Reliability and characterisation of the 20-maximum trampoline jump test

Natalie Dyas[a,b,c], ORCID ID: 0000-0002-4454-2974
David Green[b], ORCID ID: 0000-0002-0749-0087
Kevin Thomas[a], ORCID ID: 0000-0003-2973-5337
Glyn Howatson[a,d]*, ORCID ID: 0000-0001-8494-2043

[a] Department of Sport, Exercise and Rehabilitation, Northumbria University, Newcastle-upon-Tyne, UK
[b] English Institute of Sport, UK
[c] British Gymnastics, Lilleshall National Sports Centre, Shropshire, UK
[d] Water Research Group, School of Environmental Sciences and Development, Northwest University, Potchefstroom, South Africa

*Corresponding author: Glyn Howatson, Department of Sport, Exercise and Rehabilitation, Northumbria University, Northumberland Building, Newcastle upon Tyne, NE1 8ST, UK. E-mail: glyn.howatson@northumbria.ac.uk

Word count: 3468

Paper submitted: June 7, 2020 paper accepted: August 13, 2020

Author Contributions
CONCEPTION: ND, DG, KT, GH
PERFORMANCE OF WORK: ND
INTERPRETATION OR ANALYSIS OF DATA: ND, DG, KT, GH
PREPARATION OF THE MANUSCRIPT: ND, GH
REVISION FOR IMPORTANT INTELLECTUAL CONTENT: ND, DG, KT, GH
SUPERVISION: GH

Abstract

BACKGROUND: The 20-maximum trampoline jump test is a commonly used performance measure in competitive trampoline gymnastics. However, its reliability and characteristics are poorly understood.

OBJECTIVE: To determine the reliability of the 20-maximum trampoline jump test and describe its characteristics.

METHODS: Thirty-two national and international level trampoline gymnasts (18 ± 5 years) performed two, 20-maximal straight jump tests, separated by 24-72 hours. Time of flight (total, jumps 1-10 and jumps 11-20), force (average and peak), horizontal displacement and contact time were measured by a competition standard system. Test-retest reliability was assessed using intraclass correlation coefficient (ICC), typical error, and coefficient of variation (CV).

RESULTS: Total time of flight significantly decreased between trial 1 (31.80 ± 1.98 s) and trial 2 (31.43 ± 1.99 s; $p < 0.05$), however test-retest reliability was excellent (ICC = 0.96, CV = 1.3%). Other time of flight and force measures significantly decreased from trial 1 to trial 2, whereas contact time increased. All secondary measures displayed very high ICC (0.95-0.99) and low CV values (0.5-1.9%), except horizontal displacement (ICC = 0.54, CV = 20.6%).

CONCLUSION: The 20-maximum test possesses excellent reliability for the assessment of trampoline performance in a wide population of national and international level gymnasts.

Keywords

Performance; elite athlete; gymnastics; time of flight; recovery

Introduction

Trampoline gymnastics has been an Olympic event since 2000. The scoring system in trampoline gymnastics is judged on ten functional acrobatic skills, assessed for difficulty, execution, time of flight (ToF), and horizontal displacement (HD). Routines take approximately 40 s to perform, encompassing both forwards and backwards somersaults and twisting skills. The gymnasts are given 60 s to start their routine in a competition before they receive a penalty deduction [1], but this routine build-up can consist of any number of in-jumps to build the necessary momentum for optimal execution of functional skills.

In a typical routine a trampoline gymnast will perform approximately 20-30 jumps, with contact times on the trampoline bed ranging from around 0.30-0.35 s. In that time the gymnast must utilise the energy stored in the bed to propel themselves into the air, with the aim of gaining maximum flight time for two important reasons: i) to achieve a higher ToF to maximise the time available to execute complex skills, and ii) to achieve higher scores for the ToF scoring component, defined as "the overall time spent in the air" [1]. Understanding how best to optimise ToF is therefore of high importance for gymnasts, coaches, and practitioners.

Recent research has highlighted the increasing importance of the ToF metric to trampoline competition outcomes, by assessing the results from the 2011, and 2013-2015 senior World and World Age Group Championship individual finals [2]. On average, competition routine ToF scores increased across gender and age groups from 2011 to 2015 (females +5.63%; males +4.93%). Senior level gymnasts exhibited the greatest ToF scores overall (females 15.980 s; males 17.706 s). The ToF score was also reported to be a greater average contributor to overall score (40 ± 3%), compared to difficulty (24 ± 3%) and execution (36 ± 4%). These data suggest that ToF has not only increased in importance, but that it is now the biggest contributor to trampoline scoring outcomes. Further evidence of this can be seen at major competition; for example, at the London 2012 Olympic Games, the athletes in 3rd and 4th position both had difficulty scores of 14.8, and similar execution scores of 24.8 and 25.0 [3]. However, the gymnast in 3rd had a greater ToF score of 16.35 s (total score = 55.96 s), compared to the gymnast in 4th who had a lower ToF score of 16.06 s (total score = 55.86 s). Therefore, without the addition of ToF the gymnast who finished in 4th would have medalled, amplifying the significance of the ToF parameter.

In elite competition, ToF, alongside horizontal displacement (HD), is measured electronically via a force plate-based system such as the Eurotramp HD ToF and Synchronicity system (HDTS). This system also provides force and contact time data for each jump, which contribute to the calculation of ToF. The 20 maximum trampoline test (20-max) test is a commonly used monitoring tool that represents the approximate construction of a trampoline routine, whereby the first 10 jumps are, in the most-part, apportioned to building height and the last 10 jumps represent the performance of the functional skills. The 20-max test omits the performance of functional skills, to allow a specific measure of the ability of the gymnast to attain high ToF independent of their acrobatic skills.

Two studies have utilised various maximal trampoline jump tests to investigate different physical parameters [3,4]. However, only Jensen et al. [4] has assessed the reliability of a 20-max jump test, reporting that the 20-maximal trampoline jump test had very high reliability, with average CV values of 0.1% between trials. In addition, nearly perfect ($r > 0.9$) correlations were reported for all test-retest comparisons in seven Danish gymnasts across three different training days with trials separated by one week [4]. Notwithstanding, there were methodological differences compared to using an FIG approved ToF measurement system, such as the Eurotramp HDTS. Jensen et al. [4] measured ToF using a video camera and included time on the bed within their total ToF measures, alongside time spent in the air. Time on the bed would not be included within the ToF metric in a competition, therefore this version of the 20-max lacks translation to competition ToF measures. Additionally, the study only demonstrated the reliability of the 20-max test with seven male gymnasts and hence lacks some applicability to wider populations.

This aims of this study were to, firstly, assess the reliability of the 20-max test for the measurement of total ToF across a cohort of national and international trampoline gymnasts using a competition standard ToF measurement system. Secondly, to assess the reliability of ToF (jumps 1-10 and jumps 11-20), force (peak and average), HD and contact time measures. Thirdly, to describe the characteristics of the outcome measures of the 20-max test, as well as data on the characteristics in an international and national level cohort.

Methods

Participants

Thirty four national and international level trampoline gymnasts, 17 females (mean ± SD; age = 18 ± 6 years; stature = 163.4 ± 6.0 cm; mass = 57.7 ± 8.0 kg) and 17 males (age = 16 ± 3 years; stature = 171.9 ± 10.2 cm; mass = 63.0 ± 13.5 kg) were recuited from the 2019 British Gymnastics National trampoline squads. The gymnasts recruited from the Great Britain senior squad (n = 6) included gymnasts over the age of 18 who had participated at the Olympic Games, World or European Championships. Gymnasts recruited from the Great Britain senior development squad (> 17 years of age) had competed internationally at age group level (n = 5). Gymnasts recruited from the Great Britain junior and England squads had all competed nationally, and some internationally at age-group level, and were between the ages of 13-19 (n = 23). Participants were familiarised with the testing procedures having regularly performed the 20-max test as part of their normal training programme. Participants were instructed to refrain from caffeine on test days and attend testing sessions in a fed and euhydrated state. Two male gymnasts withdrew from the study due to injuries unrelated to the test. The study was approved by the Northumbria University Research Ethics Committee (HLS15980), in accordance with the Declaration of Helsinki. Written consent was given by all participants, and by parents and/or guardians of the minors involved.

Materials

To assess the 20-max jump test, we used a competition standard trampoline (Eurotramp, Premium 4x4, Weilheim an der Teck, Germany), connected to an FIG approved Eurotramp HDTS system (Eurotramp, Germany). The HDTS system was connected to the base of the trampoline, and then interfaced to a laptop containing Eurotramp ToF analysis software (Qira 1.03, Eurotramp, Germany) in order to provide measures of ToF (jumps 1-10, 11-20 and total), force (average and peak), HD (distance from the centre of the bed) and contact time for each single jump and for the total. The HDTS system has been assessed for its temporal accuracy in the measurement of ToF, and has subsequently been deemed suitable for its use in elite trampoline competitions [5]. Loads on the force plates, under each of the four feet of the trampoline, produce a deformation of the sensor causing a change in light intensity, whilst optoeletronic sensors measure normal force in a one-dimensional plane [5]. The scanning is performed effectively at 2 kHz (internally 50 kHz, mean average formation over 25 values), data is provided every 0.5 ms and the sensor has a resolution of less than 0.5 N at an accuracy of 1% [6]. The HDTS system is calibrated prior to use with no weight on the trampoline bed. Following this initial calibration, the HDTS software assesses its state and adjusts the calibration if required.

All participants performed a standardised warm-up protocol, which included 10-minutes on a cycle ergometer, followed by a series of floor-based jumps. The gymnasts then performed the

20-max test, in which they started from a static position in the centre of the trampoline bed and performed 20 maximal straight jumps. Trials were separated by 24-72 hours.

Statistical analysis

A Shapiro-Wilk test determined the outcome variables (1-10 ToF, 11-20 ToF, total ToF, average force, peak force, HD and contact time) were normally distributed. Differences in outcome measures between trials of the 20-max test were determined by paired samples t-tests. Significance was set at an alpha level of 0.05. Mean differences in outcome measures, both percentage change and standardised (effect size) between test-retest trials (95% confidence intervals; CI) were calculated. The magnitude of effect size was interpreted as: trivial = < 0.20; small = 0.2-0.59; moderate = 0.60-1.19; large = 1.20-1.99; very large => 2.0 [6]. The test-retest reliability of the 20-max was assessed using a freely available spreadsheet [7]. Test-retest reliability was calculated using the intra-class correlation coefficient (ICC), typical error of measurement (TE, raw units), and TE as a log-transformed coefficient of variation (CV), with 95% CI. Calculated ICC values were classified according to the following thresholds: 0.9 nearly perfect; 0.7–0.9 very large; 0.5–0.7 large; 0.3–0.5 moderate; and 0.1–0.3 small [6]. A CV of ≤ 5% was considered as good between-session reliability for performance tests [8,9]. The ToF plateau, representing when the gymnast had achieved a consistent maximum jump height, was calculated individually for each gymnast and then averaged. The ToF plateau was determined by calculating the percentage difference between each gymnast’s five highest averaged jumps across the two trials (99%).

Results

Test-retest reliability

The mean ± SD for each outcome measure from both trials of the 20-max test, and associated reliability statistics, are shown in Table 1. No differences in repeatability scores between males and females were seen and hence data from this homogenous group are pooled. All ToF and force variables significantly decreased from trial 1 to trial 2 ($p < 0.05$), whereas contact time significantly increased from trial 1 to trial 2 ($p = 0.006$). Although the inter-subject range (percentage decrease and increase) between trials were small; total ToF = −3.98 to +1.7%; 1-10 ToF = −6.44 to +6.55%; 11-20 ToF = −5.19 to +2.35%, average force = −8.58 to +1.50%; peak force = −7.58 to +1.97%. The HD did not significantly differ between trials ($p = 0.180$). Furthermore, effect size analyses revealed trivial to small differences between test-retest trials in all outcome measures. All ToF, force and contact time measures of the 20-max test displayed high ICC values ($r = 0.95$-0.99), whereas HD displayed a moderate value ($r = 0.54$). Furthermore, all ToF, force and contact time measures had excellent test-retest reliability, ranging from 0.5%-1.9%, whereas HD displayed a CV value of 20.6% (Table 1).

Insert Table 1 here

Characteristics of the 20-max jump test

Data were averaged across the two trials. Large inter-subject variability was evident between outcome measures; total ToF (28.940-35.650 s), 1-10 ToF (12.900-16.165 s), 11-20 ToF (15.620-19.595 s), average force (4600-10158 N), peak force (5056-11136 N), HD (21.22-59.23 cm) and contact time (0.270-0.336 s). On average, the maximum ToF of any single jump during the test was 1.735 ± 0.11 s (jump 15). 99% of the averaged maximum ToF of any single jump was calculated to determine a plateau threshold of 1.726 s. The mean ToF increased over jumps 1-10 and then plateaued from jump 11 onwards (Figure 1a). Jumps 19 and 20 declined below the plateau threshold. On average, the largest percentage changes between jumps were observed between the first seven jumps (14.2%-1.4%). From jump 7 onwards the percentage change of ToF between jumps was ≤1.0%. The average force per jump peaked at 7568 ± 1379 N (jump 13), and averaged at 7060 ± 1298 N across the 20 jumps (Figure 1b). The average contact time on the trampoline bed decreased from the initial jumps, then stabilised and averaged at 0.32 ± 0.13 s across the 20-jumps (Figure 1c).

Insert Figure 1 here

Discussion

The aim of this study was to assess the reliability of the 20-max trampoline jump test across a cohort of national and international level trampoline gymnasts. The 20-max test had excellent reliability across all ToF parameters, force and contact time measures. For HD there was no statistical difference between trials, but reliability was moderate. Importantly, no differences were identified in the 20-max repeatability scores between males and females. This is likely due to the homogenous nature of the cohort of gymnasts and the variability beteeen tests for each athelte was similar. In addition, the training status and familarity with the 20-max test was similar between sexes and hence varaibility was small. Consequently, for the purpose of assessing the reliability, the data were pooled for 20-max test variables. With a large cohort of elite athletes, these data provide important insights into the characteristics of the 20-max jump test in an international and national level cohort, and the excellent reliability supports its use as a monitoring tool in this athlete cohort.

There was a small systematic negative bias between ToF measures between trials, which was unexpected. The HD data suggests this is not attributable to performance errors, as there was no difference in HD measures between trials, indicating little deviation or movement error from the centre of the bed. The decline in total ToF between trials was small (-1.2%), and likely explained by biological rather than technical error. The HDTS system is calibrated prior to use with no weight on the trampoline bed. Following this initial calibration, the HDTS software assesses its state and adjusts the calibration if required; it is unlikely that any systematic error was present because of this process. Conversely, although all trials were separated by 24-72 hours, the study was conducted during national training camps, so it was not always possible to completely abstain exercise during this period. Twenty-seven of the gymnasts completed the two trials separated by 24 hours, in 20 of these athletes total ToF performance declined (−0.38 ± 0.52 s). The remaining 5 gymnasts (who were all members of the senior elite squad) completed the trials with 48-72 hours in between, with ToF measures decreasing in 4 gymnasts (−0.29 ± 0.45 s). Considering the typical training demands of national camps, the bias was possibly attributable to residual fatigue, although this was not possible to assess. Whilst marginal performance decrements in ToF were evident between trials, this might may heighten the appeal of the 20-max test, demonstrating its sensitivity to detect small physical performance changes in ToF. Most importantly, the random error in the 20-max test was low, as shown by high ICC (0.96) and low CV (1.3%) values. In addition, effect size analyses of total ToF revealed trivial differences between test-retest trials (−0.19). These findings are important given the paucity of literature, particularly surrounding performance testing measures for the Olympic sport of trampoline, and suggest the 20-max test can provide a reliable and sensitive measure to detect changes in ToF.

Only Jensen et al. [4] study has assessed the reliability of the 20-max jump test and reported CV values of 0.1%, which is lower than the total ToF CV values (1.3%) reported in this study. Only seven male senior and junior Danish national squad trampoline gymnasts were recruited [4], compared to 32 male and female trampoline international and national level gymnasts recruited in this study. The seven gymnasts were recruited as part of a wider study assessing physiological responses and performance in a simulated trampoline competition. Jensen et al. [4] recruited a smaller cohort of anthropometrically homogenous male gymnasts who had lower inter-subject variability in 20-max total ToF performance, as indicated by a smaller standard error of the mean (0.52 s) than the standard deviation reported in this study (1.98 s). Furthermore, gymnasts performed three trials each separated by one week. This suggests gymnasts might have been less susceptible to experiencing residual fatigue than in this study. Whilst, Jensen et al. [4] research established the reliability of the 20-max trampoline test, they only did so with a small group of anthropometrically homogenous male gymnasts.

Jensen et al. [4] measured 20-max ToF using a video camera system and analysis software, as opposed to an instrumented HDTS device. A mean total ToF value of 38.69 ± 0.58 s was reported across the three trials [4], which was considerably greater than the reported average in this study (31.80 ± 1.98 s). In the current study, ToF was measured electronically using an FIG approved force plate system, which only measured 'time spent in the air', not the contact time on the bed. The 20-max total ToF values reported in Jensen et al. [4] were greater than observed in the current study (+6.89 s), which is likely attributable to the inclusion of bed contact time added to time spent in the air. The average contact time observed in the study was 0.32 s per jump. Over the course of 19 contacts on the bed this equated to approximately 6.08 s, which explains the observed difference between the 20-max total ToF variables. Nonetheless, the high ICC and low CV values reported in the current study adds substantially to previous work, in establishing the 20-max test as a reliable method to assess ToF across a wider range of national and international level gymnasts.

The data provided here also provide new insight into the characteristics of the 20-max test in international and national level trampoline gymnasts. The first half of the 20-max test, jumps 1-10, is where the gymnast aims to increase ToF to maximum values. In the second half of the

20-max test, jumps 11-20 approximately represents the portion of the test where the gymnast is aiming to maintain their ToF. The ToF during the 20-max test was shown to plateau from jump 11, indicating ToF thereafter remained within a threshold of 99% of the gymnasts maximum ToF from any single jump until jump 18. The average ToF for jumps 18 and 19 declined marginally below the 99% threshold. In a competition, the gymnast is not limited by the number of jumps they can utilise prior to the performance of their functional skills, they are instead restricted by a time limit. However, when analysing the routines of the male and female individual trampoline medallists from the Rio Olympic Games, the average number of in-jumps used in the final voluntary routine was 11, with a range of 10-14. This supports the idea that jumps 1-10 are a good representation of competition in-jumps. Practically, using more than 10 in-jumps could be ill-advised as the data also suggested ToF marginally decreased below the 99% plateau threshold from jump 18 onwards.

This study provided new insight into force and contact time measures which are not constituent parts of the trampoline scoring system, but are inextricably linked to ToF. Force data are used to estimate the timings of take-off and landing for each jump, which, in turn is used to calculate ToF. Therefore, as expected, force data followed a similar trend to ToF. Average contact times of 0.32 s were in line with reported values of upwards of 0.30 s [10]. These additional measures of force and contact time calculated by the HDTS system were shown to have excellent reliability between sessions, enhancing the overall reliability of the 20-max test and providing alternative performance measures to collect and assess, which could help practitioners understand any training-induced changes in ToF.

Conclusion

The test-retest reliability measurements of time of flight, average and peak force, and contact time from a 20-max trampoline test across a cohort of international and national level gymnasts is excellent. This study adds to the currently sparse body of trampoline gymnastics literature, and supports the use of the 20-max test as a reliable performance monitoring tool for measuring ToF. Practically, the test might also provide a reliable outcome measure to assess physiological variables of trampoline performance, or the effect of training interventions.

Acknowledgements

This work was supported by a collaborative doctoral research project funded between Northumbria University and British Gymnastics in collaboration with the English Institute of Sport. The authors would like to thank all involved in the British Gymnastics trampoline programme for their involvement in the study.

Disclosure statement

The authors report no conflict of interest.

References

[1] Fédération Internationale de Gymnastique (FIG) Executive Committee. Code of Points 2017-2020 -Trampoline Gymnastics [Internet]. Moutier; FIG [updated 2018 January 18; cited 2020 April 1]. Available from http://www.fig-gymnastics.com/publicdir/rules/tra/TRA-CoP_2017-2020e.pdf.

[2] Heinen T and Krepela F. Evaluating routines in trampoline gymnastics. Sci Gymn J. 2015; 8(3): 229–238.

[3] Harden M and Earnest CP. The Effect of Warm-Up Modalities on Trampoline Flight Time Performance. Cent Eur J Sport Sci Med. 2015; 10(2): 33–43.

[4] Jensen P, Scott S, Krustrup P et al. Physiological responses and performance in a simulated trampoline gymnastics competition in elite male gymnasts. J Sports Sci. 2015; 31: 1761–1769.

[5] Ferger K, Hackbarth M, Mylo MD, et al. Measuring temporal and spatial accuracy in trampolining. Sports Eng. 2019; 22: 1–7.

[6] Hopkins WG. Scale of magnitudes for effect statistics [Internet]. [cited 2020 April 1] Available from http://www.sportsci.org/resource/stats/.

[7] Hopkins, WG. Spreadsheets for analysis of validity and reliability. Sportscience. 2015; 19: 36–42.

[8] Atkinson G and Nevill A. Typical error versus limits of agreement. Sports Med. 2000; 30: 375.

[9] Buchheit M, Lefebvre B, Laursen PB, et al. Reliability, usefulness, and validity of the 30-15 Intermittent Ice Test in young elite ice hockey players. J Strength Cond Res. 2011; 25(5): 1457–1464.

[10] Kraft M. A simple approach for the vertical force of the trampoline bed [online]. Braunschweig: Bundesforschungsanstalt für Landwirtschaft; 2001 [cited 2020 April 1].

Table 1

Table 1. 20-max outcome measures with reliability statistics across two trials (n= 32). Data presented as mean ± standard deviation.

	20-max test			Change in mean (95% CI)		Reliability statistics (95% CI)		
Variable	**1**	**2**	**p value**	**% change**	**Effect size**	**ICC**	**TE**	**CV%**
Total ToF (s)	31.80 ± 1.98 (28.76 – 35.49)	31.43 ± 1.99 (28.63-35.81)	0.001*	−1.2 (−1.82 to −0.53)	−0.19 (−0.30 to −0.08)	0.96 (0.92- 0.98)	0.41 (0.32-0.54)	1.3 (1.0-1.7)
1-10 ToF (s)	14.40 ± 0.96 (12.99 – 16.32)	14.24 ± 0.96 (12.65-16.13)	0.006*	−1.2 (−1.96 to −0.35)	−0.18 (−0.30 to −0.05)	0.95 (0.90- 0.97)	0.23 (0.18-0.30)	1.6 (1.3-2.2)
11-20 ToF (s)	17.40 ± 1.13 (14.06-19.52	17.19 ± 1.12 (15.59-19.67)	0.001*	−1.2 (−1.89 to −0.51)	−0.19 (−0.30 to –0.08)	0.96 (0.92-0.98)	0.24 (0.19-0.31)	1.4 (1.1-1.8)
Average force (N)	7104 ± 1272 (4688 – 10091)	6977 ± 1300 (4512 – 10226)	<0.001*	−1.9 (−2.76 to -0.99)	−0.10 (−0.15 to −0.05)	0.99 (0.98-1.00)	124.8 (100.1-166.0)	1.8 (1.4-2.4)
Peak force (N)	7757 ± 1381 (5185 – 11028)	7648 ± 1416 (4926 ± 11245)	0.005*	−1.5 (−2.45 to −0.59)	−0.08 (−0.13 to 0.03)	0.99 (0.98-1.00)	143.8 (115.3-191.2)	1.9 (1.5-2.5)
Av contact time (s)	0.320 ± 0.01 (0.27 – 0.34)	0.321 ± 0.01 (0.27 – 0.34)	0.006*	0.4 (0.10 to 0.62)	0.09 (0.03 to 0.16)	0.99 (0.97-0.99)	1.59 (1.27-2.11)	0.5 (0.4-0.7)
HD (cm)	41.26 ± 10.92 (14.51 ± 62.32)	38.79 ± 9.97 (26.97 – 62.10)	0.18	−5.2 (−13.86 to 4.26)	0.33 (−0.81 to 0.16)	0.54 (0.24-0.75)	7.2 (5.7-9.6)	20.6 (16.2-28.3)

ICC = intraclass correlation coefficient; CI = confidence intervals; TE = typical error; CV = coefficient of variation.

Figure 1 Legend

Figure 1. The mean ±SD across two trials of the 20-max trampoline test in elite trampoline gymnasts for the time of flight; the dotted line represents 99% of the maximum time of flight as an idictor of the jumps that are very close to maximum (Panel A); force (Panel B); contact time on the tramploine bed for each jump.