Pediatric Exercise Science, 2011, 23, 151-160

Reliability of Squat and Countermovement Jump Tests in Children 6 to 8 Years of Age

Rafael Martín Acero, Miguel Fernández-del Olmo, and José Andrés Sánchez
University of A Coruña

Xosé Luis Otero
University of Santiago of Compostela

Xavier Aguado
University of Castilla-La Mancha

Ferrán A. Rodríguez
University of Barcelona

The aim of this study was to determine the reliability of the squat jump test (SJ) and countermovement jump test (CMJ), in fifty-six children (30 girls and 26 boys) with ages ranging from 6 to 8 years. Each subject performed two evaluation sessions (T1, T2) with seven days between tests. The results show that the CMJ test has a high intratrial reproducibility in T1 and T2 measured through intraclass correlation coefficient (ICC ≥ 0.95). The ICC for the SJ test had a high value (0.99) only in T1. The variability for both tests among children under 9 years of age is higher than those reported for adult subjects in other studies. The intersession reliability was questionable with a high methodical error (ME= 9.86–15.1%, for the SJ and CMJ, respectively) and a significant worsening of the results of CMJ in T2 ($p < .05$).

Since the 1980s, the vertical jumps battery, this is a set of vertical jump tests, has been applied to many populations of different ages, both for athletes and-non athletes, to evaluate the explosive strength of the lower limbs (7). Squat jumps (SJ) and counter-movement jumps tests (CMJ) have been applied to young children to study neuromuscular capabilities of children as well as their development process. However, the reliability of using these test with young children, especially under 10 years of age, has not been extensively studied or reported.

Acero, Fernandez-del Olmo, and Sánchez are with the Faculty of Sport Science and Physical Education (INEF Galicia), University of A Coruña, Coruña, Spain. Otero is with the Dept. of Biostatistics and Research Methods, University of Santiago of Compostela, Compostela, Spain. Aguado is with the Faculty of Sport Science, University of Castilla-La Mancha, Spain. Rodríguez is with INEF of Catalunya, University of Barcelona, Barcelona, Spain.

Squat jumps and CMJs have reported very high test-retest reliability indices in adults (8,9,18) and have shown the best reproducibility for estimating the muscle power in physically active adults (12). However, it has been found that the reliability of these jumps depends on the age or skill of the group evaluated. The variability among repetitions of the CMJ was greater for a group of 10-year-old children than for 15-years of age (17). Papadoupoulos et al. (14) studied the height of the vertical jump in a group of boy and girl athletes from 10 to 15 years, and found an intraclass correlation coefficient (ICC) of 0.97. Another SJ study reported an ICC of 0.93 in a group of 13 year old swimmers (13). None of these studies provided necessary methodological measures for the control of extraneous variables caused by the observer and the exact instructions regarding how the jumps must be performed. This is crucial to compare data from different studies. Moreover, to date there are no test-retest reliability measures of vertical jumps in children less than 10 years of age (15). This is a crucial point in training and growth studies in which changes should be attributable to either the intervention or maturational/growth phenomenon rather than measurement error.

The aim of this study was to measure the reliability of the SJ and CMJ tests in students ranging in age from 6 to 8 years. To this end, reliability was assessed by examining: (a) the reproducibility among repetitions on each of the two assessment sessions (variation among repetitions in two sessions), and (b) the temporal reproducibility between two sessions.

Methods

Subjects

Fifty-six subjects (30 girls and 26 boys) between 6 and 8 years old participated in this study (mean ± *SD*: 7.30 ± 0.71 years). The children were students from a school in Galicia. Before testing, informed consent was obtained from the childrens' parents or guardians. The study was approved by the local Ethical Committee of University of A Coruña.

Material

Contact platform (5) ErgoJump Bosco System was used to record the heigh of the jumps. This device was a conductor carpet (dimensions L-175 x W-70 cm) connected to an electronic timing system. The timer switched on automatically when a subject takes off and switches off at the time when a foot makes contact with the plate again.

Microprocessors (5) Psion Organizer II © (Datapak 32 k) were used to record the data collected from the platform through an external connection and ErgoJump Bosco System software ©, v.05.

Jump Tests

The vertical jump tests were performed according the protocols described in Bosco (7).

Procedure and Control of Extraneous Variables

The participant performed two testing sessions (T1, T2) with seven days between tests. The seven days interval was chosen for two reasons; i) to avoid any fatigue induced for the session and ii) to make coincide the testing sessions with the school class of physical education. Thus, the two testing sessions were completed at approximately the same time of day and under similar environmental conditions. In each session all subjects performed four trials of two types of jumps (SJ and CMJ) with a 1-min-standardized interval between each jump to allow the tester to repeat the instructions for the test performance. The subjects did not have previous experiences with the jump tests and they not participated in familiarization sessions or training prior collecting the data. On the second session the testers asked verbally to the children whether they fell tired or bored. Participants were excluded from the second testing session if they answered positively any of the questions.

In the period between sessions the subjects did not perform unusual physical activity.

A control guide about all sources of variation of results was applied (1).

Control of Extraneous Variables Related to the Tester. Testers were chosen from among graduates in Physical Activity and Sport after a procedure of theoretical and practice formation in the jump test evaluation. The procedure of formation and training of the testers was systematic. The procedure included theoretical sessions about the jumps technical requirements and video sessions that displayed jumps performed correctly and incorrectly. To participate in the study as a tester, the students should reach a high intra- and interobserver reliability according with the procedure described in Anguera (2). The 5 testers with the high intra and interobserve reliability were selected for the study.

At the time of the recording sessions of this experiment the testing team had more than three years of experience. The method of systematic observation by agreement and consensus was used (2). The tester selected each jump trial according to the following categories: “correct” or “incorrect” depending on whether the jump was used for analysis. If the trial jump was “correct” the jump height was recorded; if the trial jump was “incorrect” the data were not used for analysis. A jump was considered “correct” when the criteria established in Table 1 were met.

Control of Extraneous Variables Related to the Experimental Design. According to Bosco et al. (9), with three or more standard jumps performed by adults a high degree of test-retest correlation (r = .95) is achieved. In our study the subjects performed four trials (follow a personal communication of Carmelo Bosco).

Bobbert et al. (6) found that more height is achieved through the CMJ than through the SJ, whether the angle of the knee in the starting position was controlled or freely chosen.

Bosco and Viitasalo (8) confirmed that the contribution of elastic energy in the phase of positive work during muscular contraction is related with the angle of the knee flexion during the eccentric phase. They studied different angles of the knees (low, 124.7–128.7 degrees; and large, 92.7–90.8 degrees) and they found that subjects with the lowest percentage of flight time (37.4 ± 8.4) jumped almost equally with a low angle (approx. 120 degrees) than with a large angle (approx. 90 degrees). However, the subjects with the highest percentage of flight time (56.1± 90

Table 1 Criteria Established for Considering as "Correct" a Vertical Jump

	SJ	CMJ
Starting Position	soles on the platform	
	feet parallel at a distance equal to the width of shoulders. good balance	
	upright position, trunk remain as vertical as possible	
	hands kept on the hips throughout the test	
	knee angle around 90–120 degrees	knee angle straight (180 degrees)
	position held 3–4 s	position held 2–3 s
Starting Action	no counter-movement: trial not considered valid if some movement is perceived that may increase the flexion of knees when starting the jump	knee angle around 90–120 degrees at the end the counter movement: trial not considered valid if the knees do not bend to maximum intensity and quickly
		without a break, the course of movement is inverted through the extension of the knees
	vertical jump started by means of an explosive extension of legs: maximum impulse	
	trunk remain as vertical as possible	
Flight Phase	knees and ankles entirely extended	
Landing Phase	knee angle during the ground contact around 180 degrees: stretched	
	feet in extension: stretched	
	bounce on the tip of toes to encourage knees and feet were stretched	

degrees) jumped higher with little bending, and less with much bending. In order not to inhibit the capability of jumping, participants were allowed to freely choose the angle of knee flexion between 90–120 degrees (4).

Statistical Analysis

Intraindividual variability among repetitions (reproducibility of repeated measurements) in both tests sessions (T1 and T2) was determined by means of the calculation of standard deviation (*SD*), the coefficient of variation (CV), the of analysis of variance with repeated measures (ANOVA) and the ICC.

The CV provides information about changes in individual performance while the ANOVA reports changes in group performance. The ICC indicates the ranking of the participants both within the testing session and between testing sessions. The most common method of ICC calculation is based on a model of analysis of variance with repeated measures. The idea is that the total variability of the measurements can be decomposed into two components: variability due to differences between subjects and due to differences between measurements for each subject. The latter, in turn, depends on the variability between observations and random residual variability associated with any measurement error involved. The ICC is then defined as the proportion of the total variability is due to the variability of the subjects. As a proportion, CCI values range from 0 to 1, so that the maximum possible match corresponds to a value of CCI = 1. In this case, all the observed variability could be explained by differences between subjects. The final interpretation would be that high CCI means the subjects will mostly keep their same places in the ranking between tests and thus, It is useful as an indicator of the reliability of a single measure.

Temporal variability (temporal reproducibility) between both sessions (variation between T1 vs T2) was calculated by means of the Pearson's correlation coefficient (r), the ICC, and the methodical error (ME). The ME of repeated measurements can be applied to two tests conducted on different days. The ME can be expressed as a coefficient of variation (11). The significance of the differences between mean values in each session was calculated using Student's *t* test for related samples. The level of significance chosen for the statistical analysis was *p* £ .05. All data were analyzed using SPSS for Windows (version 14.0; SPSS Inc, Chicago, IL).

Results

Descriptive statistics of the results obtained in the two sessions of assessment (T1, T2) for each of the jumps performed are presented in Table 2. A previous gender differential analysis (not shown) revealed no significant differences in the jumps evaluated and thus, all the data were analyzed as one group.

Table 2 Descriptive Statistics of the Measurements of Four Trials for the SJ and CMJ Tests (*n* = 56) in Two Sessions (T1 and T2)

	T 1				T 2			
	mean	*sd*	*max*	*min*	*mean*	*sd*	*max.*	*min*
SJ (cm)	13.00	2.82	17.85	7.90	12.96	2.93	20.85	7.60
CMJ (cm)	16.06	3.56	25.00	7.80	14.10	2.80	19.53	7.70

Intratrial Reproducibility

The reproducibility in four repetitions of the jump tests in each of the two sessions (T1, T2) is presented in Table 3. The ANOVA did not show any significant effect in the trials in both sessions. Thus, the mean values for the Jump tests values, within both sessions, showed no significant differences among trials.

From a descriptive point of view, the CV values were slightly higher in T1 in comparison with T2 (11.02% vs 9.19% and 8.72% vs 8.48% for the SJ and CMJ, respectively). Regarding the ICC, the values were lower in T1 than in T2 (0.83% vs 0.99% and 0.95% vs 0.99% for the SJ and CMJ, respectively).

Table 3 Reproducibility of Vertical Jumps (SJ and CMJ) Expressed as Variability Among Four Trials in Each of the Two Test Sessions (T1 and T2): Significance (p) of Repeated-Measures ANOVA Test, Intraclass Correlation Coefficient (ICC), and Coefficient of Variation (CV)

	T 1			T 2		
	p	ICC	CV (%)	*p*	ICC	CV (%)
SJ (cm)	0.80	0.83[a]	11.02	0.57	0.99[c]	9.19
CMJ (cm)	0.46	0.95[b]	8.72	0.88	0.99[d]	8.48

95% Confidence Interval: [a] 0.70–0.90; [b] 0.88–0.97; [c] 0.96–0.99; [d] 0.97–0.99

Temporal Reproducibility

Table 4 shows the results of temporal reproducibility, expressed as the variability between T1 and T2. Significant differences ($p \leq .001$) were found in the CMJ height, the values in T2 were significant smaller than in T1. There were no significant changes in the SJ height between T1 and T2. The ICC values were 0.70 and 0.86, and the r-values 0.60 and 0.78 for the SJ and CMJ, respectively. Regarding the EM, the values The EM values were 15.07% and 9.86% for SJ and CMJ, respectively.

Table 4 Temporal Reproducibility of the Vertical Jumps (SJ and CMJ) Expressed as Variability Day by Day Between Two Sessions (T1 and T2): Significance (p) of the Student's t for Related Samples, Intraclass Correlation Coefficient (ICC), Pearson's Correlation Coefficient (r), and the Methodical Error (ME %)

	T1 *vs* T2			
	p	ICC	*r*	ME (%)
SJ (cm)	0.80	0.70[a]	0.60	15.07
CMJ (cm)	≤0.001	0.86[b]	0.78	9.86

[a]95% Confidence Interval = 0.55–0.79

[b]95% Confidence Interval = 0.80–0.93

Discussion

According to our knowledge, the current study is the first to assess the reliability of CMJ and SJ tests in boys and girls between 6 and 8 years of age. This is a relevant issue for study designs that include prepost test or for longitudinal studies, to control which changes should be attributable to either the intervention or maturational/growth phenomenon rather than measurement error.

Ours results demonstrate that both tests have acceptable reliability although both sessions of evaluation are recommended for the SJ test. However, the intersession reliability is questionable.

Intratrial Reproducibility

The differences were not significant among the four trials of every jump, in both sessions, suggesting an absence of systematic source of variability such as the produced by a learning process. The ICC values in both sessions indicate a high stability of ranking of the participants for CMJ and SJ. CMJ showed a high stability in T1 (ICC = 0.95); similar ICC values were reported by Vincent (19) and Atkinson and Nevill (3). However, SJ showed the smallest ICC which may be because, when the half flexion of knees is attained, no countermovement is possible. In this body position, the motor control is relatively complicated, even for adult subjects, because the muscular contraction is enhanced through the tension created by the prestretch or eccentric loading of the muscles. In T2, SJ and CMJ showed better and identical ICC values (0.99) than in T1.

The CV reports the variability in the individual performance of a test in all subjects; in this regard, all the jump tests showed high variability in both T1 and T2. However, the absence of significant differences reported by the ANOVA and the high ICC values suggest that the individual variability by itself is not a good indicator of the reproducibility of a test, at least, at this ages. The high ICC values reported in this study indicate that, in children between 6–8 years old, the intratrial variability is mainly explained by differences between subjects (overall for the CMJ test).

The reproducibility values of SJ and CMJ in the schoolchildren analyzed (6–8 years) are comparable to those in adult subjects, students of physical education (16), and schoolchildren 10–15 years (17), although the intersubjects variability is higher than in the adults subjects (Table 5). According to Viitasalo (17), the CMJ is

Table 5 Intratrial Reproducibility of CMJ Test in a Same Session

Age (Years)	Sample	CV (%)	Author, Year
18–24	Physical Education Students (higher education)	4.30	Viitasalo, 1985
15	Gymnastics, Free Wrestling	5.05	Viitasalo, 1988
14	Skiing, Foot orienteering, Athletics, Wrestling	7.05	Viitasalo, 1988
13	Skiing, Foot orienteering, Hockey, Gymnastics, Athletics, Free Wrestling	7.06	Viitasalo, 1988
12	Skiing, Foot orienteering, Gymnastics, Basketball, Athletics, Free Wrestling	7.80	Viitasalo, 1988
11	Skiing, Basketball, Gymnastics, Athletics, Free Wrestling	8.62	Viitasalo, 1988
10	Skiing, Foot orienteering, Athletics, Free Wrestling	13.65	Viitasalo, 1988
6–7-8	Schoolchildren	8.72	Current study: T1
6–7-8	Schoolchildren	8.48	Current study: T2

Coefficient of variation (CV) in different ages and samples is shown.

reliable test from the 11–12 age years, and sportsmen above 15 years (e.g., players of volleyball).

The results of the current study suggest that CMJ test is a test with good reproducibility in children of 6–8 years for a session test. However, for the SJ should be include a familiarization session to achieve good reliability.

Temporal Reproducibility

Significant between-session differences in average height of CMJ suggest some form of systematic variation. Theoretically, a positive trend (improvement between T1 and T2) could be attributed to a learning effect (3), but, in fact the results decreased in the second session (see Tables 1 and 3). Such a finding is difficult to explain, and given that the sources of systematic error related with the equipment and the observers are considered improbable, the decrease could be due to some form of biologic variability (intrasubject) that could include: lack of motivation; lower psychological activation; or loss of concentration in the test performance. However, we cannot discard a learning effect since it was an improvement of the consistency of the measures was observed (Table 3). One plausible explanation is that the children to execute more correctly the jumps did not perform with the maximal effort. On the other hand, a greater variability among the four repetitions in both tests (Table 3) could explain the nonsignificant differences between sessions in the SJ test.

According to the Vincent categorization (19), the temporal reproducibility of CMJ would be considered acceptable (ICC = 0.80–0.90), whereas that of the SJ would be considered questionable (ICC = 0.70–0.80). The lower intratrial reliability of the SJ in comparison with the CMJ across the 8 trials performed in the two sessions could explain the absence of significant differences between sessions in the SJ test. Nevertheless, the high ME in both jumps and the significant differences between T1 and T2 in the CMJ test, suggests that it is not reproducible. The temporal reproducibility of the jump test was lower in the schoolchildren from 6 to 8 years (r = .60–0.78) than in adult participants, specifically, students of physical education (r = .84–0.99; 16). However, only the Pearson's correlation coefficient can be considered a relative indicator of the reliability, being strongly influenced by the range of extreme values and the heterogeneity of the sample. Furthermore, the r value does not consider the number of repetitions (11). Thus, the ICC has been considered more appropriate to assess the reliability (Baumgartner, 5).

In general terms, the temporal reproducibility of the test separated by seven days can be considered questionable, because a high methodological error and a significant worsening of the results of CMJ in the second session were observed. Accordingly, we must not only question the learning curve of these jumps, but also the necessary motivation and concentration needed in boys and girls under 9 years to achieve a maximum jump power, that is consistent and reproducible in all sessions (10).

Methodological Recommendations

It is necessary to be cautious when comparing this study with others reporting the reliability of SJ and CMJ tests. Our study did not support the use of these two tests between sessions, even when we control another sources of variability such as the testers and the technical criterions for acceptable performance. Although we did not evaluated the reliability of these jumps with not trained testers it would be likely a worsening of the results. In addition, the absence of clear technical criterions for acceptable performance could provide an incorrect value for the recorded height and, if the recording system acquires the flight time, through contact with the platform, the error may be even greater. In this regard, if the subjects land with the soles simultaneously in contrast with the metatarsus, the last part of the body in taking off from the platform, the flight time registered will be greater and the jump height might exceed the real value by about 3–4 cm (7). Such measurement error should not be included in the analysis. It is important to educate the testers because they are the decision makers on whether the jump can be considered valid or not based on their expert judgment. Often these criteria are not included in the methodology of many studies that have used vertical jumps as an assessment tool. This could explain the differences in the height of the vertical jump reported in some studies, despite the similarity in the population samples.

Conclusions

Our results show that the CMJ test has a greater intratrial reliability than the SJ test, although both present an acceptable intratrial reliability in the same session in subjects aged between 6 and 8 years. Thus, it will be enough to perform one trial in the session to obtain a reliable data and to compare for example groups of different ages (e.g., transverse studies). However, this is true for only one session since the reliability between sessions is questionable, which suggests caution is needed in experimental studies that involve some form of pre- and posttest.

It is also recommended that in future studies involving vertical jumps that intrasubject and intersessions reliability should be included as well as the performance criteria that were used and the training status of the testers.

References

1. Acero, R.M., M.F. Del Olmo, X. Aguado, and L. Bergantiños. Fontes de variaçáo na investigaçáo das capacidades de salto e de corrida. *Horizonte Rev. Ed. Fisica e Desporto.* 97: 23-32, 2001.
2. Anguera, M. *T. Manual de Prácticas de Observación*. México: Trillas, 1983.
3. Atkinson, G., and A. Nevill. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.* 26:217–238, 1998.
4. Baker, D., G. Wilson, and B. Carlyon. Generality versus specificity: a comparison of dynamic and isometric measures of strength and speed-strength. *Eur. J. Appl. Physiol. Occup. Physiol.* 68:350–355, 1994.
5. Baumgarter, T.A. Norm-referenced measurement: reliability. In: *Measurement Concepts in Physical Education and Exercise Science,* J. Safrit and M. Wood (Eds.). Champaign, Illinois: Human Kinetics, 1989, pp. 45–72.
6. Bobbert, M.F., K.G. Gerritsen, M.C. Litjens, and A.J. Van Soest. Why is countermovement jump height greater than squat jump height? *Med. Sci. Sports Exerc.* 28:1402–1412, 1996.
7. Bosco, C. *La Valutazione della Forza con in Test di Bosco*. Roma: Societá Stampa Sportiva, 1992.
8. Bosco, C., and J.T. Viitasalo. Potentiation of myoelectrical activity of human muscles in vertical jumps. *Electromyogr. Clin. Neurophysiol.* 22:549–562, 1982.

9. Bosco, C., P. Luhtanen, and P.V. Komi. A simple method for measurement of mechanical power in jumping. *Eur. J. Appl. Physiol. Occup. Physiol.* 50:273–282, 1983.
10. Doré, E., P. Duché, D. Rouffet, S. Ratel, M. Bedu, and E. Van Praagh. Measurement error in short-term power testing in young people. *J Sports Sci.* 21-2:135–142, 2003.
11. MacDougall, J.D., H.A. Wenger, and H.J. Green (Eds.). *Evaluación Fisiológica del Deportista*. Barcelona: Paidotribo, 1995.
12. Marković, G., D. Dizdar, I. Jukic, and M. Cardinale. Reliability and factorial validity of squat and countermovement jump tests. *J. Strength Cond. Res.* 18:551–555, 2004.
13. Papadopoulos, C., and K. Salonikidis. Diagnose und auswertung der motorischen fähigkeiten kraft und schnelligkeit bei jungen schwimmerm. *Leistungssport.* 30:14–18, 2000.
14. Papadopoulos, C., K. Salonikidis, and D. Schmiedtbleicher. Diagnose und auswertung der motorischen fähigkeiten kraft und schnelligkeit bei kindern im alter zwischen 10-15 jahren. *Leistungssport.* 27:32–38, 1997.
15. Van Praagh, E., and E. Doré. Short-term muscle power during growth and maturation. *Sports Med.* 32:701–728, 2002.
16. Viitasalo, J.T. Measurement of force-velocity characteristics for sportmen in field conditions. In: *Biomechanics IX-A. D.A. Winter, R.W. Norman, R.P. Wells, K.C. Hayes, and A.E. Patla*. Champaign, Illinois: Human Kinetics, 1985, pp. 96–101.
17. Viitasalo, J.T. Evaluation of explosive strength for young and adult athletes. *Res. Q. Exerc. Sport.* 59:9–13, 1988.
18. Viitasalo, J.T., and C. Bosco. Electromechanical behaviour of human muscles in vertical jumps. *Eur. J. Appl. Physiol. Occup. Physiol.* 48:253–261, 1982.
19. Vincent, J. *Statistics in Kenesiology*. Champaign, Illinois: Human Kinetics, 1994.