Int. j. racket sports sci. vol. 4(2), 2022, 20-31. eISSN: 2695-4508
DOI: 10.30827/Digibug.--- 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], 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). The Oldenburg observation sheet for Table Tennis Technique (O3T) as a tool for talent identification and development: a reliability, validity and feasibility study. *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

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

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

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

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

Stroke

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.

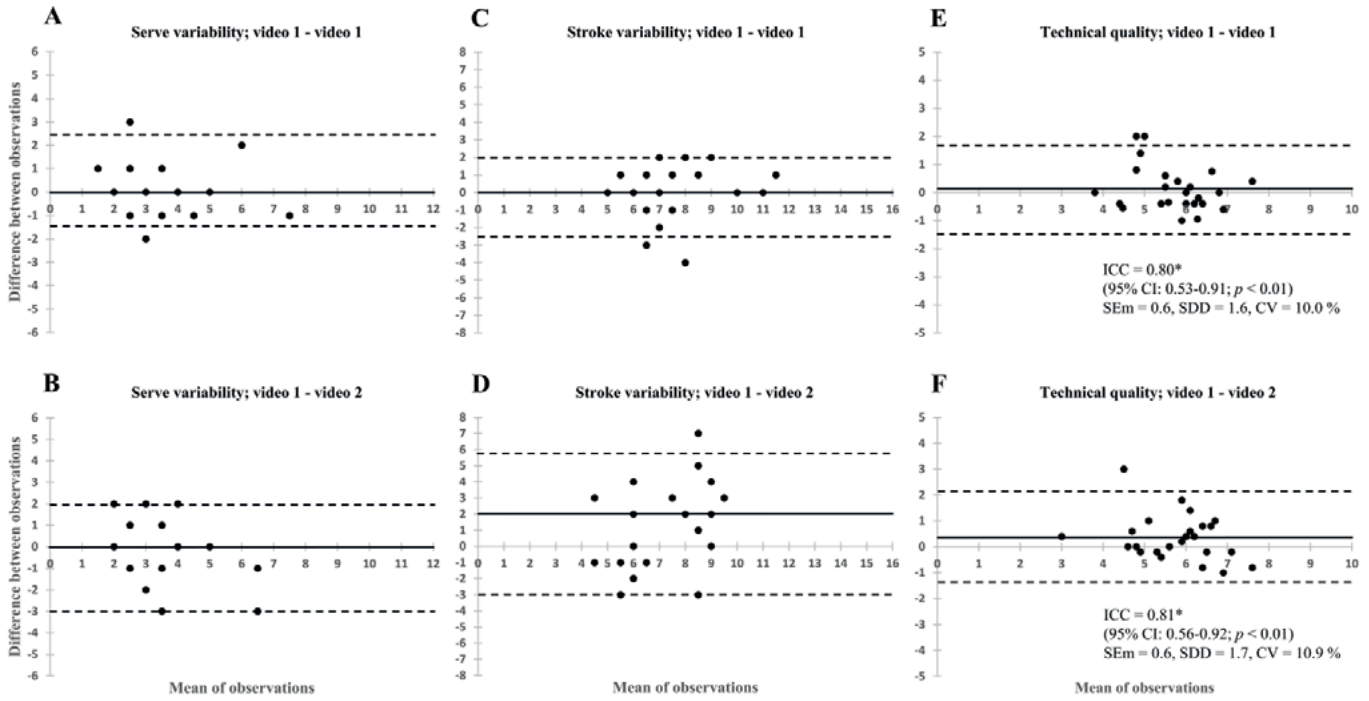

Figure 2. Bland-Altman plots for the intra-rater reliability for all three technical variables.
Legend: Solid line Median value for 'serve quantity' and 'stroke quantity', mean value for 'technical quality'. Dashed line 2.5/97.5 percentiles for 'serve quantity' and 'stroke quantity', 1.96*SD limits for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

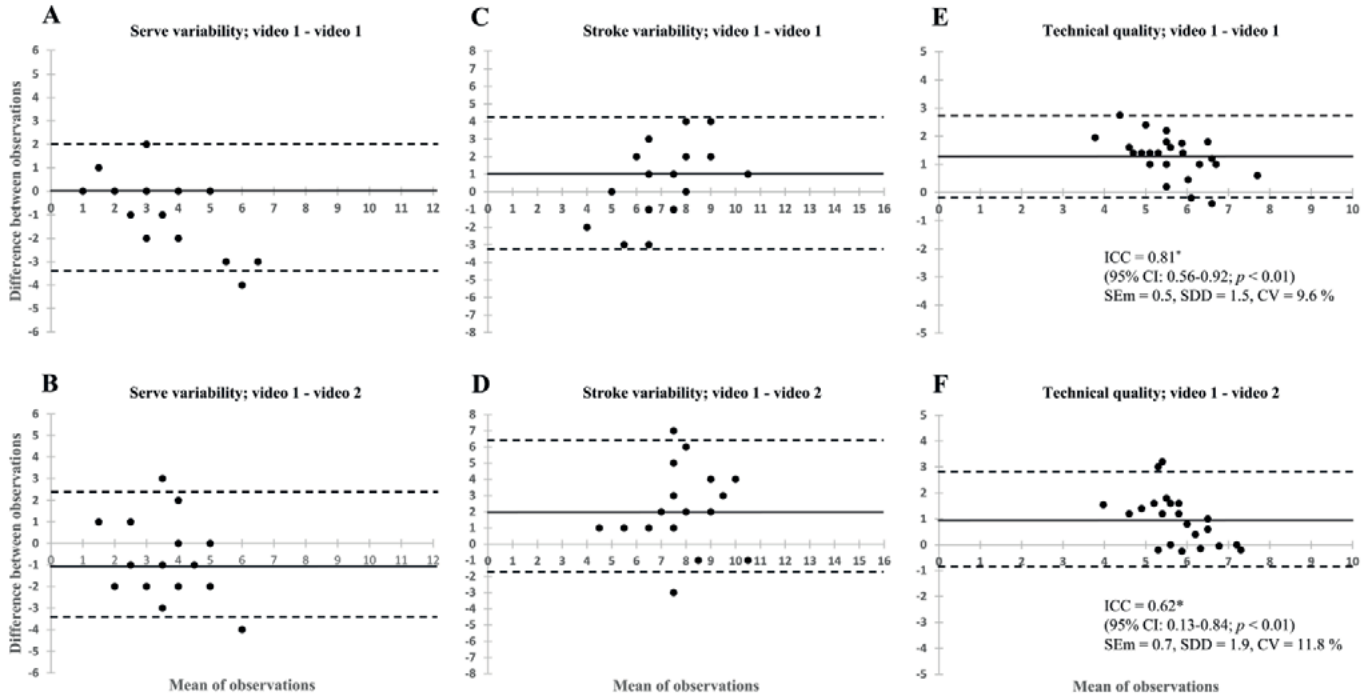

Figure 3. Bland-Altman plots for the inter-rater reliability for all three technical variables.
Legend: Solid line Median value for 'serve quantity' and 'stroke quantity', mean value for 'technical quality'. Dashed line 2.5/97.5 percentiles for 'serve quantity' and 'stroke quantity', 1.96*SD-limits for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