

ORIGINAL RESEARCH

COMPARISON OF DIFFERENT TRUNK ENDURANCE TESTING METHODS IN COLLEGE-AGED INDIVIDUALS

Michael P. Reiman, PT, DPT, OCS, SCS, ATC, LAT, FAAOMPT, CSCS^{1,4}

Amber D. Krier, DPT²

Julie A. Nelson, DPT³

Michael A. Rogers, DPT⁴

Zachariah O. Stuke, DPT, FAAOMPT⁵

Barbara S. Smith, PT, PhD⁴

ABSTRACT

Objective: Determine the reliability of two different modified (MOD1 and MOD2) testing methods compared to a standard method (ST) for testing trunk flexion and extension endurance.

Participants: Twenty-eight healthy individuals (age 26.4 ± 3.2 years, height $1.75 \pm$ m, weight 71.8 ± 10.3 kg, body mass index 23.6 ± 3.4 m/kg²).

Method: Trunk endurance time was measured in seconds for flexion and extension under the three different stabilization conditions. The MOD1 testing procedure utilized a female clinician (70.3 kg) and MOD2 utilized a male clinician (90.7 kg) to provide stabilization as opposed to the ST method of belt stabilization.

Results: No significant differences occurred between flexion and extension times. Intraclass correlations (ICCs_{3,1}) for the different testing conditions ranged from .79 to .95 ($p < .000$) and are found in Table 3. Concurrent validity using the ST flexion times as the gold standard coefficients were .95 for MOD1 and .90 for MOD2. For ST extension, coefficients were .91 and .80, for MOD1 and MOD2 respectively ($p < .01$).

Conclusions: These methods proved to be a reliable substitute for previously accepted ST testing methods in normal college-aged individuals. These modified testing procedures can be implemented in athletic training rooms and weight rooms lacking appropriate tables for the ST testing.

Key Words: Core, stabilization, trunk endurance

Level of Evidence: 3

CORRESPONDING AUTHOR

Michael P. Reiman, PT, DPT, OCS, SCS, ATC,
FAAOMPT, CSCS

Assistant Professor

Duke University School of Medicine

Department of Community and Family
Medicine

Doctor of Physical Therapy Division

DUMC 104002

Durham, NC 27710

Phone: (919) 668-3014

Fax: (919) 684-1846

Email: reiman.michael@gmail.com

¹ Duke University Medical Center, Durham, NC, USA

² Mitchell County Hospital, Beloit, KS, USA

³ Physical Therapy at Cypress, Wichita, KS, USA

⁴ Wichita State University, Wichita, KS, USA

⁵ Mid-America Orthopedics, Wichita, KS, USA

The authors declare no conflict of interest or financial disclosures. IRB approval received from Wichita State University Institutional Review Board.

INTRODUCTION

Trunk/core stability and endurance have been investigated with respect to potential contribution to injury and athletic performance in activities such as running and jumping. Workers who reported low back pain (LBP) had decreased trunk extensor endurance.¹ This lack of endurance appears to be a predictor of future occurrences of LBP.^{2,3} Imbalance between flexion and extension trunk muscle endurance times may be even more important than isolated trunk endurance deficits.⁴ The cause of LBP appears to be associated with an imbalance of flexion-extension endurance times, with the extensors having less endurance than the flexors.⁴

Expensive isokinetic testing has been used to assess strength and work of the trunk musculature.^{5,6} This type of testing, although able to provide the clinician with discrete data, is expensive, time intensive, and requires a lot of space. Less expensive isometric trunk endurance testing is more clinically available and practical.^{3,7} This type of testing also affords more clinical applicability.

Acceptable clinical tests that measure the strength or power component of trunk stability have been recently advocated because they may be more useful,⁸ and may better mimic the demands imposed by sport.⁹ While strength and power are more likely representative of athletic explosive demands, trunk strength appears to have little, or a very weak, relationship with low back health.¹⁰ Trunk endurance testing, however, continues to be warranted and necessary for at least two reasons. One, these muscles are predominantly type I muscle fibers^{11,12} that appear to become more anaerobic as a result of deconditioning.¹⁰ Two, trunk muscles' maximum isometric strength was not associated with LBP in athletes. However, the same authors found a relationship between trunk muscle endurance imbalance and LBP in the same athletes.¹³ Additionally, balance between the various trunk muscle endurance values was found to be a better predictor of LBP than strength alone.⁴

Normative values exist for trunk endurance assessments among college-aged students with no history of LBP,¹⁴ as well as college-aged male rowers.¹⁵ These methods used for testing trunk endurance demonstrate excellent reliability^{14,15} Implementation of these testing techniques in some rehabilitation and athletic training environments is, however, potentially limited because these tests require appropriate tables and the use of multiple belts. Therefore, the use of these methods in settings without the correct equipment or with large groups may be limited. The purpose of this research was, therefore, to compare the reliability of modified trunk flexion and extension testing set-ups to previously established testing procedures. Concurrent validity between endurance times (in seconds) using the standard method and the modified methods was also examined.

METHODS

Participants

An a priori power analysis showed that a sample size of 25-30 individuals would be needed to achieve a moderate effect size (.60) and 80% power. The participants consisted of 28 athletic individuals (14 females and 14 males), with no history of hip or LBP within the past 6 months, no history of hip or lumbar surgery, and no experience with the testing methods (Table 1). Participants were physically active in aerobic and/or strength training (running, basketball, and/or weightlifting) one to four times per week. The study was approved by the Wichita State University Institutional Review Board for the Protection of Human Subjects. Informed consent was obtained. Participant rights were protected and the investigation conformed to the protocol and ethical and humane research principles.

Methods

Written information and oral instructions were given before each test. All participants were timed during the performance three techniques (standard [ST],

Table 1. Demographic Characteristics of Participants (n = 28).

	Age (years)	Height (m)	Weight (kg)	BMI (kg/m ²)
Mean ± SD	26.4 ± 3.2	1.75 ± .9	71.8 ± 10.3	23.6 ± 3.4
(Minimum – maximum)	23 - 38	1.60 – 1.93	56.70 – 90.72	19.6 – 34.4
Abbreviations: SD= standard deviation; m= meters, kg= kilograms; BMI= body mass index				

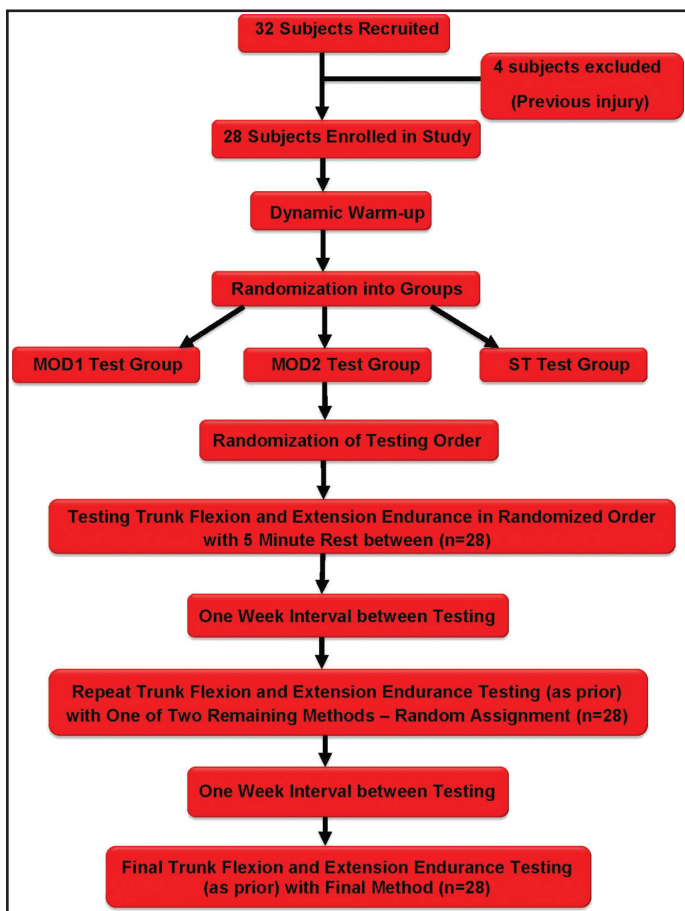


Figure 1. Flow Diagram of Study Design.

modification one [MOD1], modification two [MOD2]) for both flexion and extension at one-week intervals (Figure 1). Testing type (flexion or extension) was randomly selected at the first session. Testing order was randomly determined at each testing session. A 10-minute dynamic warm-up (easy jogging, dynamic movement drills including marching, marching toe touches, and hurdle step overs) was performed by subjects at the beginning of each session. Participants rested 5 minutes between flexion and extension testing at all sessions. The MOD1 and MOD2 techniques used two different clinicians, one female, MOD1, and one male, MOD2, as described below to provide stabilization in lieu of belts for the ST technique. The clinician not providing stabilization performed the endurance timing as per previous investigation¹⁶ and specifically described below.

Maintaining the test position as long as possible was encouraged prior to all testing, using standardized instructions. Participants were not encouraged during testing. Each participant performed one trial of

each testing method. Participants were informed of their times after study completion. Participants were encouraged to maintain their current activity level between testing sessions, specifically for trunk muscle activity. This included specific instruction not to change the current frequency or intensity of the aerobic, anaerobic, weightlifting activities.

Trunk Extension Testing Procedures

Standard testing for trunk extension was performed according to McGill et al's modification.¹⁴ The original method, from which this modification was modeled, is a reliable measure of back extensor endurance.¹ For all trunk extension testing, participants were instructed to lie prone. The lower body was fixed to the table surface via straps at the ankles, knees, and hips for standard testing. The upper body (from just above the level of the anterior superior iliac crest) was off the surface of the plinth. Participants held their upper body off the end of the table by pushing with their extended arms on a chair directly below them. Participants were instructed to maintain the horizontal position for as long as possible once testing commenced. At the initiation of the test, the upper limbs were lifted off the chair and crossed over the chest with the hands resting on the opposite shoulders. A stopwatch was used to time from the instant the upper limbs were lifted off the chair and crossed over the chest (and the participant assumed the horizontal position) as described above until the participant visually deviated from the horizontal plane. The same procedures were used for all methods of measuring extension.

The modified testing procedures used a clinician to hold the participants' lower extremities down, replacing the straps used during the standard procedure. A clinician lay across the backs of the lower extremities so that the middle of the clinician's trunk was over the middle of the participant's posterior knees (Figure 2). The clinician (either MOD1 or MOD2) not providing stabilization to the participant measured the endurance time. This clinician stood to the participant's side and used the above-mentioned criteria of the participant's visually deviating from the horizontal plane as the criterion for ending the extensor endurance test.

The MOD1 technique utilized a female clinician (age 26 years; height 1.80 m and weight 70.30 kg) and the



Figure 2. Modified Trunk Extension Testing Method.

MOD2 technique utilized a male clinician (age 25 years; height 1.93 m and weight 90.70 kg) to provide stabilization. None of the participants related any complaints regarding the modified testing procedures.

Trunk Flexion Testing Procedures

Standard testing for trunk flexion was performed according to previously published methods.¹⁴ For all trunk flexion testing, participants were positioned supine, with both hips and knees flexed to 90 degrees, trunk inclined at 60 degrees resting on a prefabricated wedge. Stabilization was achieved with a belt around the table and over the dorsum of the feet (with shoes on) for the standard method. Participants crossed their arms across the chest, placing their hands on opposite shoulders, in a manner comfortable to them. Participants maintained their body position for as long as possible after the wedge was moved back 10 cm. Time was measured from the instant the prefabricated wedge was moved back until the participant visually reestablished contact with the wedge. This was the same for all methods.

Modified testing procedures used a clinician to hold the participants' feet rather than using straps (Figure 3). The clinician (either MOD1 or MOD2) not providing stabilization to the participant measured the endurance time. This clinician stood to the participant's side and used the above-mentioned criteria of the participant's visually re-contacting the wedge as the criterion for ending the endurance test.

Statistical Analyses

Means and standard deviations for all endurance measurements were generated. Repeated measures



Figure 3. Modified Trunk Flexion Testing Method.

analyses of variance (ANOVA) with Bonferroni correction were used to analyze differences in mean flexion and extension times. The intraclass correlation coefficients (ICC_{3,1}) were used to assess reliability between test methods. Pearson's *r* was used to analyze concurrent validity between the gold standard (ST times) and times using the modified approaches. SPSS V 17.0 (Chicago, IL) was used to analyze the data. Alpha level was set at 0.05.

RESULTS

The mean endurance times for the three methods assessing flexion endurance time ranged from 304.73 ± 207.32 seconds (MOD1 flexion) to 344.26 ± 217.09 seconds (MOD2 flexion). Mean endurance times ranged from 141.12 ± 50.44 seconds (MOD1 extension) to 148.87 ± 43.68 seconds (MOD2 extension) (Table 2). Means for flexion and extension times were not significantly different. Intraclass correlations (ICCs_{3,1}) for the different testing conditions ranged from .79 to .95 (*p* < .000) and are found in Table 3. Concurrent validity using the ST flexion times as the gold standard coefficients were .95 for MOD1 and .90 for MOD2. For ST extension, coefficients were .91 and .80, for MOD1 and MOD2 respectively (*p* < .01).

DISCUSSION

Modified stabilization methods appear to be a valid and reliable substitute when standard method equipment is not available. These modified techniques may be utilized by clinicians of different sizes and in different settings. Many rehabilitation settings do

Table 2. Mean (\pm SD) time (seconds) and minimum/maximum values (time in seconds) of the three testing methods ($n = 28$).

	Mean (seconds)	Range
ST flexion	339.32 \pm 214.47	(62.6 – 966.4)
MOD1 flexion	304.73 \pm 207.32	(73.8 – 982.0)
MOD2 flexion	344.26 \pm 217.09	(84.5 – 866.2)
ST extension	147.22 \pm 50.33	(77.9 – 274.0)
MOD1 extension	141.12 \pm 50.44	(76.5 – 274.0)
MOD2 extension	148.87 \pm 43.68	(61.1 – 220.7)

Abbreviations: ST= standard testing procedure; MOD1= modified testing procedure with first clinician (female; height of 1.80 m and weight of 70.30 kg; BMI of 21.6); MOD2= modified testing procedure with second clinician (male; height of 1.93 m and weight of 90.70 kg; body mass index of 24.3).

Table 3. Reliability of Testing Conditions ($n=28$). All ICCs are significant at the $P < 0.000$ level.

Test condition comparisons	ICC _{3,1}	95% CI	SEM
ST – MOD1 flexion	0.95	.90-.98	47.15
ST – MOD2 flexion	0.90	.80-.95	68.23
MOD1 – MOD2 flexion	0.89	.77-.95	70.38
ST-MOD1-MOD2 flexion	0.91	.85-.96	63.78
ST – MOD1 extension	0.80	.62-.90	22.53
ST – MOD2 extension	0.90	.80-.95	14.86
MOD1 – MOD2 extension	0.79	.59-.90	21.56
ST-MOD1-MOD2 extension	0.83	.71-.91	20.29

Abbreviations: ST= standard testing procedure; MOD1= modified testing procedure with female clinician; MOD2= modified testing procedure with male clinician; ICC= Intraclass Correlation Coefficient; CI= Confidence Interval; SEM= standard error of measure.

not have tables that allow a belt to be wrapped under its surface, eliminating the possibility of using the original testing methods. Alternative methods, such as those used in this study, therefore, can be implemented reliably.

The ICCs for flexion and extension are considered moderate to excellent, respectively.¹⁷ Other studies' ICC values are higher because researchers used multiple measurements for each test.^{14,15} To prevent fatigue from interfering with the participants' best effort, time was measured only once for each method at weekly intervals. ICCs using average measures are consistently higher than those measures using single measurements.¹⁸

The authors of the current study are aware of only one other study that has investigated the correlation between the ST methods and any modification.¹⁶ Interrater reliability for the modified testing procedures using a clinician in the same manner as the current study was 0.97 for extension and 0.93 for flexion. Correlation of endurance times for this modified testing¹⁶ with the ST procedure was 0.90 and 0.84 for extension and flexion, respectively. A limitation in the Reiman et al. study¹⁶ was that the clinician providing stabilization was always larger than the participant.

Concurrent validity of the MOD tests with the ST test was determined by the correlation between the extension and flexion times, respectively. The absolute

value of validity is an accurate indicator of the extent of validity determined as follows: correlation coefficients $> .80$ indicate high validity, values between $.60$ and $.80$ indicate good validity, values between $.40$ and $.59$ indicate moderate validity, and values $< .40$ indicated poor validity.¹⁹ Thus, the values from this study show that modified methods are valid measures as efficient as compared to the gold standard (use of belts) and, therefore, may be used instead.

The reliability reported in the current study was lower than previously reported values.¹⁴ The testing method in the McGill et al study used only five participants tested consecutively for 5 days and then once 8 weeks later.¹⁴ The current study involved more participants ($n = 28$) and fewer test sessions, which is more likely representative of clinical situations.

Although the clinician who provided stability is not the same as a static belt, no resistance on the part of the clinician was required. The clinician simply lay over the lower extremities (or sat on the feet in the case of flexion) and remained stationary, using his/her body mass as a stabilizing force throughout the test. Neither the clinician nor the participants reported any adverse effects of this method. All of the participants subjectively reported that they were stabilized equally during all sessions.

Limitations

There are a number of limitations in this study that warrant mention. Although these modified testing methods appear to be acceptable alternatives to the ST testing procedure in a normal, asymptomatic population, their applicability in LBP subjects is unknown. Trunk extensor²⁰ and flexor^{7,20,21} muscle endurance times in subjects with LBP are less than those in the normal healthy population and it is unknown whether this would impact the reliability or validity of the methods. Trunk muscle endurance testing methods have been implemented in different LBP populations.²²⁻²⁴ It remains to be determined whether the modified procedures will compare favorably in a symptomatic population.

The criterion for test termination for measuring flexion endurance was previously established as deviating from the 60-degree angle.^{14,15} The authors of the current study determined, with pilot testing,

that reliability was lower using the criterion of breaking the 60-degree plane in any manner compared to the subject's breaking the plane by contacting the pre-fabricated wedge. We believe the latter more clearly defines test termination. In pilot testing, it was often difficult to ascertain if the subject broke the 60-degree plane throughout the entire spine. Many subjects would lose lumbar lordosis, yet appear to maintain the 60-degree angle and not contact the wedge. Due to this potential complication in test termination criterion, which was also encountered by Chan,¹⁵ the authors decided on the termination criterion being when the subject contacted the pre-fabricated wedge. Although reliability was lower than the initial standardization studies,¹⁴ it was still high. The different criterion for test termination may be a factor for the differences seen between the current results and those of Chan¹⁴ and McGill et al.¹⁵

These modified techniques were found to be reliable methods of testing trunk endurance, allowing for potential use in research to determine normative values for different populations. Future research regarding modified endurance testing methods should investigate the size relationship between the clinician and testing participant in different populations. Determining the ratio of clinician to participant size needed for sufficient stabilization in order to achieve reliable outcomes could provide additional information.

PRACTICAL APPLICATIONS

The modified testing methods employed in this study can be utilized as substitute testing methods in healthy adults when proper tables and belts are not available. Having athletes partner with someone of similar size may provide greater validity, but it is not essential. Multiple athletes can be tested at the same time when the athletes partner together and provide stabilization for each other. These procedures, therefore, could be implemented in most athletic training rooms and weight rooms and afford greater efficiency when testing large groups.

CONCLUSION

The use of the MOD1 and MOD2 procedures to test trunk flexion and extension endurance is a reliable alternative to the ST method. The use of these modified procedures will allow clinicians in settings without appropriate tables to assess trunk endurance times

accurately. Having another individual stabilize participants also allows testing of large groups, such as athletic teams, more efficiently than with standard testing. Further study is necessary to determine these testing procedures' reliability and applicability with participants with LBP.

REFERENCES

1. Alaranta H, Hurri H, Heliovaara M, Soukka A, Harju R. Non-dynamometric trunk performance tests: reliability and normative data. *Scand. J. Rehabil. Med.* 1994;26(4):211-215.
2. Biering-Sorensen F. Physical measurements as risk indicators for low-back trouble over a one-year period. *Spine.* 1984;9(2):106-119.
3. Alaranta H, Luoto S, Heliovaara M, Hurri H. Static back endurance and the risk of low-back pain. *Clin Biomech.* 1995;10(6):323-324.
4. McGill S, Grenier S, Bluhm M, Preuss R, Brown S, Russell C. Previous history of LBP with work loss is related to lingering deficits in biomechanical, physiological, personal, psychosocial and motor control characteristics. *Ergonomics.* 2003;46(7):731-746.
5. Delitto A, Rose SJ, Crandell CE, Strube MJ. Reliability of isokinetic measurements of trunk muscle performance. *Spine.* 1991;16(7):800-803.
6. Keller A, Hellesnes J, Brox JI. Reliability of the isokinetic trunk extensor test, Biering-Sorensen test, and Astrand bicycle test: assessment of intraclass correlation coefficient and critical difference in patients with chronic low back pain and healthy individuals. *Spine.* 2001;26(7):771-777.
7. Malliou P, Gioftsidou A, Beneka A, Godolias G. Measurements and evaluations in low back pain patients. *Scand J Med Sci Sports.* 2006;16(4):219-230.
8. Leetun DT, Ireland ML, Willson JD, Ballantyne BT, Davis IM. Core stability measures as risk factors for lower extremity injury in athletes. *Med Sci Sports Exerc.* 2004;36(6):926-934.
9. Cowley PM, Swensen TC. Development and reliability of two core stability field tests. *J Strength Cond Res.* 2008;22(2):619-624.
10. McGill SM. *Low Back Disorders: Evidence-Based Prevention and Rehabilitation.* Champaign: Human Kinetics, IL; 2007.
11. MacDonald DA, Moseley GL, Hodges PW. The lumbar multifidus: Does the evidence support clinical beliefs? *Man Ther.* 2006;11(4):254-263.
12. Thorstensson A, Carlson H. Fibre types in human lumbar back muscles. *Acta Physiol Scand.* 1987;131(2):195-202.
13. Renkawitz T, Boluki D, Grifka J. The association of low back pain, neuromuscular imbalance, and trunk extension strength in athletes. *Spine J.* 2006;6(6):673-683.
14. McGill SM, Childs A, Liebenson C. Endurance times for low back stabilization exercises: clinical targets for testing and training from a normal database. *Arch Phys Med Rehabil.* 1999;80(8):941-944.
15. Chan RH. Endurance times of trunk muscles in male intercollegiate rowers in Hong Kong. *Arch Phys Med Rehabil.* 2005;86(10):2009-2012.
16. Reiman MP, Krier AD, Nelson JA, Rogers MA, Stuke ZO, Smith BS. Reliability of alternative trunk endurance testing procedures using clinician stabilization vs. traditional methods. *J Strength Cond Res.* 2010;24(3):730-736.
17. Portney LG, Watkins, M.P. *Foundations of Clinical Research - Applications to Practice. 2nd ed.* Upper Saddle River, New Jersey: Prentice-Hall; 2000.
18. Nichols DP. Choosing an intraclass correlation coefficient. 1998; 67. Accessed February 5, 2011.
19. Meyer CR. *Measurement in Physical Education.* New York: Ronald Press; 1974.
20. Roy SH, De Luca CJ, Casavant DA. Lumbar muscle fatigue and chronic lower back pain. *Spine.* 1989;14(9):992-1001.
21. Ito T, Shirado O, Suzuki H, Takahashi M, Kaneda K, Strax TE. Lumbar trunk muscle endurance testing: an inexpensive alternative to a machine for evaluation. *Arch Phys Med Rehabil.* 1996;77(1):75-79.
22. Arab AM, Salavati M, Ebrahimi I, Ebrahim Mousavi M. Sensitivity, specificity and predictive value of the clinical trunk muscle endurance tests in low back pain. *Clin Rehabil.* 2007;21(7):640-647.
23. Ropponen A, Gibbons LE, Videman T, Battie MC. Isometric back extension endurance testing: reasons for test termination. *J Orthop Sports Phys Ther.* 2005;35(7):437-442.
24. Stewart M, Latimer J, Jamieson M. Back extensor muscle endurance test scores in coal miners in Australia. *J Occup Rehabil.* 2003;13(2):79-89.