EVALUATION OF BATTERY OF TESTS IN THE MODEL OF MONITIORING PHYSICAL FITNESS OF STUDNETS IN PHYSICAL EDUCATION IN SERBIA

IVANA MILANOVIĆ - SNEŽANA RADISAVLJEVIĆ-JANIĆ
Faculty of Sport and Physical Education
University of Belgrade
Serbia

ABSTRACT

Within the research project „*Establishment of the system of monitoring physical fitness of students in physical education classes*“ implemented by the Faculty of Sport and Physical Education in collaboration with the Institute for Assessment of Education Quality in Serbia from 2011-2014, the main aim was establishment of the system of monitoring physical fitness of students in elementary schools in Serbia. Based on this, we set the specific tasks: establishment of a physical fitness monitoring model for students that could be implemented in the Serbian educational system and the development of a battery of tests in that model.

After setting up the physical fitness monitoring model for students, we proceeded with evaluation of appropriate battery of tests which could be applied in that model. After a theoretical analysis of certain batteries of tests, structurally fit to the physical fitness monitoring model for students applicable in our educational system, we selected a set of the most frequently applied tests of motor abilities.

The assessment of reliability and sensitiveness of most frequently applied tests was performed in a sample of elementary school students from 3rd to 8th grade. In the same sample of students a test-retest method was applied – once in autumn and the second time in spring. In both cases remeasuring was carried after 15 days. The sample of subjects within the autumn testing involved 848 students (446 boys and 402 girls) attending elementary school, aged 9 to 14 years, while 834 students participated (444 boys and 390 girls) in the spring testing. The empirical research checked the reliability and sensitiveness of the following tests: 1600 m walk/run, 20 м - Shuttle run, Standing long jump, Bent-arm hang, Sit-ups for 30 s, Sit and reach, High jump and Shuttle run 4x10 m.

Based on the empirical verification of reliability and sensitivity of the most frequently applied tests in different models of physical fitness monitoring of children and youth worldwide, it can be concluded that they manifested satisfactory reliability and sensitiveness at all ages, in both genders from 3rd to 8thgrade of elementary school.

Based on the evaluation of the most frequently applied tests in different models of physical fitness monitoring of children and youth worldwide, we proposed the battery of tests that fits the physical fitness monitoring model for students in the existing conditions for realization in physical education classes in Serbia.

Key words: physical fitness, monitoring model, physical education classes

INTRODUCTION

The aim and the objectives of physical education in our country are no different in neither of the segments from the aims and the objectives of physical education in other developed countries in Europe or America (Hardman, 2008). Since the definition clearly emphasizes that “the aim of physical education, by various and systematic motor activities, in accordance with the rest of education areas, is to contribute to the integral (cognitive, affective, motor) development of student’s personality, development of motor abilities, acquirement, perfection and enforcement of motor skills, habits and necessary theoretical knowledge in both everyday and more specific life and work conditions“[14], it could be validly inferred that this way to achieve such an aim of physical education, would be more effective if the students were, in the right way and at the right time, directed to the development and the improvement of their own predispositions. Developed countries all around the world started to pay more and more attention to the role of PE in promotion of healthy lifestyle, so the PE classes in most countries are presented as one of the main factors in formation of positive relation of the youth towards physical exercises and sport (Carlson, 1995; Ennis, 1996; Portman, 1995; Robinson, 1990). According to the authors dealing with these problems, while monitoring the student’s physical fitness, a conceived PE instruction can only “survive” if it is in compliance with aim and tasks of PE instruction as promotion of healthy lifestyle, promotion of positive attitude towards physical activity, exercising and sport (Corbin et al., 1995). In other words, physical fitness testing of children and youth should, besides measuring and monitoring of certain student’s physical fitness components, be involved in students education about significance, way and method for improvement of each of the fitness components, and at the same time, it should promote positive attitudes towards physical exercises and physical activity in general (Corbin et al., 1995). In the recent years, in the USA and in some European countries, a completely new opinion of expediency of physical fitness monitoring has been established in PE instruction, mostly based on certain number of researches, which indicated a connection between physical fitness and health (Ortega et al., 2008). Therefore, certain changes have occurred even in the test batteries compared to the former ones. Most of today’s batteries for physical fitness testing consist of scoring instruments for measuring and monitoring those physical fitness components or performances whose results can indirectly be correlated with children health.

Within monitoring of physical fitness in the course of PE instruction in our country, the only organized system for monitoring of physical fitness in children and youth within the PE classes existed in Belgrade in the period from the seventies to the nineties of the XX century. Since than, until today, no other new system of continuous monitoring of physical fitness in students has been developed in the process of PE instruction. The PE professors are not obliged to continuously monitor physical fitness of their students, and only in one part of the PE Curriculum they are referred to follow-up physical fitness. The instructions set out in the Curriculum are drawn based on the concept of the aforesaid model from the seventies, and, when compared to the current world concepts, that concept and way of physical fitness monitoring, regardless of all its positive features, is outdated in many ways. Having in mind the aforesaid facts, Faculty of sport and physical education in collaboration with the Institute for assessment of education quality in Serbia started the research project „*Establishment of the system of monitoring physical fitness of students in physical education classes*“. In this project the main aim was establishment of the system of monitoring physical fitness of students in elementary schools in Serbia. Based on this, we set the specific tasks: the establishment of model of monitoring physical fitness of students that could be implemented in the Serbian educational system, the development of a battery of tests in that model and the determination of criterion reference values that would be used for the improvement of physical education.

Since the basic aim of PE instruction in our country is to contribute to the integral development of students’ personality primarily by stimulating students’ growth and development of physical fitness, the PE instruction should basically follow the concept of the so-called „health-related physical fitness batteries“, which means that test battery should contain the tests for monitoring and evaluating: cardiorespiratory endurance (assessment of aerobic capacity), body composition (particularly body fat), muscular strength, endurance and flexibility (assessment of agility level). The health-related testing concept complies with the syllabus and the aim of PE instruction in our country, so the application of this testing concept would contribute to enhancement of efficiency of the teaching process.

After setting up the model of monitoring physical fitness of students, the next task was development of an appropriate battery of tests which could be applied into that model. We analyzed the following battery of tests: The Chrysler/AAU Fitness Test, Prudential Fitnessgram-Activitygram, President’s Challenge Physical Fitness Test, YMCA Youth Fitness Tests; Physical Best, The Eurofit Physical

14 Official gazette of the RS – Educational Journal”, 10, Belgrade, 2004.

Fitness Test Battery and a test battery from HELENA study. After a theoretical analysis of those battery of tests which are structurally fit to the model of monitoring of physical fitness of students applicable in our educational system, we extracted the set of the most frequently applied tests of motor abilities from those batteries.

The aim of this part of the project was assessment of reliability and sensitiveness of those tests in a sample of elementary school students from our country. We checked the reliability and sensitiveness of the following tests: ***1600 m walk/run, 20 m - Shuttle run, Standing long jump, Bent-arm hang, Sit-ups for 30 s, Sit and reach*** and ***Shuttle run 4x10 m***. In addition to the assessment and evaluation of the basic metric characteristics of the tests, it was necessary to determine the fitness for different age and gender of respondents.

METHODS

The assessment of reliability and sensitiveness of the most frequently applied tests was performed in a sample of elementary school students from 3rd to 8th grade. In the same sample of students, a test-retest method was applied – the first time in autumn and the second time in spring. In the both cases re-measuring was carried after 15 days. The sample of subjects within the autumn testing involved 848 students (446 boys and 402 girls) attending elementary school „Ivo Andrić“ in Belgrade, aged 9 to 14 years, i.e., 834 students participated (444 boys and 390 girls) in the spring testing. The empirical research tested the reliability and sensitiveness of the following tests: 1600 m walk/run, 20 m - Shuttle run, Standing long jump, Bent-arm hang, Sit-ups for 30 s, Sit and reach and Shuttle run 4x10 m.

Statistical analysis

To check the reliability and sensitiveness of the tests we used several statistical analysis: Student t test (r), the Pearson coefficients of correlation (p) and a coefficient of variation (CV%). The level of statistical significance for these analyses was set up to p<.05. By using univariate analysis of variance with repeated measures, we analyzed if the tests are applicable to all age periods of both gender in this sample of students. The level of statistical significance was set to p < .01. Data were analyzed using SPSS 17.0 software (SPSS Inc. Chicago, IL, USA).

RESULTS AND DISCUSSION

Table 1 shows the boys results from test-retest method within the autumn testing. The most frequently applied tests of motor abilities performed in our sample within the autumn testing manifested satisfactory reliability. The coefficient of correlation ranged between r= 0.72 and r= 0.97. Also, this confirmed the results from Student t- test, whereas no significant difference was found between both measurements in almost tests, except on the test for measuring flexibility - ***Sit and reach***. There is statistically significant difference between both measurements at all ages. Taking into the consideration that the coefficients of correlation are very high, it can be assumed that some uncontrolled factor was influencing the results. Eventually, this may be explained with the fact that maybe the students were not adequately prepared for performing this test of flexibility, so with every new performance, there was a minimal, but statistically significant difference between repeated measurements.

In addition to the reliability and sensitivity of the tests also go the obtained data, related to the coefficient of variation (CV). When the coefficient of variation is calculated from the mean values and standard deviation of repeated measurements “within” subjects, then it can be used to assess the reliability of individual performance or as the basis to evaluate at least some significant changes caused by the “intervention” (training, fatigue, and learning). Considering that the CV is a measure of the reliability and sensitivity of the test expressed as a percentage, it can be noted that as the CV gets smaller, it indicates a good reliability and sensitivity of the test. The values from 5 to 10% are the limit at which the CV is considered acceptable (Curell & Jeukendrup, 2008). If we accept this cut-off value, it can be argued that the sample of all age groups during the autumn testing, in almost all motor tests, have acceptable coefficients of variation ranging from 2.8% in the test ***Shuttle run 4x10m*** to 14.6% in the test ***Sit and reach***. The only motor test that has an acceptable CV is the test ***Bent-arm hang***, in which the CV reaches values of up to 51.3% (at students aged 9.5 years).

Table 2 shows the girls’ results from test-retest method within the autumn testing. The most frequently applied tests of motor abilities performed in the sample of girls within the autumn testing manifested satisfactory reliability. The coefficient of correlation ranged between r=0.70 and r=0.96. Also, this confirmed the results from Student t- test, whereas no significant difference was found between both measurements in almost tests, except on the test ***Sit and reach,*** in which statistically significant differences appear in repeated measures of groups aged 9.5 and 11.5 years. The obtained data suggest that the sample of girls of all age groups during the autumn testing in almost all motor test have acceptable CV ranging from 2.9%, in the test ***Sit-ups for 30s*** to 13.1% in the test ***20m - Shuttle run***. The motor test that does not have an acceptable CV is the test ***Sit and reach***, but only in the group of girls aged 13.6 years. Since the CV for this test did not show up in all age groups, but only in certain age groups, in the general conclusion we did not take this CV as an acceptable. The only motor test that has an acceptable coefficient of variation for all age groups is the test ***Bent arm hang***, with values reaching slightly lower percentages (CV = 44.6%) than boys, but still very high.

Table 3 shows the boys results from test-retest method within the spring testing. The most frequently applied tests of motor abilities performed in all age groups within the spring testing manifested satisfactory reliability. The coefficient of correlation ranged between r= 0.76 and r= 0.97. Also, this confirmed the results from Student t- test, whereas no significant difference was found between both measurements in almost tests, except for the test ***Sit and reach*** in which in which statistically significant differences appear in repeated measures of groups aged 9.5 and 11.5 years. In addition, statistically significant difference was also found between both measurements in the test ***1600 m***, but only in the group of students aged 9.5, while for students aged 13.6 years, statistically significant difference was found for the two motor tests: ***Sit-ups for 30s*** and ***20m - Shuttle run.*** Statistically significant differences obtained in the test 1600 m possibly can be explained by different weather conditions when this group of students were tested in spring. At the second measurement, the unplanned high temperatures for that time of year had a negative effect on the performance of the test. The obtained data suggest that the sample of boys of all age groups during the spring testing in almost all motor test have acceptable CV ranging from 2.6%, in the test ***Shuttle-run 4x10m*** to 13.3% in the test ***20m - Shuttle run***. The motor test that does not have an acceptable CV in the group of girls aged 9.5 years is test ***20m - Shuttle run***. Since the CV for this test did not show up in all age groups, in the general conclusion we did not take this CV as an acceptable. The only motor test that has an acceptable coefficient of variation for all age groups is the test ***Bent arm hang***, in which the CV reaches values of up to 39.4 % (at students aged 13.6 years).

Table 4 shows the girls results from test-retest method within the spring testing. The most frequently applied tests of motor abilities performed in all age groups of girls within the spring testing manifested satisfactory reliability. The coefficient of correlation ranged between r= 0.73 and r=0.98. Also, this confirmed the results from Student t- test, whereas no significant difference was found between both measurements in almost tests, except for the test ***Sit and reach*** in which it appears statistically significant differences in repeated measures in all age groups. In addition, statistically significant difference was also found between both measurements in the test ***1600 m***, but only in the group of girls aged 9.5. As we mentioned, this can be explained by the different weather conditions between two measurements of this test for this age group of girls within the spring testing. The obtained data suggest that the sample of girls of all age groups during the spring testing in almost all motor tests have acceptable CV, ranging from 2.4%, in the test ***Shuttle run 4x10m*** to 13.6% in the test ***20m - Shuttle run***. The only motor test that has an acceptable coefficient of variation for all age groups is the test ***Bent arm hang***, with values reaching slightly lower percentages (CV = 33.6%) than boys, but still being very high.

By using univariate analysis of variance with repeated measures (2x3x2), we analyzed if the tests are applicable to all age periods of both gender in this sample of students. The univariate analysis of variance with repeated measures was applied to the results obtained in the spring testing. The factors were "Gender" (boys and girls) x "Age" (9.5 years. 11.5 years and 13.6 years) x "Measurement" (First and second measurement). In the test ***Sit-ups for 30 s***. the significant main effects of "Gender" ($F_{1.834}$=241.9; p<.01) and of "Age" ($F_{2.834}$=129.8; p<.01) were found. while the effect of factor "Measurement" was not statistically significant ($F_{1.834}$=1.8; p>.01). When it comes to the interaction of these factors, the only significant interaction was found at factors "Gender" x "Age" ($F_{2.834}$=6.2; p<.01). In the test ***Standing long jump*** the significant main effects of "Gender" ($F_{1.834}$=417.2; p<.01) and of "Age" ($F_{2.834}$=482.1; p<.01) were found, while the effect of factor "Measurement" was not statistically significant ($F_{1.834}$=1.0; p>.01). When it comes to the interaction of these factors, the only significant interaction was found at factors "Gender" x "Age" ($F_{2.834}$=48.5; p<.01). In the test ***Sit and reach*** the significant main effects of "Gender" ($F_{1.834}$=330.8; p<.01) and of "Age" ($F_{2.834}$=22.7; p<.01) were found, while the effect of factor "Measurement" was not statistically significant ($F_{1.834}$=3.3; p>.01). When it comes to the interaction of these factors, the only significant interaction was found at factors "Gender" x "Age" ($F_{2.834}$=10.7; p<.01). In the test ***Bent arm hang*** the significant main effects of "Gender" ($F_{1.834}$=57.7; p<.01) and of "Age" ($F_{2.834}$=166.9; p<.01) were found, while the effect of factor "Measurement" was not statistically significant ($F_{1.834}$=0.0; p>.01). When it comes to the interaction of these factors, the only significant interaction was found at factors "Gender" x "Age" ($F_{2.834}$=25.8; p<.01). In the test ***20 m-Shuttle run*** the significant main effects of "Gender" ($F_{1.834}$=241.8; p<.01) and of "Age" ($F_{2.834}$=152.8; p<.01) were found, while the effect of factor "Measurement" was not statistically significant ($F_{1.834}$=0.6; p>.01). When it comes to the interaction of these factors, the only significant interaction was found at factors "Gender" x "Age" ($F_{2.834}$=38.4; p<.01). In the test ***Shuttle run 4x10m*** the significant main effects of "Gender" ($F_{1.834}$=280.5; p<.01) and of "Age" ($F_{2.834}$=424.4; p<.01) were found, while the effect of factor "Measurement" was not statistically significant ($F_{1.834}$=0.9; p>.01). When it comes to the interaction of these factors, the only significant interaction was found at factors "Gender" x "Age" ($F_{2.834}$=8.1; p<.01). In the test ***1600m*** the significant main effects of "Gender" ($F_{1.834}$=211.3; p<.01) and of "Age" ($F_{2.834}$=66.2; p<.01) were found, while the effect of factor "Measurement" was not statistically significant ($F_{1.834}$=0.9; p>.01). When it comes to the interaction of these factors, the only significant interaction was found at factors "Gender" x "Age" ($F_{2.834}$=1.6; p<.01). The obtained results indicate that there is no effect of the factor "Age" and there is no effect of the factor "Gender" on the all of the observed motor variables. The lack of interaction between the factor "Measurement" and the other two factors indicates that the tests are applicable to all age periods of both genders.

CONCLUSIONS

Based on the empirical verification of reliability and sensitivity of the most frequently applied tests in different models of physical fitness monitoring of children and youth worldwide, it can be concluded that these tests manifest satisfactory reliability and sensitiveness at all ages, in both genders from 3rd to 8th grade of elementary school. The obtained results indicate that there is no effect of the factor "Age" and there is no effect of the factor "Gender" on all of the observed motor variables. The lack of interaction between the factor "Measurement" and the other two factors indicates that the tests are applicable to all age periods of both gender. Based on the evaluation of the most frequently applied tests in different models of physical fitness monitoring of children and youth worldwide, battery of tests, which fits the model of monitoring physical fitness of students in terms of implementation of physical education in Serbia, is proposed.

This paper is the result of the project "Improving the quality and accessibility of education in modernization processes in Serbia" No 47008 (2011-2014) and also it is the result of the project "The effects of physical activity application to locomotor, metabolic, psychosocial and educational status with population of the Republic of Serbia" No III47015 (2011-2014) financially supported by the Ministry of Education and Science of the Republic of Serbia.

REFERENCES

Carlson, T.B. (1995). We hate gym: Student alientation from physical education, *Journal of Teaching in Physical Education*, 14, 467-477.

Corbin, C. B., Pangrazi, R.P., & Welk, G.J. (1995). A response to "Horse is Dead; Let's Dismount".*Pediatric Exercise Science,* **7**, 347-351.

Curell, K., Jeukendrup, A.E. (2008). Validity, reliability and sensitivity of measures of sporting performance. *Sports. Med.*38 (4), 297-316.

Ennis, C.D. (1996). Student's experiences in sport-based physical education: [More than] apologies are necessary, *Quest.* 48. 453-456.

Hardman, K. (2008). Physical education in schools: A global perspective. *Kinesiology*. 40. 5-28.

Portman, P.A. (1995). Who is having fun in physical education classes? Experiences of six-grade students in elementary and middle schools, *Journal of Teaching in Physical Education.* 14, 445-453.

Robinson, D.W. (1990). An attributional analysis of students of student demoralization in physical education settings, *Quest,* 42, 27-39.

Ortega, F.,B., Ruiz, J.R., Castillo, M.J., Sjostrom, M. (2008). Physical fitness in childhood and adolescence: a powerful marker of health, *International Journal of Obesity,* 32. 1-11.

Table 1 A correlation between two consecutive measurements of motor abilities of all three age groups of boys within the autumn testing as a measure of reliability and sensitivity of the applied motor tests. Displayed are the mean and standard deviation (M ± SD) of given size (1st and 2nd measurement), correlation coefficients (r), the levels of significance (p) obtained by t-test and coefficients of variation (CV).**p<0.05**

Variables	3-4 grade / M ± SD (9.5 ±0.5years)					5-6 grade /M ± SD (11.5±0.6years)					7-8 grade /M ± SD (13.6±0.5years)				
	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%
SU (n)	21.39±4.21	21.73±4.51	.74	.22	11.9	22.58±3.78	22.81±3.74	.85*	.20	6..9	25.69±4.37	26.56±4.28	.84*	**.00**	7..3
SLJ(cm)	129.38±20.12	127.95±20.97	.88*	.10	6.4	148.66±16.50	148.56±16.95	.86*	.90	4.4	177.55±24.00	177.17±24.48	.94*	.59	3.4
SR (cm)	14.49±4.45	15.16±4.28	.92*	**.00**	10.3	13.00±5.00	13.62±5.07	.95*	**.00**	13.6	13.70±6.26	14.42±6.57	.97*	**.00**	14.6
BAH (s)	19.07±14.29	17.62±15.17	.80*	.07	51.3	22.48±16.92	21.56±17.74	.88*	.22	36.5	38.60±24.64	39.16±24.83	.86*	.60	30.2
SHR(s)	197.02±74.06	200.57±73.94	.93*	.22	13	253.2±102.53	247.57±98.34	.95*	.17	11.5	313.97±110.9	300.71±116.34	.72	.32	30.7
4x10m (s)	13.62±1.04	13.68±1.04	.72	.40	4	12.51±0.79	12.59±0.69	.80*	.06	2.8	11.91±0.88	12.00±0.81	.85*	.05	9.6
1600 m(s)	611.06±83.35	612.12±77.11	.92*	.80	3.8	564.52±77.06	570.22±76.50	.75	.36	6.7	502.13±72.92	508.51±80.36	.94*	.25	3.8
p<0.05															

Table 2 A correlation between two consecutive measurements of motor abilities of all three age groups of girls within the autumn testing as a measure of reliability and sensitivity of the applied motor tests. Displayed are the mean and standard deviation (M ± SD) of given size (1st and 2nd measurement), correlation coefficients (r), the levels of significance (p) obtained by t-test and coefficients of variation (CV).

Variables	3-4 grade/M ± SD (9.5 ±0.5years)					5-6 grade/M ± SD (11.5±0.6years)					7-8 grade/M ± SD (13.6±0.6years)				
	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%
SU (n)	18.78±4.43	18.30±4.62	.80*	.08	2.9	21.16±4.10	20.69±4.35	.77	.06	10.8	21.93±3.13	22.40±3.35	.83*	**.01**	6.5
SLJ(cm)	115.14±17.67	116.02±18.20	.84*	.32	6.5	135.98±18.96	135.65±19.13	.85*	.73	6.0	144.96±16.35	145.15±16.68	.85*	.84	4.6
SR (cm)	16.56±4.94	17.38±4.95	.94*	**.00**	9.5	16.73±4.96	17.46±4.76	.96*	**.00**	7.3	19.97±6.27	20.60±6.51	.87*	.06	20.2
BAH (s)	14.94±12.44	15.41±13.06	.86*	.51	44.6	14.27±10.07	14.40±10.96	.84*	.82	37	18.49±12.45	19.28±14.16	.81*	.57	35.9
SHR(s)	148.91±35.22	153.30±39.16	.85*	.08	10.6	189.88±70.74	192.92±70.85	.91*	.42	13.1	205.43±50.37	202.96±46.24	.81*	.25	11.8
4x10m (s)	14.52±1.09	14.38±1.03	.70	.06	4.1	13.22±0.97	13.29±0.87	.77	.17	3.4	13.15±0.86	13.09±0.80	.77	.28	3
1600 m(s)	675.95±78.50	681.72±76.58	.84*	.33	4.7	639.41±87.57	645.46±94.73	.94*	.12	3.6	610.76±78.74	605.54±79.57	.73	.70	6.1
p<0.05															

Table 3 A correlation between two consecutive measurements of motor abilities of all three age groups of boys within the spring testing as a measure of reliability and sensitivity of the applied motor tests. Displayed are the mean and standard deviation (M ± SD) of given size (1st and 2nd measurement), correlation coefficients (r), the levels of significance (p) obtained by t-test and coefficients of variation (CV).

Variables	3-4 grade/ M ± SD (9.5 ±0.5years)					5-6 grade/M ± SD (11.5±0.6years)					7-8 grade/M ± SD (13.6±0.5years)				
	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%
SU (n)	23.55±4.08	23.15±4.20	.81*	.06	8.8	24.79±3.97	25.09±4.36	.76	.25	9.7	27.57±4.16	28.35±4.06	.88*	**.00**	5.4
SLJ(cm)	131.90±20.99	131.85±21.66	.93*	.95	4.9	159.78±20.14	160.79±20.33	.94*	.12	3.3	185.84±25.40	186.91±25.21	.97*	.05	2.6
SR (cm)	14.62±4.48	15.15±4.32	.96*	**.00**	7.2	13.04±4.55	13.67±4.77	.97*	**.00**	8.6	15.32±6.63	15.61±6.50	.96*	.06	10.9
BAH (s)	13.61±9.98	14.07±10.44	.88*	.21	28.9	21.62±13.73	22.24±14.72	.88*	.32	31.7	39.95±25.19	38.53±23.45	.89*	.15	39.4
SHR(s)	208.23±76.90	212.83±79.54	.87*	.084	16.7	279.04±101.0	276.12±104.10	.91*	.51	13.3	358.39±116.88	367.97±116.52	.97*	**.00**	6.2
4x10m (s)	13.25±0.91	13.34±0.93	.83*	.06	3.0	12.21±0.78	12.15±0.85	.84*	.16	2.7	11.52±0.94	11.45±0.89	.89*	.06	2.6
1600 m(s)	610.00±95.10	592.25±88.58	.80*	**.00**	6.6	570.10±82.93	570.66±84.97	.90*	.87	4.8	500.86±78.71	501.49±74.75	.92*	.94	4.0

Table 4 A correlation between two consecutive measurements of motor abilities of all three age groups of girls within the spring testing as a measure of reliability and sensitivity of the applied motor tests. Displayed are the mean and standard deviation (M ± SD) of given size (1st and 2nd measurement), correlation coefficients (r), the levels of significance (p) obtained by t-test and coefficients of variation (CV).

Variables	3-4 grade/M ± SD (9.5 ±0.5years)					5-6 grade/M ± SD (11.5±0.6years)					7-8 grade/M ± SD (13.6±0.6years)				
	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%	1. measurment (M ± SD)	2. measurment (M ± SD)	r	p	CV%
SU (n)	20.36±4.40	20.31±4.14	.84*	.82	9.8	22.44±3.36	22.66±3.74	.73	.33	8.9	23.60±4.06	24.01±4.00	.80*	.10	7.3
SLJ(cm)	120.26±19.53	120.74±19.57	.91*	.51	5.2	140.97±18.22	142.14±18.08	.93*	.05	3.6	148.97±20.45	148.99±20.67	.96*	.97	3.2
SR (cm)	17.91±5.28	18.47±5.06	.95*	**.00**	8.5	19.14±4.97	19.67±4.98	.97*	**.00**	5.7	21.23±7.12	21.75±6.96	.98*	**.00**	6.2
BAH (s)	11.96±7.58	12.71±7.58	.85*	.07	29.4	19.47±14.84	19.07±13.96	.91*	.30	33.3	23.92±15.84	24.88±16.93	.91*	.17	26.5
SHR(s)	174.87±57.79	184.80±61.59	.87*	**.00**	13.6	214.74±74.02	215.57±75.92	.93*	.79	10.0	231.09±63.09	232.05±62.67	.88*	.18	10.6
4x10m (s)	13.96± 0.94	13.95±0.90	.76	.92	3.3	12.86±0.88	12.78±0.88	.79*	.18	3.2	12.62±0.85	12.47±0.78	.87*	**.00**	2.4
1600 m(s)	666.48±74.97	646.80±78.52	.76	**.00**	6.0	640.58±86.50	638.33±90.12	.87*	.48	5.4	606.04±70.75	602.88±70.48	.83*	.17	5.0