Journal of Physical Activity & Health, 2006, 3(Suppl. 2), S78-S89

Reliability and Validity of Low Back Strength/Muscular Endurance Field Tests in Adolescents

Norman S. Hannibal III, Sharon Ann Plowman, Marilyn A. Looney, and Jason Brandenburg

Background: Strength, muscular endurance, and flexibility are important components of healthy back function. This study determined the reliability and evaluated the validity of selected low back field tests (*FITNESSGRAM*® Trunk Extension [FG-TE] and Box 90° Dynamic Trunk Extension [B-90° DTE]) when compared to laboratory tests (Parallel Roman Chair Dynamic Trunk Extension [PRC-DTE], Parallel Roman Chair Static Trunk Extension [PRC-STE], and Dynamometer Static Back Lift [DSBL]). *Methods:* Forty males age 15.1 ± 1.2 yr and 32 females age 15.5 ± 1.2 yr participated. *Results*: Intraclass test-retest reliability coefficients (one-way ANOVA model for a single measure) ranged from .940 to .996. Validity coefficients determined by Pearson product moment correlation coefficients for males and females, respectively, were as follows: B-90° DTE vs. PRC-DTE = .82, .62 ($p < .05$); B-90° DTE vs. PRC-STE = .55, .38 ($p < .05$); B-90° DTE vs. DSBL = –.29, –.23; FG-TE vs. PRC-DTE = .23, –.11; FG-TE vs. PRC-STE = –.15, .33; and FG-TE vs. DSBL = –.04, –.36. *Conclusions*: B-90° DTE was shown to be a valid field test when compared to PRC-DTE, but only for the males. Further research on the PRC-DTE and PRC-STE items for adolescents is recommended.

Key Words: health-related physical fitness assessment, *FITNESSGRAM*, low back pain

"Balanced, healthy functioning of the musculoskeletal system requires that muscles be able to exert force or torque (measured as strength), resist fatigue (measured as muscular endurance), and move freely through a full range of motion (measured as flexibility). Because of this, strength, endurance and flexibility are viewed as important dimensions of health related fitness"[1] (p. 3). Of utmost concern is healthy back functioning, because a frequent cause of activity limitation among young adults is low back pain (LBP).[2] Children as young as preteens have been diagnosed

Hannibal was a graduate student, Plowman is Prof. Emeritus, and Looney and Brandenburg are faculty in the Dept. of Kinesiology and Physical Education, Northern Illinois University, DeKalb, IL 60115.

with LBP.[3–8] Given the ever growing problem, it is essential to find ways to raise awareness of what constitutes healthy back function. A first step is to determine relatively easy means for assessing back function. Though multiple causes and risk factors have been speculatively linked with LBP, this condition has usually (based on anatomical and physiological logic) been associated with impaired low back lumbar, hamstring, and hip flexor flexibility, as well as strength and muscular endurance (strength/endurance) of the abdominals and trunk extensors.[9,10]

Currently the *FITNESSGRAM®* test battery[1] includes items to assess abdominal strength/endurance (curl-ups), hamstring flexibility (back saver sit-and-reach), and back extensor strength and flexibility (trunk lift). Assessment of back extensor strength/endurance may be the most important because this is the only fitness component shown to predict both first-time and recurring LBP.[1,10] Retrospective studies have consistently revealed significant relationships between trunk extensor strength (static and dynamic)/endurance and LBP.[9,11–14] Prediction of recurrent pain has been linked to static and dynamic trunk extension strength utilizing a maximal voluntary contraction (MVC) and endurance time to fatigue using ~60% MVC.[9,11–14] In longitudinal prospective studies, the only strength/endurance item found to be predictive of first-time LBP was the 240-s static contraction test of trunk extension.[9,11,12,15] No studies have directly linked performance on the *FITNESSGRAM* trunk extension (FG-TE) with the presence or prediction of LBP, nor has the FG-TE been validated against any of the predictive test items in children or adolescents.[16]

One way to document validity of data requires the use of a criterion measure. Unfortunately, there is no universally acknowledged criterion test for measuring trunk extension strength/endurance. Lumbar extension using a MedX dynamometer could be a criterion measure because it allows for isolation of the erector spinae musculature,[17] but the MedX is expensive, requires skilled technicians, is time consuming, and of limited availability. A study by Kearns et al.[18] determined, through integrated electromyographic (iEMG) measurements, that trunk extension performed on a Parallel Roman Chair (PRC), and to a lesser extent during a straight leg dead lift (SLDL), requires similar recruitment of the erector spinae muscles when compared to the MedX.

Based on these findings, the PRC and an activity close to the SLDL were used as quasi-criterion measures (labeled as laboratory tests) in this study. A dynamometer static back lift (DSBL) was substituted for the dynamic SLDL because the SLDL is contraindicated for the age group being tested.[19] Both static and dynamic PRC test items were utilized. The fact that trunk extension on the PRC activates the hamstrings and gluteals to a greater extent than either the MedX or SLDL[18] is acknowledged as a limitation.

A box 90° dynamic trunk extension endurance test has been proposed for consideration as an alternate item in the *FITNESSGRAM* test battery (C.B. Corbin, *FITNESSGRAM* Advisory Council minutes, June 16, 2001) because biomechanically and physiologically it is more similar to the trunk strength/endurance tests that have been linked to LBP than the FG-TE, and it need not involve specialized equipment. Therefore this item was also included as a field test.

The purposes of this research were to determine the reliability of the selected dynamic and static trunk extension measures and the validity of the box 90° dynamic trunk extension field and FG-TE test items for high school age students.

Methods

Participants

Seventy-two (32 female and 40 male) adolescents, 14 to 18 years of age, were recruited from a local fitness/sports center and neighboring high schools. All were nonsmoking, physically active, and free of any low back/lower extremity pain or injury within the previous 6 months. Prior to participation and following approval from the university's institutional review board (IRB), a parent/legal guardian of each participant under 18 years of age gave his/her informed consent. Written assent was also obtained from all minors. Participants 18 years of age gave their own consent to participate. Each participant then completed medical history, activity, and injury questionnaires.

Procedures

Each participant's height (to the nearest .25 in.) and weight (to nearest .2 lb) were measured using standardized procedures[20] on a Health-O-Meter scale with height attachment (Bridgeview, IL) and were later converted to metric units (cm and kg, respectively). Each participant was familiarized with performance of the five strength/endurance test items prior to any data collection. All testing was completed at a fitness/sport center by the same investigator. Participants were asked to wear similar, if not the same, shorts and T-shirt for each data collection trial, to refrain from eating or drinking for 3 hrs, and to avoid strenuous exercise for 24 hrs prior to testing. Compliance was verbally confirmed prior to each trial. Each participant performed all three laboratory (criterion) tests and both field tests for a total of five separate items. Verbal encouragement was given throughout each test.

Parallel Roman Chair Dynamic Trunk Extension (PRC-DTE). The participant began in a neutral alignment (180° between the back and the legs with both arms folded and held across the chest, see Figure 1) as measured with a goniometer. From there the individual flexed forward at the waist until a 90° angle was reached between the legs and upper torso. Finally the participant extended his/her back in a controlled manner until returning to the original neutral alignment. Each extension followed a system of 3-second beeps played on a tape recording. Each participant's shirt was tucked into his/her shorts and an alligator clip was attached to his/her shirt distal to the xiphoid process. A lightweight chain and a paper clip were attached to this clip and were adjusted so that when in neutral alignment the clip was slightly off the floor. A second paper clip was attached to the same chain so that when the participant reached the 90° position, it was in contact with the floor signaling an ending to the movement.

The horizontal distance from the participant's lateral malleolus of his/her left foot to the center of the lower leg pad was measured and recorded to ensure consistent positioning for each trial. The height of the lower leg pad was also recorded for each participant. The test was terminated when the participant could no longer continue at the predetermined pace of 20 contractions per minute or when he or she voluntarily stopped. The total number of full range-of-motion correctly paced repetitions was recorded. Two spotters were used throughout the testing process, one spotting the upper body and head and the other observing the lower body and leg attachment to the PRC, ensuring the safety of each individual. This served as the laboratory (criterion) test for dynamic trunk extension endurance.

Figure 1 — Starting position for the PRC-DTE and testing position for the PRC-STE.

Parallel Roman Chair Static Trunk Extension (PRC-STE). Using the PRC, each individual began in the neutral position in the same manner as for the PRC-DTE (Figure 1). Once the test was started, the 180° contraction was held and timed while the participant attempted to remain in this neutral position. The test was terminated when the paper clip came in contact with the floor for more than two consecutive seconds or a contraction (hyperextension) beyond 180° developed in the participant's back for the same time period. The total time (seconds) of the static contraction was recorded. This item served as the laboratory (criterion) test for static trunk extension endurance.

Dynamometer Static Back Lift (DSBL). Each individual was instructed to stand on a platform with knees fully extended and head and trunk erect. The participant grasped the hand bar using an alternating grip and the hand bar was positioned across the thighs. The participant was instructed to pull the hand bar straight upward using the back muscles and to roll the shoulders backward during the pull, without leaning backward. Each pull lasted approximately 3 seconds.[21] A plumb line was hung from the ceiling directly behind the participant. If his/her back came in contact with this line, indicating that the participant was starting to lean backward, the test was terminated immediately. Two trials were administered with a 1-minute recovery period between each.[21] The higher of the two measurements was recorded. Each participant was encouraged to exhale throughout the entire contraction to avoid the Valsalva maneuver. The dynamometer was calibrated prior to the start of data collection to ensure that each measurement was accurate. All statistical analyses utilized corrected DSBL scores based on the calibration. This item served as the laboratory (criterion) test for static back strength.

Box 90° Dynamic Trunk Extension (B-90°DTE). For this test, a ½-inch foam pad was placed across a 42" × 36" × 36" box. Each participant was asked to lie in the prone position with his/her upper torso hanging off the edge of the box; the lower legs were then strapped to the box via three broad straps, one across the middle of the gluteal region, one just proximal to the popliteal fossa, and one across the ankle. The participant was then asked to flex forward until reaching a 90° angle, defined as when the upper torso was parallel with the side of the box. He/she then returned to the neutral starting position of 180°. One contraction was completed every 3 seconds as determined by a prerecorded set of beeps and the test ended when this cadence could no longer be kept or the participant voluntarily stopped (C.B. Corbin, *FITNESSGRAM* Advisory Council minutes, June 16, 2001). Each full contraction was recorded. This item served as the field test for dynamic trunk extension endurance.

FITNESSGRAM Trunk Extension (FG-TE). Each participant lay in the prone position on a mat with hands under the thighs. The participant then lifted his/her trunk off the floor as high as possible in a very slow and controlled manner, keeping the feet in contact with the floor and the head in a straight alignment with the spine. This contraction was held until the tester was able to measure the distance from the floor to the participant's chin (2 to 5 seconds). Two trials were performed and the highest score was recorded.[22] This item served as the field test of static trunk extension strength.

Once each participant demonstrated the ability to safely perform each test, the first trial (T_1) was performed. The FG-TE, which took approximately 5 s, was always performed first followed by a 5-min rest period. In order to minimize the impact of fatigue on the three muscular endurance tests (PRC-DTE, PRC-STE, and B-90° DTE), these were assigned in random order to each participant. Each endurance test was followed by a 15-min recovery period, during which the participant was asked to perform light activity. The DSBL, which took < 5 s, was always performed last.

Test-Retest Reliability

At least 2 days, but no more than 7 days after T_1, each participant returned for the second data collection session (T_2). The same random order as assigned during T_1 was completed with the same rest periods allotted. Following T_2, the data from each of the first two trials were compared. For quality control in obtaining maximal effort, if the results differed by more than 12%,[23] a third complete testing session (T_3) was performed.

Statistical Analyses

SPSS 10.0 for Windows was used for all statistical analyses. Data were analyzed for outliers via boxplots and scattergrams. Reported results are based on all 72 participants. Subject descriptive characteristics were as follows: age (yrs), F = 15.5 ± 1.19; M = 15.08 ± 1.21; height (cm), F = 164.12 ± 6.52; M = 172.61 ± 7.6; weight (kg), F = 58.53 ± 8.15; M = 66.86 ± 9.76; and body mass index ($kg \cdot m^{-2}$), F = 21.67 ± 2.25; M = 22.45 ± 3.11.

Reliability was determined by computing intraclass reliability coefficients utilizing a one-way repeated-measures analysis of variance (one-way ANOVA) model for a single measure and average measures across days for either T_1 and T_2 or T_2 and T_3 for all participants. T_2 and T_3 were used for only 12 participants (6 females and 6 males) who performed the testing protocol on three separate occasions owing to one or more tests being more than 12% different from the previous trial.[23] Whichever two trials were used, they were henceforth referred to as T_1 and T_2. The reliability of a specific test was considered acceptable if the intraclass reliability coeffiecient (R) was ≥.80[23,24] and the root mean square error (RMSE), the square root of the mean square within people (MSWP), was less than 10% of the T_1 mean.[25]

Pearson product moment correlation coefficients were used to determine the validity of the field tests (FG-TE and B-90° DTE) evaluated against the laboratory tests (PRC-DTE, PRC-STE, and DSBL) using data from T_1. A strong and acceptable correlation coefficient giving evidence of validity between a field test and the corresponding laboratory test was considered to be between 0.80 and 1.0.[24] Dependent *t*-tests determined whether there were any differences between T_1 and T_2 by sex. The alpha level was set at .05.

Results

Performance data by trial and sex are presented in Table 1. In each test item, higher scores indicate better performance. There were several significant differences between Trials 1 and 2 by sex. The female group scored lower on T_1 vs. T_2 for the PRC-DTE (reps), B-90° DTE (reps), PRC-STE (s), and DSBL (kg), while the male group scored significantly lower for the B-90° DTE (reps), PRC-STE (s), and DSBL (kg) at T_1 than T_2.

Intraclass test-retest reliability coefficients for a single measure and average measure for each test are presented in Table 2. The reliability of each test's data (range = .940–.999) was significant and acceptable.

Validity evidence for each test's data is presented in Table 3. Significant correlations were achieved between the PRC-DTE and the B-90° DTE for both the male and female groups and for the males between PRC-STE and B-90° DTE. However, the coefficient between the PRC-DTE and the B-90° DTE for the males was the only one to reach the predetermined level for acceptance. There were no significant correlations between FG-TE and any laboratory test.

Discussion

This study was designed with two purposes in mind: first, to determine the test-retest reliability of selected laboratory and field tests' data of dynamic and static low back function, and second, to evaluate the validity of the field tests' data (B-90° DTE and FG-TE) when compared to laboratory tests' data (PRC-DTE, PRC-STE, and DSBL).

Test-Retest Reliability

In order for a test to be valid, it must first be reliable and reproducible across days.[24] All the intraclass correlation coefficients (.940–.999) and the RMSE values achieved the preestablished criteria for acceptance for all five performance tests. Despite the high reliability statistics, females exhibited a significant mean difference between T_1 and T_2 on four of the five tests (PRC-DTE, B-90° DTE, PRC-STE, and DSBL)

Table 1 Descriptive Data for Participants

Variable	Mean	*SD*	Range
Parallel Roman Chair Dynamic Trunk Extension (reps)			
Trial 1 (Female)	50.59*	22.64	18–140
Trial 2	52.56	23.68	17–145
Trial 1 (Male)	57.05	34.34	15–160
Trial 2	57.73	34.11	14–165
Box 90° Dynamic Trunk Extension (reps)			
Trial 1 (Female)	44.03*	19.77	18–113
Trial 2	46.53	20.17	19–110
Trial 1 (Male)	53.38*	29.99	14–165
Trial 2	54.60	30.00	13–165
Parallel Roman Chair Static Trunk Extension (s)			
Trial 1 (Female)	132.03*	46.04	62.51–235.84
Trial 2	137.41	47.22	65.07–246.84
Trial 1 (Male)	146.92*	65.14	23.97–340.22
Trial 2	150.17	66.99	24.33–361.32
FITNESSGRAM Trunk Extension (cm)			
Trial 1 (Female)	36.37	7.24	26.67–54.61
Trial 2	36.50	7.21	26.67–53.34
Trial 1 (Male)	36.07	9.70	22.86–66.04
Trial 2	36.22	9.86	22.86–66.04
Dynamometer Static Back Lift (kg)			
Trial 1 (Female)	68.97*	13.42	46.82–109.10
Trial 2	71.57	13.27	50.91–110.0
Trial 1 (Male)	123.83*	24.63	87.73–192.27
Trial 2	127.22	25.43	86.36–200.45

Note: Female $n = 32$; Male $n = 40$.
*Significant difference Trial 1 vs. Trial 2, $p < .05$.

while males did so on three of the five tests (B-90° DTE, PRC-STE, and DSBL). Although each participant was familiarized with each item prior to data collection, and on some occasions a third testing trial was performed, there could still have been a slight learning curve. That is, previous experience with a relatively novel item in an actual testing situation may have aided the participant's performance on a subsequent trial. There was no significant change in the presumably more familiar FG-TE item. Close examination of the data, however, reveals that the differences between trial means were very small. The female group improved only 1.97 reps, 2.5 reps, 5.38 s, and 2.6 kg for the PRC-DTE, B-90° DTE, PRC-STE, and DSBL, respectively. The male group improved only 1.22 reps, 3.25 s, and 3.39 kg on the

Table 2 Intraclass Reliability Coefficients for a Single and Average Measures Across Days for Each Test and Root Mean Square Error

Variable	Single measure	Average measure	RMSE	% of Mean
Parallel Roman Chair Dynamic Trunk Extension (reps)				
Female	.992	.996	2.11	4.2
Male	.995	.998	2.35	4.1
Box 90° Dynamic Trunk Extension (reps)				
Female	.99	.993	2.44	5.5
Male	.996	.998	1.81	3.4
Parallel Roman Chair Static Trunk Extension (s)				
Female	.990	.995	4.89	3.7
Male	.994	.997	5.19	3.5
FITNESSGRAM Trunk Extension (cm)				
Female	.998	.999	0.35	0.96
Male	.998	.999	0.43	1.19
Dynamometer Static Back Lift (kg)				
Female	.940	.970	3.25	4.7
Male	.98	.99	4.92	4.0

Note: Female $n = 32$; Male $n = 40$. All coefficients were significant, $p < .05$; RMSE = Root mean square error (square root of the mean square within people); % of Mean = Percentage of the T_1 mean value represented by the RMSE.

Table 3 Pearson Product Moment Correlation Coefficients for Field Tests With Laboratory Tests

Test items	Sex	B-90° DTE	FG-TE
PRC-DTE	Female	.62*	–.11
	Male	.82*	.23
PRC-STE	Female	.38*	.33
	Male	.55*	–.15
DSBL	Female	–.23	–.36
	Male	–.29	–.04

Note: Female: $n = 32$; Male: $n = 40$. PRC-DTE = Parallel Roman Chair Dynamic Trunk Extension; B-90° DTE = Box 90° Dynamic Trunk Extension; PRC-STE = Parallel Roman Chair Static Trunk Extension; FG-TE = *FITNESSGRAM* Trunk Extension.
*Significant, $p < .05$

B-90° DTE, PRC-STE, and DSBL tests. These slight improvements, although statistically significant, do not carry much practical meaning.

These reliability results indicate that when assessing the muscular strength and endurance of the low back with these test items, only one trial needs to be performed as long as the participants have had a sufficient familiarization period with the tests. Given that not all participants' scores were within 12% after two trials, the most conservative recommendation is to have a familiarization period for safety and two practice test trials on different days before testing once and recording the score.

Validity

Validity evidence evaluating the field tests (B-90° DTE and FG-TE) against the laboratory tests (PRC-DTE, PRC-STE, and DSBL) was highest between the B-90° DTE and PRC-DTE for both the male ($r = .82$; $p < .05$) and female ($r = .62$; $p < .05$) groups (Table 3). These correlations undoubtedly reflect both the similarity in the basic movement pattern and component of fitness (dynamic muscular endurance) of the two tests plus the minor differences related to the equipment variation. However, only the validity values for the male group achieved the preset standard of ≥.80. In addition to the high correlations between these two tests, the mean number of repetitions performed between tests was not statistically different.

It was expected that the validity coefficients would be lower between the B-90° DTE and the PRC-STE than with the PRC-DTE, primarily because a static test was being compared to a dynamic test and the equipment was different. This was the case (Table 3). Although the relationship between the static test and the B-90° DTE was significant for both the male ($r = .55$; $p < .05$) and female ($r = .38$; $p < .05$) groups, neither of these values achieved the preselected acceptable standard of .80–1.0.

The comparisons between the B-90° DTE and the DSBL resulted in low (<.80), negative, nonsignificant relationships for both males and females (Table 3). There are two probable reasons for these low values. First, the B-90° DTE is a dynamic measurement while the DSBL is primarily a static measurement. Second, the B-90° DTE test is a measure of muscular endurance while the DSBL is a measure of muscular strength. The correlations among the various test items support this contention. The two muscular endurance items (PRC-DTE and PRC-STE) were more highly related to each other ($r = .26$ F; $r = .43$ M, $p < .05$) than the strength item (DSBL) to either PRC-DTE ($r = –.13$ F; $r = –.15$ M) or PRC-STE ($r = –.19$ M and F), and only the PRC-DTE vs. PRC-STE correlation for the males was significant. Obviously each item was measuring, as intended, different components of neuromuscular function.

The relationship between the FG-TE and the PRC-DTE was low and negative for females and low and positive for males. The relationship between the FG-TE and the PRC-STE was low and positive for females and low and negative for males (Table 3). The FG-TE is primarily a measure of abdominal and lumbar flexibility. Minimal back extensor strength and no endurance are required, as the participant performs only one short contraction.[22] However, both the PRC-DTE and the PRC-STE are muscular endurance tests.

The relationships between the FG-TE and the DSBL for both the male and female groups were low, negative, and nonsignificant (Table 3). Even though both tests involve a single contraction, the FG-TE is initially a dynamic contraction whereas the DSBL is a static contraction. Also, the FG-TE requires only minimal muscular strength while the DSBL is a maximal measurement of muscular strength.

Taken together, the validity evidence indicates that the B-90° DTE and FG-TE are not measuring the same thing, and that the FG-TE is not an acceptable test of either static or dynamic muscular endurance or static muscle strength. Previous research by Johnson et al.[26] on just 12 young adults reported an R of .89 and that isokinetic back strength, but not muscular endurance or passive flexibility, was a significant predictor of the FG-TE. Conversely, Patterson et al.[27] showed that the modified version of the FG-TE used in this study (no 12-inch limit imposed) was both reliable (R = .93 F; R = .95 M) and significantly correlated with goniometer measurement of lumbar flexibility in 43 male (r = .70; p < .01) and 45 female (r = .68; p < .01) high school students.

Although these correlations would not have achieved the preselected standard for validity of .80 set in the current study, they are clearly higher than any of the correlations between FG-TE and the muscular strength/endurance measures in the current study. The FG-TE test item obviously requires a minimum of extensor strength to lift the trunk against gravity. However, if it is desirable to test the strength/endurance of this musculature, another item probably should be used for this age group. Strength, muscular endurance, and flexibility are separate components of neuromuscular function, each of which can be expressed either statically or dynamically. Comprehensive back function testing would require a series of test items to cover all components. This is neither practical nor feasible in a school setting. However, inclusion of a more specific strength/endurance trunk extension test at the high school level would be feasible.

A major purpose of the present study was to evaluate an alternative field test of trunk extension to the current FG-TE. The data showed reliability and produced significant relationships between the B-90° DTE and both the PRC-DTE and PRC-STE, although only the B-90° DTE and PRC-DTE coefficient for males achieved the preselected level for validity. These results are promising, but there are several issues that need to be resolved prior to full endorsement of the B-90° DTE. Several participants mentioned discomfort from the box (which additional padding or the use of a standard weight bench may or may not be able to eliminate) as a reason for stopping. No participant complained about this type of discomfort from the PRC apparatus.

It is important that the testing apparatus allow for fatigue rather than discomfort to determine the endpoint. Although labeled for purposes of this study as a laboratory/criterion test, in practice the PRC is equipment frequently found in weight rooms. At approximately $200–300, it is not prohibitively expensive for many high schools, can serve a multiple purpose as a training device, and was well received by and functioned well for individuals in this age group. Therefore, it seems reasonable to pursue the use of the PRC test items at the high school level. While the static version of the PRC test would be more directly related to the prediction of LBP,[9,11,12,15] it is recommended that both the PRC-DTE and PRC-STE be further investigated to determine how they each related to back function and LBP in this age group.

Conclusions and Recommendations

The PRC-DTE, PRC-STE, B-90° DTE, FG-TE, and DSBL were shown to be reliable tests for a single measure and average measure across days for both sexes. Validity evidence for the B-90° DTE was limited to males, and only when the comparison test was the PRC-DTE. There was no evidence of validity of the FG-TE as a test of trunk extensor strength/endurance, and its designation as such in the *FITNESSGRAM* battery should be reconsidered. The low correlation between the PRC-DTE and PRC-STE indicates that these items cannot be used interchangeably. It is recommended that both the PRC static and dynamic test items be researched further as possible test items for healthy back function in high school age students.

References

1. Plowman SA. Muscular strength, endurance, and flexibility. http://www. cooperinst.org/ftgrefintro.asp, 2001.
2. Walters PH. Back to the basics, strengthening the neglected lower back. *ACSM's Health and Fitness J.* 2000; 4:19-25.
3. Balagué F, Damidot P, Nordin M, Parnianpour M, Waldburger M. Cross-sectional study of the isokinetic muscle trunk strength among school children. *Spine.* 1993; 18: 1199-1205.
4. Brattberg G, Wickman V. Prevalence of back pain and headache in Swedish school children: A questionnaire survey. *The Pain Clinic.* 1992; 5:211-220.
5. Burton AK, Clarke RD, McClune TD, and Tillotson KM. The natural history of low back pain in adolescents. *Spine.* 1996; 21:2323-2328.
6. Fairbank JCT, Pynsent PB, Van Poortvliet JA, Phillip H. Influence of anthropometric factors and joint laxity in the incidence of adolescent back pain. *Spine*. 1984; 9:461-464.
7. Salminen JJ, Maki P, Oksanen A, Pentti J. Spinal mobility and trunk muscle strength in 15-year-old schoolchildren with and without low-back pain. *Spine.* 1992; 17:405-411.
8. Swärd L, Hellstrom M, Jacobsson B, Peterson L. Back pain and radiologic changes in the thoraco-lumbar spine of athletes. *Spine.* 1990; 15:124-129.
9. Biering-Sorensen F. Physical measurements as risk indicators for low-back trouble over a one-year period. *Spine.* 1984; 9:106-119.
10. Plowman SA. Physical activity, physical fitness, and low back pain. In: Holloszy JO, ed. *Exercise and Sport Sciences Reviews*. Baltimore: Williams & Wilkins 1992; 221-242.
11. Hultman G, Nordin M, Saraste H, Ohlsentt H. Body composition, endurance, strength, cross-sectional area, and density of MM erector spinae in men with and without low back pain. *J Spinal Disord.* 1993; 6:114-123.
12. Jorgensen K, Nicolaisen T. Trunk extensor endurance: Determination and relation to low-back trouble. *Ergonomics*. 1987; 30:259-267.
13. Latimer J, Maher CG, Refshauge K, Colaco I. The reliability and validity of the Beiring-Sorensen test in asymptomatic subjects and subjects reporting current or previous nonspecific low back pain. *Spine*. 1999; 24:2085-2090.
14. Nicolaisen T, Jorgensen K. Trunk strength, back muscle endurance and low-back trouble. *Scand J Rehab Med.* 1985; 17:121-127.
15. Luoto S, Heliovaara M, Hurri H, Alaranta H. Static back endurance and the risk of low-back pain. *Clin Biomech*. 1995; 10:323-324.
16. Martin SB, Jackson AW, Morrow JR Jr., Liemohn WP. The rationale for the sit and reach test revisited. *Meas Phys Edu. Exerc Sci*. 1998; 2:85-92.

17. Graves JE, Pollock ML, Carpenter DM, Leggett SH, Jones A, MacMillan M, et al. Quantitative assessment of isometric lumbar extension strength. *Spine.* 1990; 15:289-294.
18. Kearns CF, Brechue WF, Bauer J, Pollock ML, Fulton M. Muscle activation during isolated and non-isolated lumbar extension exercise [Abstract]. *Med Sci Sports Exerc.* 1997; 29(Suppl.): S165.
19. Fleck SJ, KraemerWJ. *Designing Resistance Training Programs* (2nd ed). Champaign, IL: Human Kinetics; 1997.
20. Gordon CC, Chumlea WC, Roche AF. Stature, recumbent length, and weight. In: Lohman TG, Roche AF, Martorell R, eds. *Anthropometric Standardization Reference Manual*. Champaign, IL: Human Kinetics; 1988: pp. 3-8.
21. Heyward VH. *Advanced Fitness Assessment and Exercise Prescription* (4th ed). Champaign, IL: Human Kinetics; 2000.
22. Cooper Institute for Aerobic Research. *FITNESSGRAM™ Test Administration Manual* (2nd ed). Champaign, IL: Human Kinetics; 1999.
23. Fenstermaker KL, Plowman SA, Looney MA. Validation of the Rockport Fitness Walking Test in females 65 years and older. *Res Q Exerc Sport.* 1992; 63:322-327.
24. Hyllegard R, Mood DP, Morrow JR Jr. *Interpreting Research in Sport and Exercise Science.* St. Louis, MO: Mosby; 1996.
25. Looney MA. When is the intraclass correlation coefficient misleading? *Meas Phys Educ Exerc Sci.* 2000; 4:73-78.
26. Johnson KR, Miller MA, Liemohn WP. An examination of factors contributing to performance on the FITNESSGRAM™ trunk lift test [Abstract]. *Med Sci Sports Exerc.* 1997; 29(Suppl.):S9.
27. Patterson P, Rethwisch N, Wiksten D. Reliability of the trunk lift in high school boys and girls. *Meas Phys Educ Exerc Sci*. 1997; 1:145-151.