different performance variables
O'Donoghue	2012	Study 1: 92 men's singles matches in the 2011 US Open from the official website Study 2: world top four players in Grand Slam tournament between 2008 and 2011	Study 1: 2011 US Open website Study 2: official Grand Slam website	expected and observed break points per receiving game probability of winning points during receiving game
Breznik	2013	male (N=16,732) and female (N=16,432) players between 1968 to 2011 obtained from ATP and WTA website	ATP website WTA website	number of matches won by handedness results of PageRank algorithm
Ma et al.	2013	18,288 performances between 1991 and 2008 from the website of the ATP	ATP website	predicting winner or loser by logistic regression model with three variables (match characteristics, personal characteristics, skills and performance)
O'Donoghue	2013	men's and women's matches from 2012 Grand Slam tournaments	official Grand Slam website	propotion of points won on serving probability of rare events occurred
Vaverka and Cernosek	2013	players participated in all four Grand Slam in 2008	official yearbooks and website of the ITF official Grand Slam website	correlation coefficients between body height and serve speed
Bane et al.	2014	rankings data and date of birth information from 1985 to 2010	ATP website	the age of first ATP ranking the time to reach Top 100 from first ranked the time between first entry and exit from Top 100 the time between first ATP ranked and exit from Top 100
Kovalchik	2014	498 competitors in end-of-year ATP rankings of 104 or higher between 1991 and 2012	ATP website	trends in player characteristics (30 and over, teenagers, the age of peak performance, etc) with local polynomial regression curves
Tudor et al.	2014	all the matches from main draws of Roland Garross, Wimbledon and US Open in 2010 and 2011	no show	match statistics
Filipcic et al.	2015	male players ranked in top 300 on ATP ranking in 1991, 2000 and 2010 match statistics of 1961 matches from 1991, 2363 matches from 2000, 2660 matches from 2010	ATP website	match statistics
Kim et al.	2015	2012 Australian Open SF video on the web a men's match and a women's match	video on a website (site information is no show) A coordinate system by Matlab	location of the ball bounce time series of ball anble differences
Kovalchik	2016	53,442 matches played by ATP top 100 players in 2004-2014 and 1,377 matches from 2015	ATP website	fitted model of the Pythagorean theorem

Table 5. Data collection methods of public data on websites (Continuation).

Prieto-Bermejo et al.	2016	ATP top 10 players on four Grand Slams between 1990 and 2012	ATP website	relationships between ranking position and the results on tournaments
Kovalchik and Reid	2017	match activity from 2000-2015 of junior players from ITF website professional men's and women's players from Tennis Abstract website point-by-point data for Grand Slam matches from FlashScore website more detailed point-by-point data from Hawk-Eye	ITF website Tennis Abstract website FlashScore website Hawk-Eye	relative importance of match statistics for winning
Kovalchik et al.	2017	877 player trajectories entered WTA rankings between 1989 and 2016	WTA website	the mean peak ranking in the first ranking year, the number of years during which the majority of progression occurred (the progression stage), and the rate of rankings gained during the progression stage
Cui et al.	2017	1188 players in 594 matches collected from four 2015-2017 Grand Slams men's singles	official Grand Slam website	relationships between match statistics and relative quality (RQ)
Cui et al.	2020b			difference of performance indicators between seeded players and non-seeded players
Cui et al.	2018	1369 matches in four Grand Slams women's singles	official Grand Slam website	between match variables and the relative quality (RQ) performance profiles
Kovalchik and Ingram	2018	1582 men's matches and 966 women's matches from 2010 to the present 33,788 points across 161 men’s matches and 21,450 points across 170 women’s matches at the 2015 and 2016 Australian Opens by Hawk-Eye	the Match Charting Project (www. tennisabstract.com) Hawk-Eye	point distribution by match format time distribution by match format impact of match format on match durations and upsets
Sogut	2018	male (n=60) and female (n=59) players in 2017 Wimbledon	ATP website WTA website	correlation between body height and match outcomes
Vaverka et al.	2018	men (n=72-92) and women (n=70-98) at four Grand Slams in 2008, 2012 and 2016	official Grand Slam website	differences in the serve speed of Grand Slams
Fagan et al.	2019	handedness data as well as match- play results from ATP Tennis in 2014	ATP Tennis Navigator (http://www. tennisnavigator.com/)	the advantage of left-handedness probability of match-play results
Fitzpatrick et al.	2019a	244 men's matches and 250 women's matches from 2016 and 2017 French Open	2016 and 2017 Roland Garros website	relationships between performance characteristics and PWOL (Percentage of matches in which the Winner Outscored the Loser)
Fitzpatrick et al.	2019b	244 men's matches and 250 women's matches from 2016 and 2017 French Open 241 men's matches and 249 women's matches from 2016 and 2017and Wimbledon	Roland Garros website and the Wimbledon information System by IBM	relationships between performance characteristics and PWOL (Percentage of matches in which the Winner Outscored the Loser)
Gu and Saaty	2019	82987 matches from 1990 for ATP and 35886 matches from 2003 for WTA	online sites	predicted the outcome of 2015 US OPEN
Ingram	2019	2208 matches from ATP 2014 season	MatchStat.com (scraping)	a point-based Bayesian hierarchical model for predicting the outcome of tennis matches (the probability of winning a point on serve given surface, tournament and match date)
Martin et al.	2019	50 five-set matches from 2014 Grand Slams	official Grand Slam website	effect of pacing strategies on match outcome effect of players’ ATP ranking on pacing strategies effect of Grand Slam tournament on pacing strategies
Cui et al.	2020a	146 men's matches from 2016-2017 US Open and Australian Open	official website of each tournament	set-to-set differences of match performance

Table 5. Data collection methods of public data on websites (Continuation).

Damani et al.	2020	127 men's matches from 2020 Australian Open	2020 Australian Open website	differences of match statistics among entire tournament, initial rounds (1R-4R) and intense rounds (QF, SF and F)
Fernandez-Garcia et al.	2020	546 matches by professionals and U-18 in three Grand Slams	official Grand Slam (Australian Open, Roland Garros and Wimbledon) website	differences of match statistics between professionals and U-18 players
Grambow et al.	2020	1772 men's matches from 2002-2015 Wimbledon	Wimbledon information System by IBM	serve performance comparisons by tournaments year and tournament week
Li et al.	2020	professional players of mens (n=180) and womens (n=193) within top 300 ranking between 2010 and 2018	ATP website WTA website	relationships between the age and their ranking milestones
Makino et al.	2020	4230 points on three surfaces (Hard, Clay, Grass) of four players (Federer, Nadal, Murray, Djokovic)	Match Charting Project (https://github.com/JeffSackmann/tennis_MatchChartingProject)	match winner predictions using machine learning
Stefani et al.	2020	almost 5000 men's and 5000 women's matches of four Grand Slams from 2006-2019	no show	percent of matches by the higher-seeded players
Grambow et al.	2021	1771 ladies' matches from 2002-2015 Wimbledon	Wimbledon information System by IBM	serve performance comparisons by tournaments year and tournament week

Table 6. Data collection methods of broadcasting studies.

Authors	Year	Subject	Methods	Output data
O'Donoghue	2001	men's and women's 252 matches from Grand Slam tounaments between 1997 and 1999 from terrestrial and satellite television coverage	a computerized data management system	proportion of points won when serving proportion of games won
O'Donoghue and Ingram	2001	men's and women's 175 matches from Grand Slam tounaments between 1997 and 1999 from terrestrial and satellite television coverage	a specially designed computerized notational analysis system for tennis	differencese of timing factors and strategy data among tournaments and gender
Gillet et al.	2009	116 men's matches from French Grand Slam tournament in 2005 and 2006 from terrestrial television coverage	a computerized notational system	serve characteristics and point winning serve-return characteristics and point winning
Yu et al.	2009	broadcast tennis video	a frame grouping technique	3D virtual content insertion application ball detection and tracking application
Nowak and Panfil	2012	the match by Federer and Djokovic of 2007 US Open final and 2008 Australian Open semi-final from broadcasts by Eurosport	data recorded with Microsoft Excel	relationships among type of shot, ball placement on court and fixed or dynamic elements of play
Carboch et al. Carboch et al. Carboch et al.	2018a 2018b 2019	23 women's matches from 2017 Australian Open 7 men's and 23 women's matches from 2017 Australian Open 24 men's matches from Austrarian Open, French Open and Wimbledon in 2017 from television or internet broadcast	a spreadsheet for observed variables	comparisons of point duration, number of rally shots, time between the points, rally pace and work to rest ratio
Carboch et al.	2020	23 women's matches from 2017 Australian Open and 24 men's matches from Austrarian Open, French Open and Wimbledon in 2017 from television or internet broadcast	a spreadsheet for observed variables	comparisons of match characteristics between new and used balls
Martínez-Gallego et al.	2020	34 men's doubles matches from ATP tournaments in 2018 from Tennistv.com	a registration system created with Microsoft Excel	point ending situations

As mentioned above, there was a limitation to the use of the Hawk-Eye system data; thus, video images obtained by the manual method were used. The manual method of collecting video images was a general methodology. Especially the studies that targeted junior matches, such as those without the Hawk-Eye system, adopted the manual method to collect video images (Schmidhofer et al., 2014; Fitzpatrick et al., 2017; Klaus et al., 2017). Recently, advances in image processing have made it easier to calculate parameters from images obtained with video cameras than before 2000. As the use of video images is a relatively simple method to collect data in environments where it is difficult to employ a high-tech system, such as Hawk-Eye, the use of methods currently employed in other sports events and image processing techniques used in other fields as well as developing original systems for automatic collection of parameters from video images appropriate for the objective of the study using existing techniques as references may be solutions for the establishment of a method for data collection from video images.

We confirmed that two studies used data mining theory (O'Donoghue and Simmonds, 2019; Li et al., 2021). As mentioned below, there were many studies that used published data on the Internet. The field of data mining was prospected to develop a technique for predicting or simulating the results of matches with published data on the Internet.

2) Studies of secondary data collection

Many studies of secondary data collection were carried out by collecting data from websites. On the present website of the ATP Tour (ATP TOUR.com, online), a wide variety of data, including the summary of points scored and the decisive shot at each score called MATCH BEATS, detailed results of rallies called RALLY ANALYSIS, and, on the page called the second-screen, positions where the ball was hit, positions where the ball fell, distance run, and speed of the ball hit, in addition to conventional stats, such as the first-service percentage and first-service scoring rate, are provided. Such detailed data has the same quality as the vision-based tracking data described in this review, and proceeding with exploratory research using such open data may lead to further development of research in the field of performance analysis in tennis. In particular, many studies analyzing such data from a long-time perspective have been conducted, and they are expected to provide findings that will aid in the 4) development of a database and modelling, and 5) educational use for both coaches and players among the 5 viewpoints suggested by Hughes (1998) by making studies from both cross-sectional and longitudinal viewpoints possible.

However, public data from tournaments and matches are limited, and only data of particular tournaments are available. In addition, it was only after 1991 that stats began to be provided and after 2018 that detailed stats began to be released. Therefore, caution is needed in the use of data.

Recently, data collected from broadcasting have become available on websites as streaming services. Data collected by such methods will continue to be used for research.

Most of the studies by secondary data collection targeted men's singles matches of world top-ranked players. There were few studies of female players, doubles matches, and junior players. It is necessary to perform studies to obtain data about these categories. As mentioned below, studies that targeted junior matches, such as those that were played without the Hawk-Eye system, adopted the manual method to collect video images, especially data from online streaming video for doubles matches.

CONCLUSIONS

As a result of search of reports concerning performance analysis of tennis published after 2000 with particular interest in data collection methods, 90 papers were retrieved. The data collection methods were classified into primary and secondary methods, and subclassified into 6 categories, i.e., tracking, video recording, data mining, the observations of coaches, Internet, and broadcasting. This review of the studies in different categories suggests the importance of considering vision-based tracking technologies, the increased use of manual video-recordings, the possibility of data mining, the use of official websites, and performing studies focusing on female players, doubles teams, and junior players.

REFERENCES

ATP TOUR.com (online) Nitto ATP Finals | Results | ATP Tour | Tennis. https://www.atptour.com/en/scores/current/nitto-atp-finals/605/results?, (accessed 2021-11-17).

Bane, M., Reid, M., Morgan, S. (2014). Has player development in men's tennis really changed? An historical rankings perspective. *Journal of Sports Sciences, 32*(15), 1477-1484.

Bartlett, R. (2001). Performance analysis: can bringing together biomechanics and notational analysis benefit coaches? *International Journal of Performance Analysis in Sport, 1*(1), 122-126. https://doi.org/10.1080/24748668.2001.11868254

Breznik, K. (2013). On the gender effects of handedness in professional tennis. *Journal of Sports Science and Medicine, 12*(2), 346-353.

Carboch, J., Blau, M., Sklenarik, M., Siman, J., Placha, K. (2020). Ball change in tennis: How does it affect match characteristics and rally pace in Grand Slam tournaments? *Journal of Human Sport and Exercise, 15*(1), 153-162. https://doi.org/10.14198/jhse.2020.151.14

Carboch, J., Placha, K. (2018a). Development of rally pace and other match characteristics in women's matches in the Australian open 2017. *Journal of Physical Education and Sport, 18*(2), 1079-1083. https://doi.org/10.7752/jpes.2018.s2161

Carboch, J., Placha, K., Sklenarik, M. (2018b). Rally pace and match characteristics of male and female tennis matches at the Australian Open 2017. *Journal of Human Sport and Exercise, 13*(4), 743-751. https://doi.org/10.14198/jhse.2018.134.03

Carboch, J., Siman, J., Sklenarik, M., Blau, M. (2019). Match characteristics and rally pace of male tennis matches in three grand slam tournaments. *Physical Activity Review, 7*, 49-56. https://doi.org/10.16926/par.2019.07.06

Carvalho, J., Araújo, D., Travassos, B., Esteves, Pedro., Pessanha, L., Pereira, F., Davids, K. (2013). Dynamics of players' relative positioning during baseline rallies in tennis. *Journal of Sports Sciences, 31*(14), 1596-1605. https://doi.org/10.1080/02640414.2013.792944

Carvalho, J., Araújo, D., Travassos, B., Fernandes, O., Pereira, F., Davids, K. (2014). Interpersonal Dynamics in Baseline Rallies in Tennis. *International Journal of Sports Science & Coaching, 9*(5), 1043-1056. https://doi.org/10.1260/1747-9541.9.5.1043

Connaghan, D., Moran, K., O'Connor, NE. (2013). An automatic visual analysis system for tennis. *Proceedings of The Institution of Mechanical Engineers Part P-Journal of Sports Engineering and Technology, 227*(4), 273-288. https://doi.org/10.1177/1754337112469330

Cui, Y., Gomez, MA., Goncalves, B., Liu, H., Sampaio, J. (2017). Effects of experience and relative quality in tennis match performance during four Grand Slams. *International Journal of Performance Analysis in Sport, 17*(5), 783-801. https://doi.org/10.1080/24748668.2017.1399325

Cui Y, Gómez MÁ, Gonçalves B, Sampaio J. (2018). Performance profiles of professional female tennis players in grand slams. *PLOS One, 13*(7), e0200591. https://doi.org/10.1371/journal.pone.0200591

Cui, Y., Gomez, MA., Goncalves, B., Sampaio, J. (2019). Clustering tennis players' anthropometric and individual features helps to reveal performance fingerprints. *European Journal of Sport Science, 19*(8), 1032-1044. https://doi.org/10.1080/17461391.2019.1577494

Cui, Y., Liu, H., Gómez, MÁ., Liu, H., Gonçalves, B. (2020a). Set-to-set Performance Variation in Tennis Grand Slams: Play with Consistency and Risks. *Journal of Human Kinetics, 73*(1), 153-163. https://doi.org/10.2478/hukin-2019-0140

Cui Y, Zhao Y, Liu H, Gómez MÁ, Wei R, Liu Y. (2020b). Effect of a seeding system on competitive performance of elite players during major tennis tournaments. *Frontiers in Psychology, 11*, 1294. https://doi.org/10.3389/fpsyg.2020.01294

Damani, C., Damani, B., Bagchi, A. (2020). Match statistics significant to win the initial and intense rounds of a tennis tournament. *TRENDS in Sport Sciences, 27*(4), 225-231. https://doi.org/10.23829/TSS.2020.27.4-6

Djurovic, N., Lozovina, V., Pavicic, L. (2009). Evaluation of tennis match data - new acquisition model. *Journal of Human Kinetics, 21*, 15-21. https://doi.org/10.2478/v10078-09-0002-9

Edelmann-Nusser, A., Raschke, A., Bentz, A., Montenbruck, S., Edelmann-Nusser, J., Lames, M. (2019). Validation of sensor-based game analysis tools in tennis. *International Journal of Computer Science in Sport, 18*(2), 49-59. https://doi.org/10.2478/ijcss-2019-0013

Fagan, F., Haugh, M., Cooper, Hal. (2019). The advantage of lefties in one-on-one sports. *Journal of Quantitative Analysis in Sports, 15*(1), 1-25. https://doi.org/10.1515/jqas-2017-0076

Fernandes, M. A. (2017). Using Soft Computing Techniques for Prediction of Winners in Tennis Matches. *Machine Learning Research, 2*(3), 86-98. https://doi.org/10.11648/j.mlr.20170203.12

Fernández-García, A. I., Giménez-Egido, J. M., and Torres-Luque, G. (2021). Differences in Grand Slam competition statistics between professional and U-18 players according to the sex. *Revista Internacional de Ciencias del Deporte, 17*(63), 25-37. https://doi.org/10.5232/ricyde2021.06303

Figueroa, P. J., Leite, N. J., & Barros, R. M. L. (2006). Tracking soccer players aiming their kinematical motion analysis. *Computer Vision and Image Understanding, 101*(2), 122–135. https://doi.org/10.1016/j.cviu.2005.07.006

Filipcic, A., Zecic, M., Reid, M., Crespo, M., Panjan, A., Nejc, S. (2015). Differences in performance indicators of elite tennis players in the period 1991-2010. *Journal of Physical Education & Sport, 15*(4), 671-677. https://doi.org/10.7752/jpes.2015.04102

Fitzpatrick, A., Davids, K., Stone, JA. (2017). Effects of Lawn Tennis Association mini tennis as task constraints on children's match-play characteristics. *Journal of Sports Sciences, 35*(22), 2204-2210. https://doi.org/10.1080/02640414.2016.1261179

Fitzpatrick, A., Stone, JA., Choppin, S., Kelley, J. (2019a). A simple new method for identifying performance characteristics associated with success in elite tennis. *International Journal of Sports Science & Coaching, 14*(1), 43-50. https://doi.org/10.1177/1747954118809089

Fitzpatrick, A., Stone, JA., Choppin, S., Kelley, J. (2019b). Important performance characteristics in elite clay and grass court tennis match-play. *International Journal of Performance Analysis in Sport, 19*(6), 942-952. https://doi.org/10.1080/24748668.2019.1685804

Floyd, CM., Hoffman, M., Fokoue, E. (2020). Shot-by-shot stochastic modeling of individual tennis points. *Journal of Quantitative Analysis in Sports, 16*(1), 57-71. https://doi.org/10.1515/jqas-2018-0036

Ganser, A., Hollaus, B. and Stabinger, S. (2021). Classification of Tennis Shots with a Neural Network Approach. *Sensors , 21*(17), 5703. https://doi.org/10.3390/s21175703

Gillet E, Leroy D, Thouvarecq R, Stein JF. (2009). A notational analysis of elite tennis serve and serve-return strategies on slow surface. *Journal of Strength Conditioning and Research, 23*(2), 532-539. https://doi.org/10.1519/JSC.0b013e31818efe29

Grambow, R., O'shannessy, C., Born, P., Meffert, D., Vogt, T. (2020). Serve efficiency development at Wimbledon between 2002 and 2015: A longitudinal approach to impact tomorrow's tennis practice. *Human Movement, 21*(1), 65-72. https://doi.org/10.5114/hm.2020.88155

Grambow, R., O'Shannessy, C., Born, P., Meffert, D., Vogt, T. (2021). Serve efficiency development indicates an extended women's tennis world class cohort: Analysing 14 years of Ladies Wimbledon Championships - implications for coaching. *Human Movement, 22*(2), 43-52. https://doi.org/10.5114/hm.2021.100011

Gu, W., Saaty, TL. (2019). Predicting the outcome of a tennis tournament: based on both data and judgments. *Journal of Systems Science and Systems Engineering, 28*(3), 317-343. https://doi.org/10.1007/s11518-018-5395-3

hawkeyeinnovations.com (online) Hawk-Eye innovations. https://resources.platform.pulselive.com/HawkEye/document/2016/07/29/ec84be34-2375-4b5a-8854-917e0e7021f0/HawkEyeinTennis2016.pdf, (accessed 2021-10-8).

Hizan, H., Whipp, P., Reid, M. (2010). Validation of Match Notation (A Coding System) in Tennis. *Journal of Quantitative Analysis in Sports, 6*(3), 1-13, https://doi.org/10.2202/1559-0410.1223

Hizan, H., Whipp, P., Reid, M. (2011). Comparison of serve and serve return statistics of high performance male and female tennis players from different age-groups. *International Journal of Performance Analysis in Sport, 11*(2), 365-375. https://doi.org/10.1080/24748668.2011.11868556

Hizan, H., Whipp, P., Reid, M. (2015). Gender differences in the spatial distributions of the tennis serve. *International Journal of Sports Science & Coaching, 10*(1), 87-96. https://doi.org/10.1260/1747-9541.10.1.87

Hizan, H., Whipp, P., Reid, M; Wheat, J. (2014). A comparative analysis of the spatial distributions of the serve return. *International Journal of Performance Analysis in Sport, 14*(3), 884-893, https://doi.org/10.1080/24748668.2014.11868765

Hughes, M. (1998). The application of notational analysis to racket sports. In Science and Racket Sports II (edited by Lees, A., Maynard, I., Hughes, M. and Reilly, T.), pp211-220. London: E & FN Spon.

Hughes, M. and Franks, I. (2004a). Notational analysis – a review of the literature. In Notational Analysis of Sport Second edition (edited by Hughes, M. and Franks, I.), pp59-106. London: Routledge.

Hughes, M. and Franks, I. (2004b). From analysis to coaching. In Notational Analysis of Sport Second edition (edited by Hughes, M. and Franks, I.), pp257-271. London: Routledge.

Ingram, M. (2019). A point-based Bayesian hierarchical model to predict the outcome of tennis matches. *Journal of Quantitative Analysis in Sports, 15*(4), 313-325. https://doi.org/10.1515/jqas-2018-0008

Jans, W. (2007). The Time Structure of Female Tennis Players During the Grand Slam Tournaments. *Research Yearbook, 13*(2), 230-234.

Johnson, CD., McHugh, MP. (2006). Performance demands of professional male tennis players. *British Journal of Sports Medicine, 40*(8), 696-699. https://doi.org/10.1136/bjsm.2005.021253

Kashiwagi, R., Okamura, S., Iwanaga, S., Murakami, S., Numata, K. and Takahashi, H. (2021). The differences in the ball speed and the spin rate depending on the results of a tennis serve. *Malaysian Journal of Movement, Health & Exercises, 10*(1), 48-50. https://doi.org/10.4103/2231-9409.328217

Kim, H., Cai, F., Ryu, J., Haddad, JM., Zelaznik, HN. (2015). Tennis match time Series do not exhibit long term correlations. *International Journal of Sport Psychology, 46*(6), 542-554. https://doi.org/10.7352/IJSP.2015.46.542

Klaassen, FJGM., Magnus, JR. (2001). Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model. *Journal of the American Statistical Association, 96*(454), 500-509. https://doi.org/10.1198/016214501753168217

Klaassen, FJGM., Magnus, JR. (2003). Forecasting the winner of a tennis match. *European Journal of Operational Research, 148*(2), 257-267. https://doi.org/10.1016/S0377-2217(02)00682-3

Klaassen, FJGM., Magnus, JR. (2009). The efficiency of top agents: An analysis through service strategy in tennis. *Journal of Econometrics, 148*(1), 72-85. https://doi.org/10.1016/j.jeconom.2008.09.036

Klaus, A., Bradshaw, R., Young, W., O'Brien, B., Zois, J. (2017). Success in national level junior tennis: Tactical perspectives. *International Journal of Sports Science & Coaching, 12*(5), 618-622. https://doi.org/10.1177/1747954117727792

Kolbinger, O., Lames, M. (2013). Ball trajectories in tennis - Lateral and vertical placement of right handed men's singles serves. *International Journal of Performance Analysis in Sport, 13*(3), 750-758. https://doi.org/10.1080/24748668.2013.11868686

Kovalchik, SA. (2014). The older they rise the younger they fall: age and performance trends in men's professional tennis from 1991 to 2012. *Journal of Quantitative Analysis in Sports, 10*(2), 99-107. https://doi.org/10.1515/jqas-2013-0091

Kovalchik, SA. (2016). Is there a Pythagorean theorem for winning in tennis? *Journal of Quantitative Analysis in Sports, 12*(1), 43-49. https://doi.org/10.1515/jqas-2015-0057

Kovalchik, SA., Albert, J. (2017). A multilevel Bayesian approach for modeling the time-to-serve in professional tennis. *Journal of Quantitative Analysis in Sports, 13*(2), 49-62. https://doi.org/10.1515/jqas-2016-0091

Kovalchik, SA., Reid, M. (2017). Comparing matchplay characteristics and physical demands of junior and professional tennis athletes in the era of big data. *Journal of Sports Science and Medicine, 16*(4), 489-497.

Kovalchik, S., Bane, M., Reid, M. (2017). Getting to the top: an analysis of 25 years of career rankings trajectories for professional women's tennis. *Journal of Sports Sciences, 35*(19), 1-7. https://doi.org/10.1080/02640414.2016.1241419

Kovalchik, SA., Ingram, M. (2018). Estimating the duration of professional tennis matches for varying formats. *Journal of Quantitative Analysis in Sports, 14*(1), 13-23. https://doi.org/10.1515/jqas-2017-0077

Kovalchik, S., Reid, M. (2018). A shot taxonomy in the era of tracking data in professional tennis. *Journal of Sports Sciences, 36*(18), 2096-2104. https://doi.org/10.1080/02640414.2018.1438094

Lara, JPR., Vieira, CLR., Misuta, MS., Moura, FA., de Barros, RML. (2018). Validation of a video-based system for automatic tracking of tennis players. *International Journal of Performance Analysis in Sport, 18*(1), 137-150. https://doi.org/10.1080/24748668.2018.1456886

Lees, A. (2003). Science and the major racket sports: a review. *Journal of Sports Sciences, 21*, 707-732.

Li, M., Li, Q., Li, Y., Cui, Y., Zhao, X., Guo, L. (2021). Analysis of characteristics of tennis singles matches based on 5G and data mining technology. *Security and Communication Networks*, 5549309. https://doi.org/10.1155/2021/5549309

Li, P., Weissensteiner, JR., Pion, J., Bosscher, VD. (2020). Predicting elite success: Evidence comparing the career pathways of top 10 to 300 professional tennis players. *International Journal of Sports Science & Coaching, 15*, 793-802. https://doi.org/10.1177/1747954120935828

Liebermann, D. G., Katz, L., Hughes, M. D., Bartlett, R. M., McClements, J. and Franks, I. M. (2002). Advances in the application of information technology to sports performance. *Journal of Sports Sciences, 20*(10), 755-769. https://doi.org/10.1080/026404102320675611

Loffing, F., Hagemann, N., Strauss, B. (2010). Automated processes in tennis: do left-handed players benefit from the tactical preferences of their opponents? *Journal of Sports Sciences, 28*(4), 435-443. https://doi.org/10.1080/02640410903536459

Ma, SM., Liu, CC., Tan, Y., Ma, SC. (2013). Winning matches in Grand Slam men's singles: An analysis of player performance-related variables from 1991 to 2008. *Journal of Sports Sciences, 31*(11), 1147-1155. https://doi.org/10.1080/02640414.2013.775472

Makino, M., Odaka, T., Kuroiwa, J., Suwa, I., Shirai, H. (2020). Feature selection to win the point of ATP tennis players using rally information. *International Journal of Computer Science in Sport, 19*(1), 37-50. https://doi.org/10.2478/ijcss-2020-0003

Martin-Lorente, E., Campos, J., Crespo, M. (2017). The inside out forehand as a tactical pattern in men's professional tennis. *International Journal of Performance Analysis in Sport, 17*(4), 429-441. https://doi.org/10.1080/24748668.2017.1349528

Martin, C., Bideau, B., Touzard, P., Kulpa, R. (2019). Identification of serve pacing strategies during five-set tennis matches. *International Journal of Sports Science & Coaching, 14*(1), 32-42. https://doi.org/10.1080/24748668.2017.1349528

Martínez-Gallego, R., Crespo, M., Ramón-Llin, J., Micó, S., Guzmán, JF. (2020). Men's doubles professional tennis on hard courts: Game structure and point ending characteristics. *Journal of Human Sport & Exercise, 15*(3), 633-642. https://doi.org/10.14198/jhse.2020.153.13

Martinez-Gallego, R., Guzman, JF., Crespo, M., Ramon-Llin, J., Vuckovic, G. (2019). Technical, tactical and movement analysis of men's professional tennis on hard courts. *Journal of Sports Medicine and Physical Fitness, 59*(1), 50-56. https://doi.org/10.23736/S0022-4707.17.07916-6

Martinez-Gallego, R., Guzman, JF., James, N., Pers, J., Ramon-Llin, J., Vuckovic, G. (2013a). Movement characteristics of elite tennis players on hard courts with respect to the direction of ground strokes. *Journal of Sports Science and Medicine, 12*(2), 275-281.

Martínez-Gallego, R., Guzmán, JF., James, N., Ramón-Llin, J., Crespo, M., Vuckovic, G. (2013b). The relationship between the incidence of winners/errors and the time spent in different areas of the court in elite tennis. *Journal of Human Sport & Exercise, 8*(3), S601-607, https://doi.org/10.4100/jhse.2013.8.Proc3.05

Martínez-Gallego, R., Vives, F., Guzmán, JF., Ramón-Llin, J., Crespo, M. (2021). Time structure in men's professional doubles tennis: does team experience allow finishing the points faster? *International Journal of Performance Analysis in Sport, 21*(2), 215-225. https://doi.org/10.1080/24748668.2021.1872218

Mecheri, S., Rioult, F., Mantel, B., Kauffmann, F., Benguigui, N. (2016). The serve impact in tennis: First large-scale study of big Hawk-Eye data. *Statistical Analysis and Data Mining: The ASA Data Science Journal, 9*, 310-325. https://doi.org/10.1002/sam.11316

Mergheş, PE., Simion, B., Nagel, A. (2014). Comparative analysis of return of serve as counter-attack in modern tennis. *Timisoara Physical Education & Rehabilitation Journal, 6*(12), 18-22. https://doi.org/10.2478/tperj-2014-0023

Meurs, EV., Buszard, T., Kovalchik, S., Farrow, D., Reid, M. (2021). Interpersonal coordination in tennis: assessing the positional advantage index with Australian Open Hawkeye data. *International Journal of Performance Analysis in Sport, 21*(1), 22-32. https://doi.org/10.1080/24748668.2020.1843213

Murata, M. and Takahashi, H. (2020). Verification of the accuracy and reliability of the TrackMan tennis radar. Proceedings of the Institution of Mechanical Engineers, Part P: *Journal of Sports Engineering and Technology. 235*(2), 154-160. https://doi.org/10.1177/1754337120953005

Myers, NL., Kibler, WB., Axtell, A.H., Uhl, TL. (2019). The Sony Smart Tennis Sensor accurately measures external workload in junior tennis players. *International Journal of Sports Science & Coaching, 14*(1), 24-31. https://doi.org/10.1177/1747954118805278

Newton, P., Aslam, K. (2009). Monte Carlo Tennis: A Stochastic Markov Chain Model. *Journal of Quantitative Analysis in Sports, 5*(3), 7. https://doi.org/10.2202/1559-0410.1169

Nowak, M., Panfil, R. (2012). Scoring abilities in the game of tennis - a pragmatic study of unique cases. *Human Movement, 13*(4), 313-322. https://doi.org/10.2478/v10038-012-0036-z

O'Donoghue, PG. (2001). The most important points in Grand Slam singles tennis. *Research Quarterly for Exercise and Sport, 72*(2), 125-131. https://doi.org/10.1080/02701367.2001.10608942

O'Donoghue, P. (2004). Match analysis in racket sports. Science and Racket Sports III (edited by Lees, A., Kahn, J. -F. and Maynard, I. W.), pp155-162. London: Routledge.

O'Donoghue, P. (2010). Research methods for sports performance analysis. London: Routledge, pp1-28.

O'Donoghue, P. (2012). Break points in Grand Slam men's singles tennis. *International Journal of Performance Analysis in Sport, 12*(1), 156-165. https://doi.org/10.1080/24748668.2012.11868591

O'Donoghue, P. (2013). Rare events in tennis. *International Journal of Performance Analysis in Sport, 13*(2), 535-552. https://doi.org/10.1080/24748668.2013.11868668

O'Donoghue, P., Girard, O. and Reid, M. (2013). Racket Sports. Routledge Handbook of Sports Performance Analysis (edited by McGarry, T., O'Donoghue, P. and Sampaio, J.), pp404-414. London: Routledge.

O'Donoghue, P., Ingram, B. (2001). A notational analysis of elite tennis strategy. *Journal of Sports Sciences, 19*(2), 107-115. https://doi.org/10.1080/026404101300036299

O'Donoghue, P., Simmonds, Emma. (2019). Probability of winning and match length in Tiebreak Ten tennis. *International Journal of Performance Analysis in Sport, 19*(3), 402-416. https://doi.org/10.1080/24748668.2019.1615296

Pereira, TJC., Nakamura, FY., de Jesus, MT., Vieira, CLR., Misuta, MS., de Barros, RML., Moura, FA. (2017). Analysis of the distances covered and technical actions performed by professional tennis players during official matches. *Journal of Sports Sciences, 35*(4), 361-368. https://doi.org/10.1080/02640414.2016.1165858

Pereira, TJC., van Emmerik, REA., Misuta, MS., Barros, RML, Moura, FA. (2018). Interpersonal coordination analysis of tennis players from different levels during official matches. *Journal of Biomechanics, 67*, 106-113. https://doi.org/10.1016/j.jbiomech.2017.11.036

Pickering, C. and Byrne, J. (2014). The benefits of publishing systematic quantitative literature reviews for PhD candidate and other early-career researchers. Higher Education Res Dev., 33(3): 534-548.

Polk, T., Yang, J., Hu, YQ., Zhao, Y. (2014). TenniVis: Visualization for tennis match analysis. *IEEE Transactions on Visualization and Computer Graphics, 20*(12), 2339-2348. https://doi.org/10.1109/TVCG.2014.2346445

Pollard, G., Cross, R., Meyer, D. (2006). An analysis of ten years of the four grand slam men's singles data for lack of independence of set outcomes. *Journal of Sports Science and Medicine,* 5(4), 561-566.

Prieto-Bermejo, J., Gómez-Ruano, MÁ. (2016). Entering tennis men's Grand Slams within the top-10 and its relationship with the fact of winning the tournament. *Revista Internacional de Ciencias del Deporte, 12*(46), 411-422. https://dx.doi.org/10.5232/ricyde2016.04605

Prieto-Lage, I., Prieto, MA., Curran, TP., Gutierrez-Santiago, A. (2018). An accurate and rapid system to identify play patterns in tennis using video recording material: Break point situations as a case study. *Journal of Human Kinetics, 62*(1), 199-212. https://doi.org/10.1515/hukin-2017-0170

Reid, M., McMurtrie, D., Crespo, M. (2010). The relationship between match statistics and top 100 ranking in professional men's tennis. *International Journal of Performance Analysis in Sport, 10*(2), 131-138. https://doi.org/10.1080/24748668.2010.11868509

Reid, M., Morgan, S., Whiteside, D. (2016). Matchplay characteristics of Grand Slam tennis: implications for training and conditioning. *Journal of Sports Sciences, 34*(19), 1791-1798. https://doi.org/10.1080/02640414.2016.1139161

Schmidhofer, S., Leser, R., Ebert, M. (2014). A comparison between the structure in elite tennis and kids tennis on scaled courts (Tennis 10s). *International Journal of Performance Analysis in Sport, 14*(3), 829-840. https://doi.org/10.1080/24748668.2014.11868761

Sogut, M. (2018). Stature: Does it really make a difference in match-play outcomes among professional tennis players? *International Journal of Performance Analysis in Sport,18*(2), 255-261. https://doi.org/10.1080/24748668.2018.1466259

Stare, M., Žibrat, U., Filipčič, A. (2015). Stroke effectiveness in professional and junior tennis. *Kinesiologia Slovenica, 21*(2), 39-50.

Stefani, R. (2020). A longitudinal analysis of the differential performances of seeded male and female Grand Slam tennis players. *The Sport Journal, 21*. https://thesportjournal.org/article/a-longitudinal-analysis-of-the-differential-performances-of-seeded-male-and-female-grand-slam-tennis-players/

Torres-Luque, G., Fernández-García, ÁI., Cabello-Manrique, D., Giménez-Egido, JM., Ortega-Toro, E. (2018). Design and validation of an observational instrument for the technical-tactical actions in singles tennis. *Frontiers in Psychology, 9*: 2418. https://doi.org/10.3389/fpsyg.2018.02418

Tudor, PB., Zecic, M., Matkovic, B. (2014). Differences between 2010 and 2011 performance indicators of tennis play at the grand slam tournaments. *Kinesiology, 46*, 101-106.

Vaverka, F., Cernosek, M. (2013). Association between body height and serve speed in elite tennis players. *Sports Biomechanics, 12*(1), 30-37. https://doi.org/10.1080/14763141.2012.670664

Vaverka, F., Nykodym, J., Hendl, J., Zhanel, J., Zahradnik, D. (2018). Association between serve speed and court surface in tennis. *International Journal of Performance Analysis in Sport, 18*(2), 262-272. https://doi.org/10.1080/24748668.2018.1467995

Wei, XY., Lucey, P., Morgan, S., Sridharan, S. (2016). Forecasting the next shot location in tennis using fine-grained spatiotemporal tracking data. *IEEE Transactions on Knowledge and Data Engineering, 28*(11), 2988-2997. https://doi.org/10.1109/TKDE.2016.2594787

Whiteside, D. and Reid, M. (2017). Spatial characteristics of professional tennis serves with implications for serving aces: A machine learning approach, *Journal of Sports Sciences, 35*:7, 648-654, https://doi.org/10.1080/02640414.2016.1183805

Yamaguchi, S., Oshimi, D. and Fukuhara, T. (2018). The impact of sport events on a host region: A literature review. Japan J. Phys. Educ. Hlth. *Sport Sci., 63*, 13-32. https://doi.org/10.5432/jjpehss.17065

Yu, XG., Jiang, NJ., Cheong, LF., Leong, HW., Yan, X. (2009). Automatic camera calibration of broadcast tennis video with applications to 3D virtual content insertion and ball detection and tracking. *Computer Vision and Image Understanding, 113*(5), 643-652. https://doi.org/10.1016/j.cviu.2008.01.006

DOI: 10.30827/Digibug.80901

Validation of wearables for technical analysis of tennis players

Validación de sensores inerciales para el análisis técnico de tenistas

Gabriel Delgado García [1], Jesús Siquier Coll [1], Santiago Castro Infantes [2], Emilio J. Ruiz Malagón [2], Berta Benito Colio [1] * and Francisco Tomás González Fernández [3]

1 Department of Physical Activity and Sports Sciences, Comillas Pontifical University. CESAG, Mallorca, Spain.
2 Faculty of Physical Activity and Sport. University of Granada, Granada, Spain.
3 Department of Physical Activity and Sports Sciences, University of Granada. Ciencias de la Educación y del Deporte, Melilla University, Spain.

Received: 06-05-22

Accepted: 17-03-23

Abstract

The aim of the study was to analyze the validity of three well-known commercial sensors (Zepp1, Zepp2 and Qlipp) by comparing the speed data they provide with a speed radar and a 3D photogrammetric system. Thirteen tennis players of different levels were part of the present study: In the first experiment, performed in the tennis field, 4 players executed a total of 100 strokes (serves and groundstrokes), in the groundstrokes using a ball throwing machine to standardize throws at a speed of 70 km/h and with the minimum spin effect allowed by the machine. The ball speed measured with the Zepp1 sensor and with the Qlipp sensor was compared with the speed recorded by a radar (Stalker Pro II, USA) and with a photogrammetric system composed by 4 USB cameras (ELP, China) recording at 100 Hz. The ball and the end of the racket frame were digitized on the video using the freeware Kinovea and their real 3D coordinates were obtained by applying the DLT algorithm, using the Kinemat tool in the mathematical analysis software GNU Octave. The velocity was calculated by deriving the 3D coordinates using a fifth degree spline. In the second experiment, performed inside the laboratory, 9 players executed 20 forehand and backhands each one (n = 360 groundstrokes). Ball speed was computed with the Zepp2 device and with an highly accurate photogrammetric device (Qualisys), considered as the reference. The data of the present work indicate that the hitting kinematics of each player and the speed of the stroke affects the accuracy of the sensor, so we consider that further studies are required to evaluate the error in players of different levels and playing styles. The Zepp1 and Zepp2 inertial sensors evaluated in this work seem adequate to measure ball speed in intra-subject studies and the Lin CCC values in the first study and the adjusted values in the second study were almost all greater than 0.75.

Keywords: *Tennis, performance, validation, racket sports, photogrammetry, Zepp, Qlipp.*

Resumen

El objetivo del estudio fue analizar la validez de tres sensores comerciales conocidos (Zepp1, Zepp2 y Qlipp) comparando los datos de velocidad que proporcionan con los de un radar de velocidad y con los de un sistema fotogramétrico 3D. Trece tenistas de diferentes niveles formaron parte del presente estudio. En el primer experimento, realizado en una pista de tenis, 4 tenistas realizaron un total de 77 golpeos (saques y golpeos de fondo), en el caso de los golpeos de fondo se usó una máquina lanza-pelotas para estandarizar los lanzamientos a una velocidad de 70 km/h y con el mínimo efecto liftado permitido por la máquina. La velocidad de la pelota medida con el sensor Zepp1 y con el sensor Qlipp se comparó con la velocidad registrada por un radar (Stalker Pro II, USA) y con un sistema fotogramétrico compuesto por 4 cámaras USB (ELP, China) grabando a 100 Hz. La pelota y el extremo de la raqueta fueron digitalizados en el vídeo utilizando el freeware de análisis de vídeo Kinovea y se obtuvieron sus coordenadas 3D reales aplicando el algoritmo DLT, usando la herramienta Kinemat en el software de análisis matemático GNU Octave. La velocidad fue calculada derivando las coordenadas 3D mediante un spline de quinto grado. En el segundo

Corresponding author: Berta Benito Colio, bbenitocolio@gmail.com

Cite this article as:
Delgado García, G., Siquier Coll, J., Castro Infantes, S., Ruiz Malagón, E. J., Benito-Colio, B., & González-Fernández, F. T. (2022). Validation of wearables for technical analysis of tennis players. *International Journal of Racket Sports Science*, 4(2), 56-61.

experimento, realizado en el laboratorio, 9 jugadores de tenis ejecutaron 20 derechas y 20 reveses cada uno (n = 360 golpeos) y la velocidad de la pelota se midió con el Zepp2 y con un sistema fotogramétrrico de alta precisión (Qualisys), considerado como la referencia. Los datos del presente trabajo indican que la cinemática de golpeo y la velocidad de golpeo de cada jugador afectan la precisión del sensor, por lo que consideramos que se requieren más estudios para evaluar el error en jugadores de diferentes niveles y estilos de juego. Los sensores Zepp1 y Zepp2 evaluados en este trabajo parecen adecuados para medir la velocidad de pelota en estudios intra-sujeto y los valores Lin CCC en el primer estudio y los valores ajustados en el segundo estudio fueron casi todos mayores de 0.75.

Palabras clave: *Tenis, rendimiento, validación, deportes de raqueta, fotogrametría, Zepp, Qlipp.*

INTRODUCTION

The use of wearable technology for technical analysis of tennis players is becoming increasingly common (Shan et al., 2015; Kos et al., 2016; Delgado et al., 2021; Ruiz-Malagón et al., 2022; Ruiz-Malagón et al., 2023). These technologies in addition to performance enhancement allow the quantification of training load, thus being able to help prevent overuse injuries such as epicondylitis (Edelmann-Nusser, 2019; Keaney & Reid, 2018). Some brands that market these sensors are Babolat, Zepp, Qlipp or Sony. These devices usually provide information of the stroke speed (either they estimate the speed of the racket or the ball), the spin of the stroke, the type of stroke and the impact point of the ball on the racket. We have found only two scientific works indexed in the Journal Citation Report, concerning the validity of the Babolat sensor and the (Edelmann-Nusser, 2019; Keaney & Reid, 2018). In the research by Keaney & Reid (2018) the sample consisted of a single athlete, so more studies validating these devices with a more heterogeneous sample are required. In other racket sports there are also similar publications and for example Jaitner and Gawin (2010) found high correlations between racket speed measured with an inertial sensor and badminton shuttlecock speed.

There are other publications showing other inertial sensors for technical analysis oriented to racket sports. Yang et al. (2017) develop a sensor (TennisMaster), and evaluate its performance by collecting the acceleration and angular velocity data of 1030 serves performed by 12 subjects of different playing levels. The results showed that the TennisMaster device achieves an accuracy in serve detection of 96% and an accuracy in splitting the phases of the stroke of 95%. Kos et al. (2016) also obtained high accuracy (above 95%) using algorithms for classification of forehand, backhand and serve strokes.

Considering that the quantification of training load is fundamental for both training improvement and musculoskeletal injury prevention the aim of the study was to study the validity of three known commercial sensors (Zepp1, Zepp2 and Qlipp) by comparing the speed data they provide with those of a speed radar and with those of a 3D photogrammetric system, including tennis players of different levels of play.

METHODS

Participants

The study sample for the first experiment consisted of 4 tennis players. Who performed a total of 100 strokes. One of the subjects was of competition level and the other three were beginners (Table 1). In the second experiment 9 players were included (5 of competitive level [one included in the sample of the first experiment]) and 4 beginners and the study complied with the guidelines established in the Declaration of Helsinki for research in humans.

Procedures

Part 1: On-track evaluation

Different types of strokes were performed (services and groundstrokes). In the case of the groundstrokes the ball was launched by a ball throwing machine (Lobster GrandSlam 4, see figure 1) at a speed of 70 km/h and with the minimum spin effect allowed by the device. Table 1 shows the strokes made by each player.

Table 1. Players included in the study and strokes made by each player.

Player number	Level	Characteristics	Analyzed strokes
1	Comp.	Male, 28 y.o.	30 forehands*
2	Beg.	Male; 48 y.o.	16 forehands
3	Beg.	Male, 28 y.o.	16 serves
4	Beg.	Female, 26 y.o.	16 forehands, 12 backhand & 10 serves.

Notes: Comp.: Competition; Beg.: Beginner.
*The competition player performed forehands varying the hitting effect (flat, slice or topspin).

The ball velocity measured with the Zepp1 (classic) sensor and with the Qlipp sensor was compared with the velocity recorded by a radar (Stalker Pro II, USA, see figure 1) and with a photogrammetric system composed of 4 USB cameras (ELP, China) recording at 100 Hz. The ball and the end of the racket were digitized using the freeware Kinovea and their real 3D coordinates were obtained by applying the DLT algorithm using the Kinemat tool (Reinschmidt &

van den Bogert, 1997) in the mathematical analysis software GNU Octave. The velocity was calculated by deriving the 3D coordinates using a fifth-degree spline (and computing the average speed of five frames just after the impact of the ball).

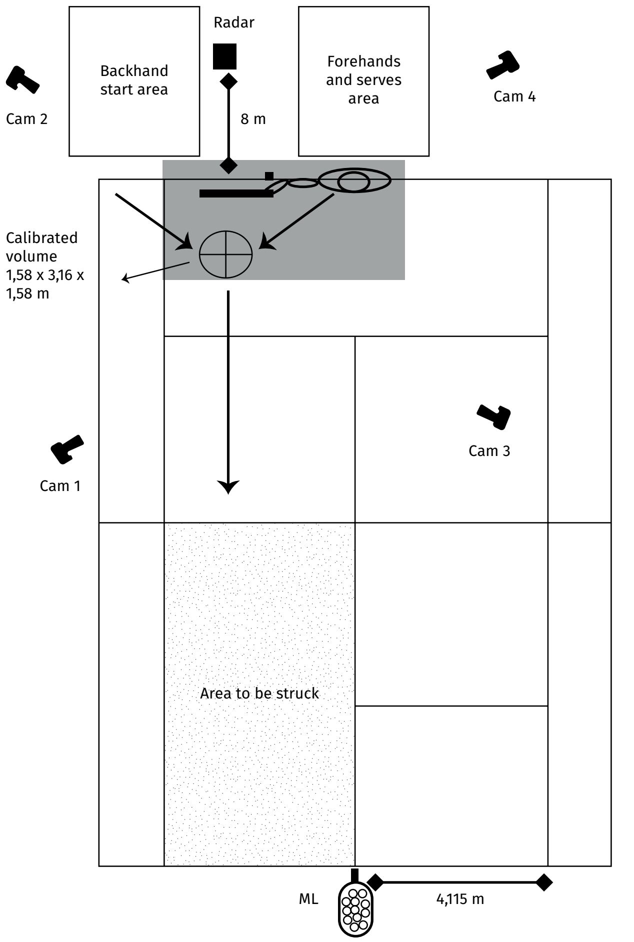


Figure 1. Scheme of the experiment carried out on track for the validation of the Zepp1 and Qlipp devices.
ML: Ball machine. Cam 1 and Cam 2 allow to analyze the serve and forehand and Cam 3 and Cam 4 the backhand.

Part 2: Laboratory evaluation

The Zepp2 (new version) device was placed on the racket grip, following manufacturer indications. The player was asked to perform 20 forehand and 20 backhand strokes against a ball attached to a flexible stick with a retroreflective marker below the ball, so a total of 360 strokes were collected (9 players x 2 types of strokes [forehand and backhands]) x 20 strokes of each type). The speed of the retroreflective marker was computed straightly after each stroke with an highly accurate photogrammetric system composed by 8 Qualisys cameras, used as the reference (Delgado-García et al., 2020).

Statistical procedures

The following statistical parameters were used to evaluate the validity of the sensor: RMSE, MAE, Pearson's r, Lin CCC and Bland-Altman (BA) plots. In order to analyze the quality of the correlations, the Evans scale (1996) was used.

In the second study the type of stroke (forehand or backhand) was considered in the statistical analysis. Both the whole sample and each groundstroke (forehand or backhand) independently were taken into account. In addition, the databases (n = 357 for the groundstrokes; n = 177 for the forehand [only three strokes were not stored] and n = 180 for the backhand) were divided in two: I) the first three databases called training databases (n = 179 for the groundstrokes; n = 89 for the forehand and n = 90 for the backhand) allowed the calculation of a ridge regression line (including the slope and the intercept at the y-axis of the line) that allowed to compute the racket speed based on the Zepp2 estimated racket speed (slope and ordinate at the origin); II) the rest of the data, called test databases were fitted based on the calculated regression equation and compared with the gold standard.

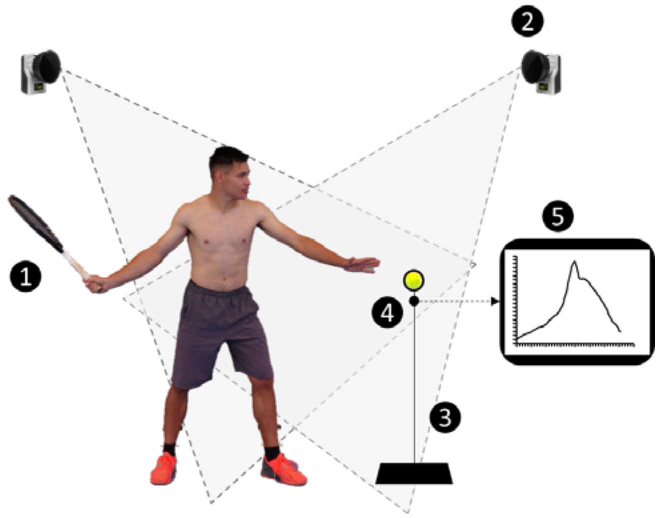


Figure 2. Set-up of the experiment number 2. The key elements are indicated with numbers: (1) tennis racket with the Zepp2 device; (2) photogrammetric system composed by 8 Qualisys cameras ; (3) Flexible stick with a tennis ball in the extreme to be hit by the player; (3) retroreflective marker for estimating ball speed with the photogrammetric system in the moment of the impact; (5) computer connected to Qualisys that allow to compute the retroreflective marker maximum speed just after the stroke.

RESULTS

Part 1: On-field evaluation

The racket velocity measured with the Zepp1 device had a high correlation score with the velocity determined with the other devices, while in the case of the Qlipp sensor the correlations were moderate (see table 2).

The values of *MAE* were (V = Velocity):

- V Radar vs. V *Zepp* = 23 km/h; V Radar vs. V *Qlipp* = 18 km/h; V Radar vs. V Ball *3D* = 5 km/h.
- V Racket *3D* vs. V *Zepp* = 7 km/h; V Racket *3D* vs. V *Qlipp* = 22 km/h.

- V Ball *3D* vs. V *Zepp* = 25 km/h; V Ball *3D* vs. V *Qlipp* = 21 km/h.

Figure 3 shows the BA plot of the racket speed measured with the Zepp1 and the racket speed measured with the 3D system. Differences in error are observed as a function of player and type of stroke (only in player 4).

Table 2. Lin CCC and Pearson's r between the speed measurements taken with different.

	VB Rad (km/h)	VR (3D) (km/h)	VB (3D) (km/h)	VR Qlipp (km/h)	VR Zepp1 (km/h)
VB Rad (km/h)	1	0.58	0.98	0.72	0.57
VR (3D) (km/h)	0.86	1	0.55	0.49	0.91
VB (3D) (km/h)	0.99	0.83	1	0.64	0.55
VR Qlipp (km/h)	0.75	0.71	0.66	1	0.57
VR Zepp1 (km/h)	0.85	0.95	0.83	0.8	1

*Above the diagonal the Lin CCC values are shown and below the diagonal the Pearson's R values are shown.
V: velocity; R: racket; B: ball.

Part 2: Laboratory evaluation

This section shows the data for the unadjusted values and the data for the adjusted values in parentheses. In the case of the groundstrokes sample (forehands and backhands) the ridge regression equation to compute the Qualisys ball speed (reference) based on the Zepp ball speed was: y = x - 6.99 (km/h) (lambda = 0.5; r = 0.76; $p < 0.001$). In the case of the forehand the ridge regression equation was y = x - 5.89 (km/h) (lambda = 10.63; r = 0.80; $p < 0.001$) and in the case of the backhand it was y = 0.859x + 5.69 (lambda = 0.89; r = 0.62; $p < 0.001$). If one doesn´t one to consider the type of stroke the adjustment proposed simply consist on substracting the value of 7 km/h to the ball speed provided by the Zepp2 device. This correction must be considered with caution as the retroreflective marker wasn´t placed exactly on the ball but a little down in the flexible stick, and considering the relation between angular and linear speed it is obvious that the speed in the extreme (ball measured with the Zepp2) will be higher, with the same angular speed. When all strokes were taken into account the Lin CCC value was 0.66 (0.75) and the MAE value was approximately 9 km/h (7 km/h). The mean error was approximately -7 km/h ± 10 km/h (0 ± 9.62 km/h), with the Zepp2 device measuring higher velocity values than Qualisys. At the intra-subject level, the highest MAE value found was 18 km/h (13 km/h) and the lowest was 4 km/h (4 km/h). When the strokes were evaluated according to the type of stroke, the following data were obtained for the forehand stroke:

- *Lin CCC* = 0.75 (0.85).
- *MAE* ~ 8 km/h (6 km/h).
- Maximum *MAE* ~ 15 km/h (10 km/h).
- Minimum *MAE* ~ 4 km/h (3 km/h).
- Mean error ~ -8 km/h ± 8 km/h (0 ± 7 km/h).

In the case of the backhand stroke the data were as follows:

- *Lin CCC* = 0.56 (0.67).
- *MAE* ~ 11 km/h (9 km/h).
- Maximum valor *MAE* ~ 20 km/h (13 km/h).
- Minimum valor *MAE* ~ 4 km/h (3 km/h).
- Mean error ~ -8 km/h ± 11 km/h (1 ± 11 km/h).

The BA plots showed heterodasticity for the groundstrokes, forehands and backhands, and the error has a positive tendency regression line while the stroke speed increases (Figures 4).

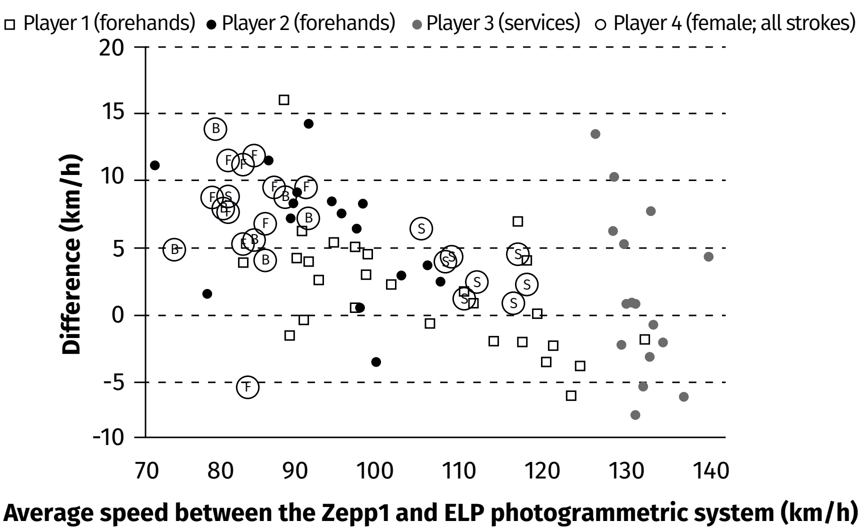


Figure 3. Bland-Altman (BA) plots of Zepp1 vs. 3D (racket) speed comparisons. * For player 4, each type of stroke is indicated by letters (F being forehand, B being backhand and S being serve).

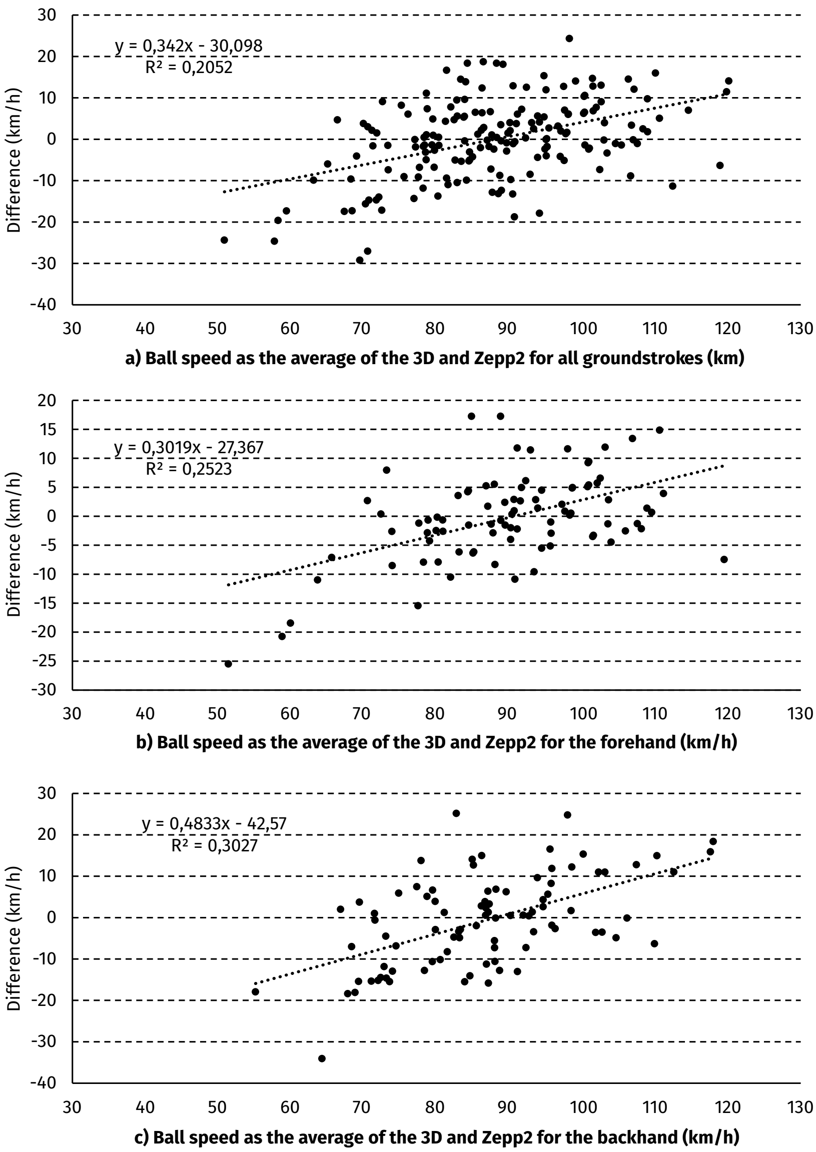


Figure 4. Bland-Altman plots that relate the average speed with the ball speed measured difference between the Zepp2 and the photogrammetric system for the total groundstrokes (a), for the forehand (b) and for the backhand (c)

DISCUSSION

The use of wearable devices for technical analysis is becoming increasingly common both in the field of training and in research. Although there are numerous companies that have developed this type of devices in tennis, the studies that analyze their validity and reliability are scarce, this experiment being one of the few in this regard. It is suggested that the error of the devices is sufficient for use in training, but not for research, where it is advised the use of photogrammetric systems.

We have only found one research paper in a journal indexed in the Journal Citation Report studying the validity of the Zepp device (Keaney & Reid, 2018). Although a high precision photogrammetric system was used as the gold standard the sample consisted of a single player and only 24 strokes were analyzed. The data of the present work indicate that the stroke kinematics of each player affects the accuracy of the sensor (for example, in Figure 3 it is observed that in the player 1 represented with white squares the magnitude of the error for the forehands is lower than that of the player 2 represented with black circles), where the error seem to be positive in almost all forehands, as well as the ball speed, as can be deduced from the Figure 4 were the speed of the strokes executed at lower speeds seem to be underestimated by the Zepp2 device while the speed of the ball of the strokes exerted at high speed seem to be overestimated (the error has a positive tendency regression line, relative to the stroke speed) so we consider that more studies are required to evaluate the error in players of different levels and styles of play. The type of stroke also seems to affect accuracy and for example in the player 4 (Figure 3) the Zepp1 overestimated the speed of the serve less than the speed of the groundstrokes. The aforementioned article indicates that the Zepp sensor and the Babolat branded smart racket, determined the volume and intensity of the strokes with good accuracy (mean error for stroke speed was 2.69 ± 5.63 km/h), but were less effective in identifying the type of stroke or the location of the impact on the racket.

Keaney & Reid (2018) point out that quantifying training using these types of sensors is critical, but that further validation studies are required. They also indicate that there is a need to improve inertial sensors for technical analysis of tennis players so that they can accurately measure impact location. This is of great interest, both for performance improvement and injury prevention, taking into account that this variable (point of impact of the ball on the racket) is related - in addition to the delivery speed of the ball after impact - with the vibrations transmitted from the racket to the arm and therefore with musculoskeletal injuries such as epicondylitis.

PRACTICAL APPLICATIONS

Despite the importance of further research, inertial sensors seem to be suitable for measuring tennis ball velocity in intrasubject studies and for trainning in the case of beginners were change in velocity after a trainning program could be sustantials.

CONCLUSIONS

The inertial sensors evaluated in this work (Zepp1, Zepp2 and Qlipp) seem adequate for measuring ball velocity in intra-subject studies (the Lin CCC values in the first study and the adjusted values in the second study were almost all greater than 0.75). Specifically, the Zepp brand sensor obtained higher values. However, the Zepp2 errors were approximately 10 km/h when evaluating the unadjusted data and approximately 7 km/h for the adjusted data (in the laboratory study). These values are quite similar to those obtained in the Keaney & Reid (2018) study. It is suggested that the measurement error of the Zepp is high in case of use with high-level players, where changes in velocity after a training program may be unnoticeable. In the case of beginner players, it could be useful since the changes after a training program will surely be more evident. It is necessary to validate the rest of the variables provided by these sensors (type of stroke, location of the impact on the racket, and stroke effect) and to include a larger number of players, taking into account that the stroke pattern could affect the sensor measurements.

REFERENCES

Delgado-García, G., Vanrenterghem, J., Mildenberger, C., Gallego, L. R., Chicano-Gutiérrez, J. M., & Soto-Hermoso, V. (2020). Accuracy and reliability of a low-cost methodology to assess 3D body posture based on commercial cameras and Excel templates. *Measurement*, 108638-108638. https://doi.org/10.1016/j.measurement.2020.108638

Delgado-García, G., Vanrenterghem, J., Ruiz-Malagón, E. J., Molina-García, P., Courel-Ibáñez, J., & Soto-Hermoso, V. M. (2021). IMU gyroscopes are a valid alternative to 3D optical motion capture system for angular kinematics analysis in tennis. Proceedings of the Institution of Mechanical Engineers, *Part P: Journal of Sports Engineering and Technology, 235*(1), 3-12. https://doi.org/10.1177/1754337120965444

Edelmann-Nusser, A., Raschke, A., Bentz, A., Montenbruck, S., Edelmann-Nusser, J., & Lames, M. (2019). Validation of sensor-based game analysis tools in tennis. *International Journal of Computer Science in Sport, 18*(2), 49-59. https://doi.org/10.2478/ijcss-2019-0013

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Thomson Brooks/Cole Publishing Co.

Jaitner, T., & Gawin, W. (2010). *A mobile measure device for the analysis of highly dynamic movement techniques, 2*(2), 3005-3010. https://doi.org/10.1016/j.proeng.2010.04.102

Keaney, E. M., & Reid, M. (2018). Quantifying hitting activity in tennis with racket sensors: new dawn or false dawn?. *Sports Biomechanics, 19*(6), 831-839. https://doi.org/10.1080/14763141.2018.1535619

Kos, M., Zenko, J., Vlaj, D., & Kramberger, I. (2016). Tennis stroke detection and classification using miniature wearable IMU device. *2016 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 1-4.

Reinschmidt, C., & Van den Bogert, T. (1997). *KineMat. A MATLAB Toolbox for Three-Dimensional Kinematic Analyses. Calgary, Canada: Human Performance Laboratory, The University of Calgary.; 1997* [cited 2020 Jul 22]. https://isbweb.org/software/movanal/kinemat/

Ruiz-Malagón, E. J, Vanrenterghem, J., Ritacco-Real, M., González-Fernández, F. T., Soto-Hermoso, V. M., & Delgado-García G. (2023) Field-based upper-body motor variability as determinant of stroke performance in the main tennis strokes. In *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*. https://doi.org/10.1177/17543371231156266

Ruiz-Malagón, E. J., Delgado-García, G., Ritacco-Real, M., & Soto-Hermoso, V. M. (2022). Kinematics differences between one-handed and two-handed tennis backhand using gyroscopes. An exploratory study. *International Journal of Racket Sports Science, 4*(1), 16-24. https://doi.org/10.30827/Digibug.76982

Shan, C. Z., Sen, S. L., Che Fai, Y., & Su Lee Ming, E. (2015). Investigation of Sensor-based Quantitative Model for Badminton Skill Analysis and Assessment. *Journal Teknologi, 72*(2). https://doi.org/10.11113/jt.v72.3891

Yang, D., Tang, J., Huang Y., Xu, C., Li, J., Hu, L., Shen, G., Liang, C-J., & Liu, H. (2017). *TennisMaster: an IMU-based online serve performance evaluation system. In Proceedings of the 8th Augmented Human International Conference (AH '17)*. Association for Computing Machinery, New York, NY, USA, (17), 1–8. https://doi.org/10.1145/3041164.3041186