# Methodological aspect of evaluation of the reliability the 3-Minute Burpee Test

**Robert Podstawski[1ABCDE], Piotr Markowski[2CE], Dariusz Choszcz[2BC], Jarosław Klimczak[3BD], Oscar Romero Ramos[4DE], Rafael Merino Marban[4DE]**

[1] Department of Physical Education and Sport, University of Warmia and Mazury, Olsztyn, Poland
[2] Faculty of Technical Sciences, Department of Heavy Duty Machines and Research Methodology, University of Warmia and Mazury, Olsztyn, Poland
[3] Faculty of Environmental Sciences, Department of Tourism, Recreation, and Ecology, University of Warmia and Mazury, Olsztyn, Poland
[4] Faculty of Educational Sciences, University of Málaga, Malaga, Spain

**Authors' Contribution:**
A Study Design
B Data Collection
C Statistical Analysis
D Manuscript Preparation
E Funds Collection

**Source of support:** Departmental sources

**Received:** 30 July 2016; **Accepted:** 26 August 2016; **Published online:** 20 September 2016

**AoBID:** 11441

## Abstract

**Background & Study Aim:** Reliability is one of the key adequacy criteria for evaluating motor ability tests and their relevance for physical fitness evaluations and also a verified "training trend". The aim of this study was the reliability of the 3-Minute Burpee Test (3MBT) as criterion of verification of "training trend" young women.

**Material & Methods:** The study involved 72 female students of the University of Warmia and Mazury in Olsztyn (Poland), aged 19-21 years (mean 20.6 ±0.66 years). The reliability of 3MBT was determined based on the calculated values of correlation coefficients between the average number of cycles completed by the participants in 11 repetitions of the test. The «training trend» hypothesis was verified by the Friedman test and post-hoc tests. Heart rate (HR) was measured in the first and eleven repetitions (immediately before and after the test, and 30- , 60- and 90 seconds after the test), and the differences between the measured values, including an increase in HR after the test and a decrease in HR 90 s after the test, were determined by the Wilcoxon signed-rank test.

**Results:** The values of correlation coefficients between successive repetitions of the 3MBT are within the range of very low reliability (0.70 to 0.79) to perfect reliability (0.95 to 0.99). Statistically significant differences between the number of completed cycles in successive trials of the 3MBT point to the presence of a "training trend". These differences are observed up to the fifth repetition of the 3MBT.

**Conclusions:** The 3MBT can be reliably used to assess endurance and strength abilities among young women. The "training trend" has to be eliminated before the test is used in evaluations of endurance-strength abilities. A minimum of 5 repetitions should be performed for the final test to deliver reliable results.

**Keywords:** endurance and strength abilities • MET • training trend • validity

**Corresponding author:** Robert Podstawski, Department of Physical Education and Sport, University of Warmia and Mazury in Olsztyn, Prawocheńskiego 7, 10-720 Olsztyn, Poland; e-mail: podstawskirobert@gmail.com



This is an open-access article distributed under the terms of the Creative Commons Attribution-Non-commercial 4.0 International (http://creativecommons.org/licenses/by-nc/4.0/), which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non-commercial and is otherwise in compliance with the license.

**Test-retest method** – method of determining stability in which a test is given one day and then administered exactly as before a day or so later [3, p. 200].

**Accuracy** – the quality or state of being correct or precise (*Oxford Dictionaries*); **accuracy** in parts of methodological books is synonymous of **validity** (degree to which a test or instrument measures what it purports to measure; can be categorized as *logical, content, criterion, or construct validity* [3, p. 193]), whereas **validity** (relevance) include **accuracy** and **reliability** [5].

**MET** – *abbreviation* **metabolic equivalent** – *noun* a unit used for expressing the resting [38].

**Position** – *noun* **1**.the place where a player is standing or playing **2.** the way in which a person's body is arranged [38].

**Posture** – *noun* the position in which a body is arranged, or the way a person usually holds his or her body when standing [38].

**Body mass index** – *noun* an index that expresses adult weight in relation to height, calculated as weight in kilograms divided by height in metres squared. Abbreviation BMI (Note: A body mass index of less than 25 is considered normal, and one of over 30 implies obesity) [38].

## Introduction

Reliability is one of the key adequacy criteria for evaluating motor ability tests and their relevance in physical fitness (PF) evaluations [1-6]. According to Osiński [7], a measurement can be a reliable indicator of motor ability if it meets test adequacy criteria such as validity, objectivity, standardization and normalization. According to Pilicz [8], test adequacy criteria also include time-effectiveness and cost-effectiveness.

Reliability is defined as the overall consistency of a standard measure or the error of measurement. The reliability of a test provides information about the magnitude of measurement error; therefore, the concept of measurement reliability is linked with the consistency of results scored by an individual or a group in a motor test during repeated measurements [2, 4, 7]. A reliable test should generate identical or very similar results if it is repeated at least twice by the same individuals and under the same conditions [8]. The consistency of a measurement can be determined by calculating the reliability coefficient when a random group of participants perform a given motor task several times. In most cases, reliability is determined by calculating the coefficient of correlation between two repeated measurements [7]. According to Starosta [9], motor abilities should be assessed in motor tests based on the following estimated intervals of reliability (correlation) coefficients: 0.95-0.99 perfect reliability, 0.90-0.94 high reliability, 0.80-0.90 acceptable reliability, 0.70-079 very low reliability, 0.60-0.69 doubtful reliability for individual assessments, acceptable for group comparisons. A correlation coefficient of 0.70 is generally regarded as satisfactory.

The reliability of a test can be improved by increasing the number of repetitions (test-retest method) [1, 7]. In reliability assessments performed with the use of the test-retest method, special attention should also be given paid the "training trend" or the "fatigue trend" which was first observed by Denisiuk [10, 11]. The cited author demonstrated that daily measurements of physical fitness levels, performed over a period of 12 days, produced different results in the same individuals. A "training trend" is often observed when an individual's motor abilities improve as a result of repeated fitness trials despite relatively high values of the correlation coefficient in tests with a low reliability coefficient [8].

This phenomenon was also reported by Podstawski et al. [12, 13]. A "training trend" is also noted when correlation coefficients are analysed simultaneously with significant differences between means. To eliminate such errors, successive trials should take place at intervals which ensure that the participants' fitness levels will not improve significantly with every repetition and will not deteriorate due to fatigue.

Numerous researchers have noted that new evaluation methods should be tested and that the reliability of the existing motor ability tests has to be verified [2, 4, 7, 14]. This follow-up study investigates the usefulness of the 3-Minute Burpee Test (3MBT) in evaluations of the endurance-strength abilities of individuals from different age groups. The 3MBT has been used to test elementary school students, university students and early education teachers [15]. The results scored by female university students were used to develop classification standards for the 3MBT [16]. A study of female university students demonstrated that body mass and BMI are significantly negatively correlated with 3MBT scores [17]. Successive studies confirmed that body mass, BMI and height are negatively correlated with the number of completed cycles, and that 3MBT is a highly intense physical activity during which the average heart rate (HR) of female university students was determined at 181.92 b·min$^{-1}$ [18].

The aim of this study was the reliability of the 3-Minute Burpee Test (3MBT) as criterion of verification of "training trend" young women.

## Materials and methods

### Participants

The study was conducted on 72 first-year full-time female students of the University of Warmia and Mazury in Olsztyn (Poland) aged 19-21 years (mean 20.6 ±0.66 years). The students who agreed to participate in the study did not attend obligatory physical education classes (90 minutes per week), but performed one 3MBT per week (experiment classes). Differences in the physical activity levels of students participating in extracurricular sports could significantly distort the examined relationships, therefore, students involved in additional sports activities and students released from physical education classes for medical reasons were excluded from the study. All participants resided permanently in the Region of Warmia and Mazury in Poland.

### Procedures

The study was performed in compliance with the Declaration of Helsinki and upon the prior consent of the Bioethical Committee of the University of Warmia and Mazury in Olsztyn. All participants gave their written consent to the study.

### Measurements

The research goal was pursued by determining: the values of correlation coefficients between the results scored in the eleventh repetition of the 3MBT; significant differences between the results (average number of claps) scored in each of the 11 repetitions that would point to the presence of a "training trend"; the number of cycles and HR values (before the test, immediately after the test, and 30-, 60- and 90 seconds after the test) in the first and eleventh repetition of the 3MBT.

Female students' physical activity levels were evaluated with the use of the Polish version of the standardized and validated International Physical Activity Questionnaire (IPAQ) [19]. The energy expenditure of participants did not exceed 600 METs per week, and none of the tested female students met the "sufficient physical activity" category. Questionnaire results revealed that the analysed female students were characterized by low levels of physical activity and sedentary behaviour. IPAQ was used only to select a homogenous sample of female students, and its results are not presented in this study. The students' endurance-strength abilities were evaluated in the 3MBT [17]. All participants performed the test 11 times at weekly intervals.

### Description of trials – phases of the 3MBT

**Phase I:** from the initial upright standing position, the participant assumes a supported squatting position (posture) with both hands on the ground;

**Phase II:** from the supported squat, the feet are thrust backwards into the push-up position with straightened arms;

**Phase III:** from this position, the participant once again returns to the supported squat;

**Phase IV:** the cycle is completed by the participant returning to the upright standing position and simultaneously clapping his or her hands over the head, making sure that the arms remain extended. The cycle is performed as many times as possible in the given time limit (3 minutes).

### Comments

The participants should be closely monitored to ensure that the exercise is performed properly and that each phase of the exercise is performed in the correct order. The participants should not bend their arms or arch their backs in the push-up position, although an exception can be made for those with low upper body strength. The legs must be fully extended every time the participant is in the push-up position. A cycle is not counted if any of the phases is not performed correctly [17].

Heart rate was measured with the Polar RS 100 HR monitor with a chest strap. The scores measured immediately before the test, immediately after the test, and 30-, 60- and 90 seconds after resting in a sitting position were read out by the tested student. Every student was shown how to perform the 3MBT during the experiment class before the test and was given ample time to practice. The participants performed an active warm up for 10 minutes before the test. The warm-up routine was identical for all students, and it comprised selected physical exercises such as jogging, arm and hip circles, leg swings, balance exercises, 10-meter dashes, stretching and corrective drills [20].

### Statistical analysis

The results were processed statistically to calculate minimum, mean and maximum values, median values, standard deviation, coefficient of variation for dependent variables (*HR immediately before and after the test, and 30-, 60- and 90 seconds after the test, increase in HR and decrease in HR after the first and eleventh repetition of the test*). The significance of correlations between the average number of cycles completed in each trial was determined based on the values of correlation coefficients.

The presence of correlations between successive trials (independent variable) of the 3MBT, performed at weekly intervals, and the average number of cycles (dependent variable) completed in a given trial, which would testify to the presence of a "training trend", was also investigated. Analysis of variance (ANOVA) was performed for the dependent variable (number of cycles).The results were processed by repeated measures ANOVA (parametric test) or the Friedman (non-parametric) test, subject to the fulfilment of adequacy criteria.

Additional analyses were performed to determine the significance of differences in the average number of cycles completed in each repetition. The relevance of a given test (parametric or non-parametric) for calculations was verified by checking the fulfilment of adequacy criteria. Research hypotheses were verified by determining whether every test (parametric or non-parametric) fulfilled the adequacy criteria for calculations.

If statistically significant differences were found, post-hoc tests were used to determine differences in the average number of completed cycles between successive repetitions. The significance of differences between the increase and decrease in HR in the first and eleventh repetition was verified in the Wilcoxon signed-rank test (comparison of two dependent samples). Duncan's post-hoc test was used to identify trials that differed significantly in the number of cycles. Calculations were performed at a significance level of a = 0.05 in the Statistica v. 12 software package (Statsoft, US) [21].

## Results

Significant differences in HR were not noted only before the test in the first (T1) and the eleventh (T11) trial (Table 1). Immediately after the test and 30 s after the test, HR values were significantly higher (p=0.0000) in the eleventh trial (T11, $T11_{30}$) than in the first trial. HR values determined 60 s and 90 s after the test ($T11_{60}$, $T11_{90}$) were significantly lower (p=0.0000) in the eleventh trial than in the first trial. The students completed significantly more cycles in the eleventh trial (45.18 cycles) than in the first trial (36.79 cycles). The increase in HR was significantly lower (p = 0.0000) in the first trial ($T1 – T1_0$ = 71.87 b×min$^{-1}$) than in the eleventh trial ($T11 – T11_0$ = 79.06 b×min$^{-1}$). The decrease in HR was also significantly lower (p = 0.0000) in the first trial ($T1_0 – T1_{90}$ = 31.04 b×min$^{-1}$) than in the eleventh trial ($T11_0 – T11_{90}$ = 37.73 b×min$^{-1}$).

**Table 1.** Statistical indicators of heart rate measured immediately before and after the test.

| Variable | Descriptive statistics | | | | | |
|---|---|---|---|---|---|---|
| | Mean | Median | Min | Max | SD | CV |
| **T1** //before the 1st trial (beats per min) | 110.13 | 114.00 | 78.00 | 144.00 | 16.94 | 15.38 |
| **$T1_0$** //immediately after the 1st trial (beats per min) | 182.00 | 183.50 | 145.00 | 210.00 | 12.87 | 7.07 |
| **$T1_{30}$** //30 s after the 1st trial (beats per min) | 171.32 | 172.00 | 140.00 | 195.00 | 12.11 | 7.07 |
| **$T1_{60}$**//60 s after the 1st trial (beats per min) | 161.92 | 163.00 | 132.00 | 186.00 | 12.35 | 7.63 |
| **$T1_{90}$**//90 s after the 1st trial (beats per min) | 150.96 | 151.50 | 112.00 | 176.00 | 13.10 | 8.68 |
| **P1** // number of cycles in 3MBT | 36.79 | 36.50 | 12.00 | 60.00 | 9.60 | 26.09 |
| **T11** //before the 11th trial (beats per min) | 109.50 | 112.00 | 78.00 | 142.00 | 15.87 | 14.49 |
| **$T11_0$** //immediately after the 11th trial (beats per min) | 186.56 | 186.50 | 149.00 | 210.00 | 12.82 | 6.87 |
| **$T11_{30}$** //30 s after the 11th trial (beats per min) | 172.64 | 174.00 | 140.00 | 192.00 | 12.34 | 7.15 |
| **$T11_{60}$** //60 s after the 11th trial (beats per min) | 161.10 | 162.00 | 129.00 | 186.00 | 12.50 | 7.76 |
| **$T11_{90}$** //90 s after the 11th trial(beats per min) | 148.83 | 149.00 | 116.00 | 171.00 | 12.87 | 8.64 |
| **P11** //number of cycles in 3MBT | 45.18 | 46.00 | 14.00 | 69.00 | 9.51 | 21.05 |
| **Mean values of the measured indicators in female students** | | | | | | |
| HR before the test | p = 0.2080; | | | | | t = 1.2706 |
| HR after the test | p = 0.0000; | | | | | t = –10.5356 |
| HR 30 s after the test | p = 0.0000; | | | | | t = 7.8236 |
| HR 60 s after the test | p = 0.0000; | | | | | t = 12.2981 |
| HR 90 s after the test | p = 0.0000; | | | | | t = 19.3847 |
| Number of cycles | p = 0.0000; | | | | | t = –12.1496 |
| Difference in HR increase between P1 and P2 | p = 0.0000; | | | | | t = 6.298640 |
| Difference in HR decrease between P1 and P2 | p = 0.0000; | | | | | t = 7.205269 |

***Note:*** **SD** standard deviation, **CV** coefficient of variation, **p** probability computed in t-test t, **t** -value

**Table 2.** Coefficients of correlation between successive repetitions of the 3MBT.

| Repetition | Coefficients of correlation between successive test trials, data are significant at p<0.5, N = 72 (cases with missing data were eliminated) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | R1 | R2 | R3 | R4 | R5 | R6 | R7 | R8 | R9 | R10 | R11 |
| R1 | 1.0000 | 0.8443 | 0.8762 | 0.8425 | 0.8415 | 0.8095 | 0.7730 | 0.8239 | 0.8232 | 0.7998 | 0.8104 |
| R2 | | 1.0000 | 0.9317 | 0.8917 | 0.8892 | 0.8465 | 0.8131 | 0.8572 | 0.8421 | 0.8561 | 0.8694 |
| R3 | | | 1.0000 | 0.9153 | 0.9222 | 0.8643 | 0.8472 | 0.8933 | 0.8692 | 0.8740 | 0.8944 |
| R4 | | | | 1.0000 | 0.9447 | 0.8982 | 0.8582 | 0.8954 | 0.8915 | 0.8982 | 0.9163 |
| R5 | | | | | 1.0000 | 0.9293 | 0.8786 | 0.9229 | 0.8865 | 0.9152 | 0.9233 |
| R6 | | | | | | 1.0000 | 0.9064 | 0.9156 | 0.8795 | 0.9285 | 0.9333 |
| R7 | | | | | | | 1.0000 | 0.8830 | 0.8688 | 0.9184 | 0.9317 |
| R8 | | | | | | | | 1.0000 | 0.9189 | 0.921052 | 0.9224 |
| R9 | | | | | | | | | 1.0000 | 0.9073 | 0.9176 |
| R10 | | | | | | | | | | 1.0000 | 0.9515 |

**Table 3.** The results of Mauchly's sphericity test – pooled sample.

| Effect | Mauchly's sphericity test, sigma-restricted parametrization, effective hypothesis decomposition | | | |
|---|---|---|---|---|
| | Kendall'sW | $\chi^2$ statistic | df | p |
| R1 | 0.079608 | 170.7342 | 54 | 0.0000 |

**Table 4.** Average number of cycles completed in successive trials.

| Variable | Friedman's ANOVA, Kendall's coefficient of concordance, chi-squared ANOVA (Women, N = 72 , df 10) = 243.4911, p = 0.00000, Kendall's W = 0.33818, mean R rank = 0.32886 | | | |
|---|---|---|---|---|
| | Mean rank | Rank sum | Mean | SD |
| R1 | 1.861111 | 134.0000 | 36.79167 | 9.597443 |
| R2 | 3.208333 | 231.0000 | 39.02778 | 9.172789 |
| R3 | 5.618056 | 404.5000 | 42.31944 | 9.110938 |
| R4 | 5.583333 | 402.0000 | 42.56944 | 8.647765 |
| R5 | 5.847222 | 421.0000 | 42.86111 | 8.480808 |
| R6 | 7.250000 | 522.0000 | 43.95833 | 8.581732 |
| R7 | 6.986111 | 503.0000 | 43.98611 | 9.589450 |
| R8 | 7.027778 | 506.0000 | 44.06944 | 9.465023 |
| R9 | 7.159722 | 515.5000 | 43.91667 | 9.295993 |
| R10 | 7.583333 | 546.0000 | 44.58333 | 9.265642 |
| R11 | 7.875000 | 567.0000 | 44.66667 | 8.767739 |

**Table 5.** Results of Duncan's post-hoc test for differences between the average number of cycles completed in successive trials – pooled sample.

| Duncan's test; variable DV_1 (cycle) Estimated probability for post-hoc tests | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **R** | **36.792** | **39.028** | **42.319** | **42.569** | **42.861** | **43.958** | **43.986** | **44.069** | **43.917** | **44.583** | **44.667** |
| **Trial** | **I** | **II** | **III** | **IV** | **V** | **VI** | **VII** | **VIII** | **IX** | **X** | **XI** |
| I | | 0.000028 | 0.000011 | 0.000003 | 0.000004 | 0.000004 | 0.000005 | 0.000001 | 0.000004 | 0.000001 | 0.000001 |
| II | | | 0.000009 | 0.000011 | 0.000003 | 0.000004 | 0.000004 | 0.000005 | 0.000004 | 0.000001 | 0.000001 |
| III | | | | 0.633656 | 0.333842 | 0.003859 | 0.003748 | 0.002499 | 0.004172 | 0.000063 | 0.000036 |
| IV | | | | | 0.578199 | 0.013665 | 0.013743 | 0.009923 | 0.013897 | 0.000396 | 0.000236 |
| V | | | | | | 0.046863 | 0.048706 | 0.038029 | 0.044195 | 0.002643 | 0.001736 |
| VI | | | | | | | 0.957773 | 0.843730 | 0.936696 | 0.283549 | 0.236139 |
| VII | | | | | | | | 0.873789 | 0.901966 | 0.286073 | 0.242054 |
| VIII | | | | | | | | | 0.795168 | 0.327256 | 0.286073 |
| IX | | | | | | | | | | 0.265638 | 0.218680 |
| X | | | | | | | | | | | 0.873789 |

An evaluation of the test's relevance for assessments of motor abilities revealed that the coefficients of correlation between successive repetitions of the 3MBT were within the range of very low reliability (0.70-0.79) and perfect reliability (0.95-0.99) (Table 2).

Significant differences in the number of completed cycles cannot be determined based on the coefficients of correlation between successive repetitions of the 3MBT. The results of preliminary calculations (Mauchly's sphericity test) revealed that the condition of *sphericity* was not met for the compared trials (Table 3).

The results indicate that the condition of sphericity was not met for the compared 3MBT trials ($p< 0.05$). Therefore, the research hypotheses were verified with the use of the non-parametric Friedman test. The test revealed significant differences ($p<0.05$) between the average number of cycles in successive trials (Table 4).

The results of the analysis (by Duncan's post-hoc test) point to significant differences in the number of completed cycles up to the fifth repetition of the 3MBT (Table 5).

## Discussion

The study conducted in line with the adopted procedural algorithm confirmed the presence of a "training trend", previously described by Denisiuk [10]. Significant differences in the average number of completed cycles were noted up to the fifth repetition of the 3MBT. The presence of a "training trend" was also validated by HR values which were significantly higher immediately after the test and 30 s after the test in the first trial than in the eleventh trial. Heart rates determined 60 s and 90 s after the test were significantly lower in the eleventh trial than in the first trial. The presence of a "training trend" was also confirmed by differences in the increase and decrease in HR, which were significantly higher in the eleventh trial. The tested students performed the 3MBT at weekly intervals and did not participate in additional PAs, therefore, based on the international PA guidelines for persons aged 19-21 years [22], an improvement in test results resulting from higher PF levels can be ruled out.

The 3MBT evaluates endurance-strength abilities which, according to Raczek's theory of motor skills, fall in the category of hybrid abilities [23]. Endurance-strength abilities can also be evaluated in a 500 m test on a rowing ergometer [24]. An evaluation of the reliability of a 500 m test on a rowing ergometer produced highly similar results. In a study of university students who rowed 500 m on an ergometer every week, a "training trend" was observed up to the fourth trial in female students and up to the fifth trial in male students [12, 13].

The duration of exercise significantly influences the reliability of a test. In long tests, such as endurance runs, the measurement error is compensated by other factors (such as cardiovascular endurance), which contributes to a test's reliability [7]. A similar dependency was observed in the 3MBT. Very high HR values (around 182 b·min$^{-1}$) in female students immediately after the 3MBT indicate that the exercise involved high levels of exertion. This assumption is validated by high HR values noted 90 s after exercise in the first and eleventh trial (150.96 and 148.83 b·min$^{-1}$).

Burpee test in 30 second version, repeated cyclically (4 times divided by a 60-second break) was used to the comprehensive assessment of an anaerobic capacity and precision skills before and during activity [25] in the framework of the SPHSA questionnaire [26]. SPHSA means sense of positive health and survival abilities. Past research reports using SPHSA questionnaire inform the declared profile of SPHSA (subjective assessment), generally young adult women [26-29] and only 69 males [30]. Developed standards for adult men and women [25] concern the measurement of somatic health (dimension A) and survival abilities (dimension D).

The second aspect of our research concerns the need to monitor the training load (workload) during test researches and training. Review of the recommended methods [31-34] shows the great diversity and accuracy of measurement. This issue is also important as monitoring the methods of selection the right people for the professional practicing sport [5, 35] and the training means not only from the perspective of sport success, but also the health [6, 36-38].

### Practical implications and limitations

The coefficients of correlation between successive repetitions of the 3MBTwere within the range of very low (0.80-0.90), to acceptable (0.90-0.94), and even perfect (0.95-0.99) reliability. Classification standards for various age groups should be developed for the 3MBT based on representative samples. To date, such standards have been developed only for women aged 19 to 26 years [16]. The development of classification standards for all age groups would increase the rank and relevance of the 3MBT.

## Conclusions

The 3MBT is a reliable test that can be used to evaluate the endurance-strength abilities of young women. The values of correlation coefficients for the 3MBT are within the range of very low to perfect reliability in the studied group of female university students.

The results of statistical analyses revealed significant differences in the number of claps between successive repetitions of the 3MBT, which points to the presence of a "training trend" (up to the fifth trial).

The "training trend" has to be eliminated before the 3MBT is used in evaluations of endurance-strength abilities. A minimum of 5 repetitions should be performed for the final test to deliver reliable results.

## Acknowledgements

The authors would like to thank all students who volunteered for the study.

## Competing interests

The authors declare that they have no competing interests.

## References

1. Thomas S, Nelson JK. Research methods in physical activity. Champaign: Human Kinetics Publishers; 1990
2. Kroes M, Vissers YLJ, Slejpen FAM et al. Reliability and validity of a qualitative and quantitative motor test for 5- to 6-year-old children. Eur J Pediatr Neurology 2004; 8(3): 135-143
3. Thomas JR, Nelson JK, Silverman SJ. Research Methods in Physical Activity :5th edition. Windsor: Human Kinetics Publishing; 2005
4. Hraski M, Hovat V, Hovat I. Reliability of instrument for assessing standing long jump performance in preschool children. In: Milanović D, Sporiš G, eds. Integrative Power of Kinesiology. Zagreb: University of Zagreb Press; 2011: 235-237
5. Kalina RM, Jagiełło W, Chodała A. The result of "testing fights in a vertical posture" as a criterion of talent for combat sports and self-defence – secondary validation (part I: the reliability). Arch Budo Sci Martial Art Extreme Sport 2015; 11: 229-238
6. Mosler D. Validity and reliability of non-apparatus and quasi apparatus flexibility tests – verification during health-related training based on judo. Arch Budo Sci Martial Art Extreme Sport. 2015; 11: 123-133
7. Osiński W. Antropomotoryka. Poznań: AWF; 2003 [in Polish]
8. Pilicz S. Pomiar Ogólnej Sprawności Fizycznej. Studia i Monografie 65. Warszawa: AWF; 1997 [in Polish]
9. Starosta W. Globalna i lokalna koordynacja ruchowa w wychowaniu fizycznym i sporcie. Warszawa: Osgraf; 2006 [in Polish]
10. Denisiuk L. Badania nad wartością niektórych prób sprawności fizycznej. Wychowanie Fizyczne i Sport 1961; 1: 327-363 [in Polish]
11. Denisiuk L. The Current state of Research in Physical Fitness of School children in Poland. Warsaw: Poland Institute for Research in Physical Culture; 1962: 69-83
12. Podstawski R, Choszcz D, Wysocka-Welanc M. Próba opracowania metody trafności testu wytrzymałości czasu krótkiego oraz analiza wpływu treningu na czas pokonania dystansu ergometrem wioślarskim na przykładzie studentów Uniwersytetu Warmińsko-Mazurskiego w Olsztynie. Antropomotoryka 2009; 46: 55-64 [in Polish]
13. Podstawski R, Choszcz D, Konopka S. The impact of training on the 500 m rowing ergometer time and the assessment of the applied test's relevancy. Human Movement 2011; 12(3): 264-272

14. Žuvela F, Božanić A, Miletić D. POLYGON – A new fundamental movement skills test for 8 year old children: construction and validation. J Sport Sci Med 2011; 10: 157-163
15. Podstawski R, Honkanen A, Boraczyński T et al. Physical fitness classification standards for Polish early education teachers. SAJRSPER 2015; 37(1): 113-130
16. Podstawski R, Mańkowski S, Raczkowski M. The level of strength and Endurance-strength abilities of the female early education teachers as examined by the medicine ball forward throw and the 3 min. Burpee Test: A comparative analysis. LASE J Sport Sci 2014; 4(2): 93-109
17. Podstawski R, Kasietczuk B, Boraczyński T et al. Relationship Between BMI and Endurance-Strength Abilities Assessed by the 3 Minute Burpee Test. Int J Sports Sci 2013; 3(1): 28-35
18. Podstawski R, Markowski P, Choszcz D et al. Correlations between anthropometric indicators, heart rate and endurance-strength abilities during high-intensity exercise of young women. Arch Budo Sci Mar Arts Extr Sports 2016; 12: 17-24
19. Biernat E, Stupnicki R, Gajewski, K. Międzynarodowy Kwestionariusz Aktywności Fizycznej (IPAQ) – wersja polska [International Physical Activity Questionnaire (IPAQ) Polish version. Wychowanie Fizyczne i Sport 2007; 51(1): 47-54 [in Polish]
20. Frandkin AJ, Zazryn TR, Smoliga JM. Effects of warming-up on physical performance: a systematic review with meta-analysis. J Strentg Con Res 2010; 24(1): 140-148
21. Rabiej M. Statystyka z programem Statistica – Wersja Polska. Gliwice: Helion S.A.; 2012 [in Polish]
22. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med Sci Sport Exerc 2011; 43(7): 1334-1359
23. Raczek J. Antropomotoryka. Teoria motoryczności człowieka w zarysie. Warszawa: PZWL; 2010 [in Polish]
24. Podstawski R, Choszcz D, Konopka S et al. Anthropometric determinants of rowing ergometer performance in physically inactive collegiate females. Biol Sport 2014; 31(4): 315-321
25. Kalina RM, Kalina A. Methods for measurement of somatic health and survival abilities in the framework of the SPHSA questionnaire – methodological aspects. Arch Budo Sci Martial Art Extreme Sport 2013; 9: 17-30
26. Kalina RM. The profile of Sense of Positive Health and Survival Abilities indices (subjective assessment) as a diagnostic tool used in health-related training. Arch Budo 2012; 8(3): 179-188
27. Jagiełło W, Sawczyn S, Jagiełło M. The subjective profile of positive health and survival abilities in women differing as to physical activity. Arch Budo 2012; 8(4): 219-224
28. Jagiełło W, Sawczyn S, Jagiełło M et al. The subjective profile of positive health and survival abilities in female students differing as to physical activity. Teoriâ i Praktika Fizičeskoj Kultury 2013; 5: 15-18
29. Tsos A, Oliynyk OV, Szepeluk A. Subjective health profiles among Ukrainian students of medical vocational school. Health Problems of Civilization 2016; 10(1): 32-38
30. Bergier B. The diversity of the profiles involving the sense of positive health and survival abilities of Polish students of paramedical sciences. Arch Budo 2015; 11: 17-25
31. Kalina RM. Methodology of measurement, documentation and programming optimal workload continuous with variable intensity – applications in sports medicine, physiotherapy, geriatrics, health-related training, sport for all. Arch Budo 2012; 8(4): 235-249
32. Biernat E, Boguszewski D. The level of physical activity of the working inhabitants of Warsaw practising martial arts and combat sports. Arch Budo 2015; 11: 69-77
33. Franchini E, Julio UF, Gonçalves Panissa VL et al. Short-term low-volume high-intensity intermittent training improves judo-specific performance. Arch Budo 2016; 12: 219-229
34. Sterkowicz-Przybycień K, Blecharz J, Tyka AK et al. Changes in indices of physical capacity, coordination, psychomotor activity and selected psychological traits in the driver after preparation to the Dakar Rally. Arch Budo Sci Martial Art Extreme Sport 2015; 11: 29-40
35. Dias da Silva RA, Motta Drummond MD, Couto BP et al. Content validation of training means for taekwondo. Arch Budo 2015; 11: 305-317
36. Pedrosa GF, Soares YM, Gonçalves R et al. Elaboration and evaluation of judo training means. Arch Budo 2015; 11: 7-16
37. Pedrosa GF, Soares YM, Gonçalves R et al. Content validation of a catalog of exercises for judo. Percept Mot Skills 2016; 122(3): 933-955
38. Dictionary of Sport and Exercise Science. Over 5,000 Terms Clearly Defined. London: A & B Black; 2006

**Cite this article as:** Podstawski R, Markowski P, Choszcz D et al. Methodological aspect of evaluation of the reliability the 3-Minute Burpee Test. Arch Budo Sci Martial Art Extreme Sport 2016; 12: 137-144