OBJECTIVITY, RELIABILITY, AND VALIDITY OF THE BENT-KNEE PUSH-UP FOR COLLEGE AGED FEMALES

by

HEATHER MIRANDA WOOD

(Under the direction of Ted Baumgartner)

ABSTRACT

Push-ups have been used in many fitness tests to measure arm and shoulder girdle strength and endurance. The purpose of this study is to determine the reliability, objectivity, and validity of the bent-knee push-up for college-aged females and to determine the relationship between the modified push-up and bent-knee push-up. Eighty-seven college-aged females participated in this study. The interscorer objectivity coefficient for the bent-knee push-up was .997. A stability reliability coefficient of .83 was obtained. The correlation between the bent-knee push-up scores and the bench press scores was .67. The correlation between the revised push-up scores and the bench press scores was .68. Based on these findings, the bent-knee push-up test could be used as a test of arm and shoulder girdle strength and endurance for females to ensure proper measurement of the females with low strength.

INDEX WORDS: Push-ups, Bent-knee Push-ups, Revised Push-ups, Arm and Shoulder Girdle Strength and Endurance

OBJECTIVITY, RELIABILITY, AND VALIDITY OF THE BENT-KNEE PUSH-UP FOR COLLEGE AGED FEMALES

by

HEATHER MIRANDA WOOD

B.S.Ed., The University of Georgia, 2001

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GA

2003

OBJECTIVITY, RELIABILITY, AND VALIDITY OF THE BENT-KNEE PUSH-UP FOR COLLEGE AGED FEMALES

by

HEATHER MIRANDA WOOD

Major Professor: Ted Baumgartner

Committee: Kirk Cureton
Kathy Simpson

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2003

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Assessment of muscular strength and endurance is important in clinical, research, physical education, sport, and fitness testing situations. Enhancement of muscular strength and endurance may be helpful in sport related competition. Adequate levels of strength and endurance can decrease the chances of physical problems and injury such as lower back pain and contribute to good body mechanics such as good posture.

Muscular strength is defined as the maximum force that can be generated by a muscle group during a short period of time. Muscular endurance is the ability to exercise muscle groups over an extended period of time at moderate intensity utilizing aerobic energy and to resist fatigue. A general definition of arm and shoulder girdle strength and endurance is the ability of the arm and shoulder girdle muscles to move weight. A field-based definition is the ability to move or support the body weight against the pull of gravity (Baumgartner, Jackson, Mahar, and Rowe, 2003).

Arm and shoulder girdle strength and endurance can be measured with tests that involve either isometric or isotonic contractions of the muscles and are performed to exhaustion. Some such field-based tests include push-ups, pull-ups, chin-ups, and the flexed-arm hang. Many of the developers of these tests suggest different versions for boys and girls. The Physical Fitness Index (PFI) is a combination of tests used to measure overall strength and endurance (Mathews, 1978). The arm strength and endurance tests for the PFI are, pull-ups, dips, and push-ups. Two different versions of

the pull-up, one for boys and one for girls are administered. Boys only perform the dips; while a push-up with hands on a bench and toes on the ground is used for girls only (Mathews, 1978). In the past, to measure arm and shoulder girdle strength and endurance, the American Alliance for Health Physical Education Recreation and Dance (AAHPERD) Youth Physical Fitness Test listed two types of pull-ups in which the boys executed pull-ups while hanging from a bar, while girls executed pull-ups with the bar lowered so they perform the pull-ups with the knees bent at a 90-degree angle and the feet on the ground. The pull-ups for girls are less difficult than the pull-ups for boys (Mathews, 1978). Presently, the AAHPERD Youth Physical Fitness Test lists the pull-up test for boys and the flexed-arm hang for girls in the (Miller, 2002).

Several different forms of push-ups have been used in field-based physical fitness tests. In the past, push-ups for males were executed with the toes and hands on the ground (full-body push-up), while females performed modified versions, including those with the knees and hands on the ground (bent-knee push-up). In the Complete Guide to Youth Fitness Testing, Safrit (1995) suggests using the 90-degree push-up, in which hands and toes are on the ground and the body is lowered to the ground until the arms are at a 90-degree angle, and the bent-knee push-up. Safrit states that one advantage of the bent-knee push-up is that this version is easier to perform than the 90-degree push-up, which means that participants were more likely to complete at least one. Miller (2002) lists the full-body push-up and the bent-knee push-up as types of measures of arm and shoulder girdle strength and endurance. Miller suggests that both males and females can use the full-body push-up, while only females can use the bent-knee push-up. In the Amateur Athletic Union's (AAU) Physical Fitness Program, males perform the 90-degree

push-up while females perform the bent-knee push-up (1992). In the Indiana Physical Fitness test (Clarke, 1967) and the California Physical Performance tests (Clarke, 1967) the bent-knee push-up is listed as one of the items for girls in the fitness tests. Boys are administered full-body push-ups. These differences suggest different tests for males and females may be necessary.

Currently, the American College of Sports Medicine (ACSM) lists the full-body push-up for males and the bent-knee push-up for females in the muscular fitness portion of the ACSM Fitness Test (Miller, 2002). The Canadian Society for Exercise Physiology also suggests that males perform a full-body push-up and females perform a bent-knee push-up in the Canadian Physical Activity, Fitness & Life style Appraisal (Miller, 2002). These modified versions are used since females tend to be weaker than males and also have different weight distribution.

The President's Council on Physical Fitness and Sports used only the flexed-arm hang or the pull-up for both males and females in past versions of the President's Challenge (1993). Presently, however, the President's Challenge has a 90-degree push-up test in which the toes and hands are on the ground, or either the flexed-arm hang or the pull-up as alternatives for both males and females (2001). The 90-degree version of the push-up is the recommended test in the FITNESSGRAM for both males and females (Cooper Institute, 2001). The alternatives to this test include the flexed arm hang and the pull-up. Not much evidence is presented for using only one form of push-up for both males and females. However, convenience could be a major factor. Using only one form of the test requires less training for those who administer the test. Using only one push-

up version also allows comparison between scores from year to year and possibly between males and females.

Another form of the push-up has recently been studied. Baumgartner, Oh, Chung, and Hales (2002) developed a modified push-up test to make the scoring procedure more accurate than for the 90-degree push-up test. This test still uses only one form for both males and females. They developed the test using college students. However no research has been conducted to indicate their test is better for females and children than the bent-knee push-up test.

Purpose

The purpose of this study is to determine the objectivity, reliability, and validity for the bent-knee push-up for college-aged females and to determine the relationship between the modified push-up (Baumgartner, et al., 2002) and bent-knee push-up.

Limitations

1. Students in weight training classes were used as participants in this study. These participants may be stronger than other college students, which may lead to higher than average scores for each test and fewer zero scores.
2. All participants were volunteers. The weaker students may not have volunteered in fear of not performing well.
3. The effort given by a student to obtain a maximum score could not be controlled.

Delimitations

1. Only college females aged 18-23 were used in this study.
2. Only students in beginning and advanced weight training classes were used as participants.

CHAPTER 2

REVIEW OF THE RELATED LITERATURE

The push-up is used as a measure of arm and shoulder girdle strength and endurance. Different forms of this test appear in a variety of fitness tests and programs. The push-up has been found to be easier to perform than pull-ups and the flexed-arm hang (Cooper Institute, 2001).

Push-Up Tests

Many fitness tests include versions of the push-up test for females that appear to be easier and more discriminating among levels of ability for females than the male push-up version. For example, in the past the Physical Fitness Index required tests included a test where females perform push-ups with the feet on the floor and the hands on a bench 13 inches high, 20 inches long and 14 inches wide, touching the chest to the bench (Clarke, 1967, Mathews, 1978). The list of tests in the California Physical Performance test (Clarke, 1967), the Indiana Physical Fitness test (Clarke, 1967), and the Division for Girls' and Women's Sports test (Mathews, 1978) includes the bent-knee push-up for girls. Safrit and Wood (1995) suggest using the bent-knee push-up for female examinees, stating that the full-body push-up may not be a discriminating measure for some groups since some examinees cannot execute even one. Tritschler (2000) suggests that girls and women perform a bent-knee push-up, while boys and men perform a full-body push-up. The University of Massachusetts (2002) also includes a bent-knee push-up for female participants in the school's health and fitness assessments.

Payne, Gledhill, Katzmarzyk, Jamnik, and Ferguson. (2000) and Tremblay and Chiasson (2002) used the Canadian Home Fitness Test procedures in studies to predict physical fitness in participants. The tests for the push-up portion of the Home Fitness Test are the full-body push-up for men and the bent-knee push-up for women. Payne, et al. (2000) tested males and females aging from 15 to 69 years of age. They found the push-up was a significant predictor of health scores. Tremblay and Chiasson (2002) tested men and women aged 17 to 20 to investigate fitness characteristics in young adults. They found that push-up scores correlated well with physical activity and body weight scores. These correlations from both studies suggest that the bent-knee push-up can be used to determine or predict other fitness scores.

Recently, for some programs like FITNESSGRAM (Cooper Institute, 2001) and President's Challenge (2001) the required test is the 90-degree push-up on the hands and toes for both males and females. These programs are two of the most commonly used fitness programs for youth. Based on the FITNESSGRAM standards, most females from elementary school to college aged could complete at least one of these push-ups. In the FITNESSGRAM (Cooper Institute, 2001), intraclass reliability coefficients (ranging from.64 to .96) are reported for push-up scores of females. However, the reliability coefficient of .96 was obtained using a bent-knee push-up. The developers of the FITNESSGRAM (Cooper Institute, 2001) found from the research literature that, although the push-up scores for college-aged women had good reliability, the push-ups were done with the knees on the floor. The developers also found, based on their norm data, that 5% of both boys and girls over 8 years of age and 10% of both boys and girls ages 6-8 years completed zero 90-degree push-ups. The development of the

FITNESSGRAM was based on a review of the literature, so no research studies were conducted by the developers of FITNESSGRAM to estimate reliability.

The 90-degree push-up or the pull-up is used in the President's Challenge (2001) for both males and females. The flexed-arm hang is listed as an alternative if the participants cannot complete even one of the previous choices. This program is suggested for students aged 6 to 17. The President's Challenge physical fitness program packet does not include any references to research to provide a rationale for the use of these tests for both males and females.

McManis, Baumgartner, and Wuest (2000) conducted a study investigating the objectivity and reliability for 90-degree push-up scores of both males and females from elementary school to college. They tested 150 elementary school, 70 high school, and 84 college-aged students. The participants performed the 90-degree push-up test on two different occasions. The researchers found stability reliability coefficients of .64, .71, .86, and .50 for elementary school girls and boys and high school girls and boys, respectively. The stability reliability coefficients for college-aged females ranged from .75 to .87, while for college-aged males the coefficients ranged from .22 to .75. These reliabilities range from high to low, suggesting the scores may not be consistent from day to day for the 90-degree push-up test. The objectivity estimates were .46 and .75 for elementary school girls and boys, respectively. For college-aged participants the objectivity estimates ranged from .31 to .88 for females and from .16 to .91 for males. These low objectivity estimates may suggest that different judges or scorers do not count the push-ups in the same manner. McManis, et al. (2000) also found that it was difficult to discern correct and incorrect forms with the 90-degree push-up, and low strength

college aged females had difficulties performing the 90-degree push-up correctly. Based on these findings, they state that a modified version of the 90-degree push-up may be needed for adequate measurement of arm and shoulder-girdle strength and endurance of young children and low-strength individuals.

In light of these findings, Baumgartner, Oh, Chung, and Hales (2002) developed a revised push-up test, defining the down position as the body from the chest to the knees contacts the floor. So, their push-up is a full-body push-up. They conducted their research on male and female undergraduate college students. They found good interscorer objectivity coefficients (.75, .88, .97, .95, .98, .99) and good stability reliability coefficients (.90, .93, .95, .95). Strong validity evidence was also found with correlations of .80 for females and .87 for males between revised push-up scores and number of bench press executions with a percentage of the body weight. Despite this strong evidence, many females had scores of zero on the revised push-up test. These zero scores on the revised push-up and 90-degree push-up may indicate that the push-up test on the hands and toes is too difficult for females to perform.

Due to the difficulty of the 90-degree push-up and revised push-up tests for some participants, especially females, the tests may lack the discrimination needed in a fitness test. The bent-knee push-up is easier than the full-body push-up for females to perform. McManis, et al. (2000) found that low strength females had difficulties performing the 90-degree push-up correctly, but were more successful performing bent-knee push-ups as an exercise in their class.

Test Characteristics

Objectivity, reliability, and validity must be acceptable for a test to be considered a good test. Validity exists if the interpretation of the test scores is correct (Baumgartner, Strong, Hensley, 2002). The test scores must be reliable for validity to exist. However, reliability of the scores does not guarantee validity. Validity evidence can be obtained by the criterion approach which is determining the correlation between test scores and scores for a criterion or standard measure of the attribute being studied. The logical approach is another way of determining validity. This approach uses content validity evidence, which shows scores are logically associated with the material being tested. This evidence is based on subjective decisions. A third, more complex approach, is construct validity. This approach is used with abstract tests, which measure items that are not directly observable, such as attitudes toward physical activity (Baumgartner, Jackson, Mahar, and Rowe, 2003).

Baumgartner, et al. (2002) define reliability as the consistency of test scores over a short period of time. Reliability can be estimated by either a stability reliability approach or an internal-consistency reliability approach. With stability reliability, each person is tested on several different days using the same procedures. The stability reliability coefficient is the intraclass correlation between the scores from each day. Internal-consistency reliability is estimated from scores obtained by administering a test multiple times within one day. The internal-consistency reliability coefficient is the intraclass correlation among these scores. Scores are more reliable the closer the intraclass correlation coefficient is to positive one (Baumgartner, et al., 2003).

Objectivity is the consistency of scores assigned by two or more scorers, raters, or judges. Objectivity is necessary before scores can be reliable. The scoring system must be clearly defined and the judges must be sufficiently trained on how to assign a score in order to obtain acceptable objectivity. To estimate objectivity, an intraclass correlation between the scores of the different raters must be calculated. The closer this coefficient is to one, the more objectivity the measurements have (Baumgartner, et. al., 2003).

Criterion Score

For the measure of arm and shoulder girdle strength and endurance, a standard measure is the bench press. The bench press test could be one repetition of the maximum amount of weight a person can press (1RM) or the number of times a person can press a percentage of her body weight. Baumgartner, Oh, Chung, and Hales (2002) used the bench press as the criterion score in determining the validity for the revised push-up test. They used 70% of body weight for males and 40% of body weight for females for the bench press test. Jackson, Fromme, Plitt, and Mercer (1994) and Pate, Burgess, Woods, Ross, and Baumgartner (1993) also used the bench press to show evidence of criterion validity for the push-up. They used 1RM for the test. Since the push-up requires moving a person's body weight up and down, the bench press should also require moving a percentage of a person's body weight.

CHAPTER 3

OBJECTIVITY, RELIABILITY, AND VALIDITY OF THE BENT-KNEE PUSH-UP FOR COLLEGE AGED FEMALES

[1]Wood, H. M. and Baumgartner, T. A.. To be submitted to Measurement in Physical Education and Exercise Science

Abstract

Push-ups have been used in many fitness tests to measure arm and shoulder girdle strength and endurance. The 90-degree and revised push-up tests are currently used for both males and females. The 90-degree push-up differs from the revised push-up in that the down position requires the participant to lowers the body to the ground until the arms are at a 90-degree angle, while the down position for the revised push-up requires the participant to lower the body until the chest and thighs touch the ground. In the past, the bent-knee push-up was used as a measure of arm and shoulder girdle strength and endurance for females. Several researchers have found low interscorer objectivity and stability reliability for 90-degree push-up scores. The revised push-up test has been found to produce many zero scores. The purpose of this study is to determine the objectivity, reliability, and validity for the bent-knee push-up test for college-aged females and to determine the relationship between the revised push-up and bent-knee push-up tests. Ninety-three college-aged females participated in this study. The bent-knee push-up test was administered to all the participants the first day. Two raters were used to determine interscorer objectivity for approximately half of the participants. On the second day, the participants were split into two groups. One group was administered the bent-knee push-up test again to determine stability reliability of the scores. The other group was administered the revised push-up test. On the third day of testing, all participants were administered the bench press test using 40% of their body weight. This score was used to determine the criterion validity of both forms of the push-up tests. The interscorer objectivity coefficient for the bent-knee push-up was .997. A stability reliability coefficient of .83 was obtained**.** The correlation between the bent-knee push-

up and revised push-up scores was .75. The correlation between the bent-knee push-up scores and the bench press scores was .67. The correlation between the revised push-up scores and the bench press scores was .68. This is sufficient evidence that the interpretations of arm and shoulder girdle strength and endurance based on bent-knee push-up scores are equally as valid as interpretations based on the revised push-up scores for college-aged females. Based on these findings, the bent-knee push-up test could be used as a test of arm and shoulder girdle strength and endurance for females to ensure proper measurement of the females with low strength.

KEY WORDS: Arm and Shoulder Girdle Muscular Strength and Endurance, Push-up Test, Bent-Knee Push-Up, Revised Push-Up

Introduction

Push-ups are used as a measure of arm and shoulder girdle strength and endurance. Several different forms of push-ups are used in physical fitness tests. In the past, push-ups for males were executed with the toes and hands on the ground (full-body push-up), while females performed modified versions, including those with the knees and hands on the ground (bent-knee push-up). These modified versions were used since females tend to be weaker than males. Presently, however, fitness tests like FITNESSGRAM (Cooper Institute, 2001) and President's Challenge (2001) have a 90-degree push-up test for both males and females in which the toes and hands are on the ground while the participant lowers the body to the ground until the arms are at a 90-degree angle.

Many fitness tests include versions of the push-up test for females that appear to be easier and more discriminating among levels of ability for females than the male push-up version. For example, in the past the Physical Fitness Index test required females to perform push-ups with the feet on the floor and the hands on a bench 13 inches high, touching the chest to the bench (Clarke, 1967, Mathews, 1978). The California Physical Performance test in 1962 (Clarke, 1967), the Indiana Physical Fitness test (Clarke, 1967), and the Division for Girls' and Women's Sports test (Mathews, 1978) were several tests that included the bent-knee push-up for females. Safrit and Wood (1995) suggest using the bent-knee push-up for female examinees, stating that the full-body push-up may not be a discriminating measure for some groups since some examinees cannot execute even one. Tritschler (2000) suggests that girls and women perform a bent-knee push-up, while boys and men perform a full-body push-up. The University of Massachusetts (2002) also

includes a bent-knee push-up for female participants in the school's health and fitness assessments.

Recently, some programs like FITNESSGRAM (Cooper Institute, 2001) and President's Challenge (2001) require the 90-degree push-up for both males and females. Based on the FITNESSGRAM standards, most females performing this push-up could complete at least one. In the FITNESSGRAM, intraclass reliability coefficients are reported for push-up scores of females ranging from 0.64 to 0.96 (Cooper Institute, 2001). However, the reliability coefficient of 0.96 was obtained using bent-knee push-up scores. McManis, Baumgartner, and Wuest (2000) found low objectivity and reliability for both males and females in the ages from elementary school to college for 90-degree push-up scores. They also found that it was difficult to discern correct and incorrect forms with this push-up, and low strength college-aged females had difficulties performing the 90-degree push-up correctly. Based on these findings, they state that a modified version of the 90-degree push-up may be needed for adequate measurement of arm and shoulder-girdle strength and endurance of low strength individuals and females.

In light of these findings, Baumgartner, Oh, Chung, and Hales (2002) developed a revised push-up test, defining the down position as the body from the chest to the knees contacts the floor. So, their push-up is a full-body push-up. They found good interscorer objectivity coefficients (.75, .88, .97, .95, .98, .99) and good stability reliability coefficients (.90, .93, .95, .95). Strong validity evidence was also found with correlations between revised push-up and bench press scores of .80 for females and .87 for males. Despite this strong evidence, many females had scores of zero on this revised push-up test. The developers of the FITNESSGRAM (Cooper Institute, 2001) found that, though

the push-up scores for college-aged women had good reliability, the push-ups were done with the knees on the floor. The developers also found that 5% of both boys and girls over 8 years of age and 10% of both boys and girls ages 6-8 years completed zero 90-degree push-ups. These zero scores on the revised push-up and 90-degree push-up may indicate that the push-up test on the hands and toes is too difficult for females to perform.

Due to the difficulty of the 90-degree push-up and modified push-up tests for some participants, especially females, the tests may lack the discrimination needed in a fitness test. The bent-knee push-up may be easier than the full-body push-up for females to perform. McManis, et al. (2000) found that low strength females had difficulties performing the 90-degree push-up correctly, but were more successful performing bent-knee push-ups as an exercise in their class.

Objectivity, reliability, and validity must be acceptable for a test to be considered a good test. Baumgartner, Jackson, Mahar, and Rowe (2003) state that in order to have validity, a test must have reliability and to have reliability, a test must first have objectivity. Validity exists if the interpretation of the test scores is correct (Baumgartner, Strong, Hensley, 2002). Validity evidence can be obtained by the criterion approach which is determining the correlation between test scores and scores for a criterion or standard measure of the attribute being studied (Baumgartner, et al., 2003). Baumgartner, et al. (2002) define reliability as the consistency of test scores, and state that a test has objectivity if the scores are not dependent on who administered the test. At least two scores for each person being tested must be gathered to provide evidence of reliability or objectivity. These two scores can be collected either from two different scorers, trials in one day, or trials over days (Baumgartner, et al., 2003).

For the measurement of arm and shoulder girdle strength and endurance, a standard measure or criterion is the bench press. Baumgartner, Oh, Chung, and Hales (2002) used the bench press as the criterion score in determining the validity for the revised push-up test. Jackson, Fromme, Plitt, and Mercer (1994) and Pate, Burgess, Woods, Ross, and Baumgartner (1993) also used the bench press to show evidence of criterion validity for the push-up. Since the push-up requires moving a person's body weight up and down, the bench press should also require moving a percentage of a person's body weight. Baumgartner, et al. (2002) used 70% of body weight for males and 40% of body weight for females for the bench press test. Each person executed as many repetitions as possible for the bench press.

The purpose of this study is to determine the objectivity, reliability, and validity of the bent-knee push-up for college-aged females and to determine the relationship between the revised push-up and bent-knee push-up.

Participants

Data was collected from 87 female college-aged students who were enrolled in weight training classes. They were accustomed to executing push-up and bench presses.

Procedures

During the initial meeting with the participants, all of the participants signed an informed consent form, were told the purpose of the study, and were familiarized with tests they would be performing. The participants were tested a total of three days. The first day, all participants were weighed and performed bent-knee push-ups. The bent-knee push-up required that the participant lie with the hands under the shoulders, elbows spread slightly and knees close together, touching the ground. The feet were raised in the

air. With the body straight, the participant extended the arms fully so that her weight was resting on the hands and knees; this was the up position. Then the arms were bent, lowering the body to the ground so that the chest touched the ground; this was the down position. The participant then repeated these steps in a constant pattern until fatigued. One push-up was counted when the participant started at the up position, went to the down position, and then returned to the up position. The score was the number of push-ups executed correctly before stopping or before body position was changed (Clarke, 1967). Two raters, trained in the methods to be used, tested the participants for objectivity purposes. Both raters independently scored the number of push-ups performed correctly. The two raters were used in only half of the classes. These classes were randomly chosen.

On the second day, the participants were assigned to one of two groups. One group performed the bent-knee push-up again while the other group performed the revised push-up (Baumgartner, Oh, Chung, and Hales, 2002). The group performing the bent-knee push-up again was the same group scored by two raters on the previous testing day. The revised push-up required the person being tested to lie face down on the floor with the hands placed under the shoulders, fingers pointed forward and elbows pointed backwards along the sides of the body. The person pushed up to full arm extension; this was the up or starting position. Then, keeping the body straight, the participant lowered herself until all of the body from the chest to the thighs touched the floor. The participant then pushed up to full arm extension, back to the starting position. These down and up steps counted as one push-up. The person being tested continued these steps at a

comfortable rate until fatigued. The score was the number of push-ups executed correctly before stopping or before body position was changed.

During the third day of testing, all the participants performed a bench press. The bench press test was lifting 40% of the body weight (Baumgartner, et al., 2002). The participant's hands were placed on the bar approximately shoulder length apart. The participant pushed the bar up to full arm extension. This was the starting position. The bar was then lowered with a controlled motion to the chest. One press was counted once the bar was raised back up to the starting position. The number of full presses with correct form was recorded. An attempt to make up for any missed testing days was made after the first three days of testing were completed.

Data Analysis

Objectivity for the bent-knee push-up test scores recorded by the two raters was estimated for a criterion score which was from one rater using an intraclass correlation coefficient (R) based on a one-way ANOVA model as presented by Baumgartner, et al. (2003). Reliability for the bent-knee push-up test scores was estimated for a criterion score which was from one day using an intraclass correlation coefficient (R) based on a one-way ANOVA model as presented by Baumgartner et al. (2003). For validity, a Pearson Correlation coefficient was calculated to determine the relationship between the bent-knee push-up and bench press scores. Also, a Pearson Correlation coefficient was calculated to determine the relationship between the revised push-up and bench press and between bent-knee push-up and revised push-up scores.

Results

There were a total of 87 participants in the study. Descriptive information for the scores on the bent-knee push-up, revised push-up, and bench press are presented in Table 1. The score of zero occurred once for the bent-knee push-up and twice for the revised push-up.

There were 37 females who completed bent-knee push-ups on the first day when two raters where present. See Table 1 for descriptive information concerning the scores for raters A and B. The intraclass correlation coefficient for one rater was .997. This correlation suggests high objectivity between raters. Rater A and rater B had a mean score of 21. The score for a person varied only by one push-up when variations occurred between the two raters.

Thirty-two participants had bent-knee push-ups scores for both day 1 and day 2. The intraclass correlation coefficient for one day was .83. This is an acceptable value for the stability reliability evidence (Baumgartner, et al., 2003). The mean scores for day 1 and day 2 were 21 and 23, respectively. The same rater tested the participants both days.

Seventy-seven females performed the bent-knee push-up and bench press. For descriptive information on this group, see Table 1. A scatter gram for the bent-knee push-up and bench press scores was plotted and the relationship between the two variables was linear with no obvious outliers. The correlation between these two scores was .67. This is a moderate correlation for criterion validity evidence for the bent-knee push-up test (Baumgartner, et al., 2003). The mean score for the bent-knee push-ups was 21. The mean score for the bench press was 14. The same rater administered both tests.

The Pearson Correlation coefficient between the revised push-up and bench press score was .68. This value is moderate criterion validity for the revised push-up. Forty-one participants completed both the revised push-up and bench press tests. Table 1 also lists the descriptives for the scores in this group. The mean scores for the revised push-up and bench press tests were 9 and 13, respectively.

The correlation between the bent-knee push-up and revised push-up scores was .75. There is a moderate relationship between the two versions of the push-up. Forty-five participants had complete scores for both of these tests. The mean of the scores for the bent-knee push-up and revised push-up was 21 and 9, respectively.

Discussion

The objectivity coefficient (.997) for the bent-knee push-up scores was very good. Baumgartner, et al. (2003) state that the interscorer objectivity coefficient should be at least .80. The bent-knee push-up appears to be easily administered by different raters without the raters having different scores. The protocol for the bent-knee push-up test seems to be simple enough for the raters and the participants to follow easily.

The stability reliability coefficient obtained (.83) was moderate for the bent-knee push-up scores. Baumgartner, et al. (2003) suggest that the stability reliability coefficient should be at least .80. Baumgartner, Oh, Chung, and Hales (2002) found higher stability reliability than we did for the revised push-up. The moderate stability reliability suggests that the scores may change slightly between days for each participant even if the scorer remains the same.

The criterion validity evidence, a coefficient correlation between the bent-knee push-up scores and the bench press scores, of .67 is moderate. The correlation between

the revised push-up scores and bench press scores was .68. These similar findings suggest that interpretations of arm and shoulder girdle strength and endurance based on the bent-knee push-up test scores may be equally as valid as interpretations based on the revised push-up test scores when used on college-aged women. Baumgartner, et al. (2002) found a higher correlation (.80) than we did between the revised push-up scores and bench press scores for females.

Few zeros were found in this study. One zero was recorded for the bent-knee push-up, while two zeros were recorded for the revised push-up. Still, the number of zero scores is so small that both tests seem appropriate for use in fitness testing for college-aged females. The participant who scored the zero on the bent-knee push-up was also one of the two participants to score a zero on the revised push-up. The bent-knee push-up test seems to be slightly easier than the revised push-up test since, on the average, more bent-knee push-up than revised push-ups were executed.

In conclusion, the interscorer objectivity for the bent-knee push-up scores in this study is very good. The stability reliability of these scores is acceptable. The validity obtained in this study for the bent-knee push-up test is acceptable. Since the validity coefficient for both the bent-knee and the revised push-up are similar, either test can be used. Thus, either test rather than the 90-degree push-up should be used in fitness testing for college-aged females. However, one test should be chosen at the time of testing and used on all females in the group tested. The use of the bent-knee push-up rather than the revised push-up would allow for better measurement of the lower strength females.

References

Baumgartner, T.A., Jackson, A.S., Mahar, M.T, & Rowe, D.A. (2003). Measurement for evaluation in physical education and exercise science (7th ed.). New York, NY: McGraw-Hill.

Baumgartner, T.A., Strong, C.H., & Hensley, L.D. (2002). Conducting and reading research in health and human performance (3rd ed.). New York, NY: McGraw-Hill.

Baumgartner, T.A., Oh, S., Chung, H., & Hales, D. (2002). Objectivity, reliability, and validity for a revised push-up test protocol. *Measurement in Physical Education and Exercise Science*, 6(4), 225-242.

Clark, H. H. (1967). Application of measurement to health and physical education (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Cooper Institute for Aerobics Research. (2001). *The FITNESSGRAM Reference guide*. Retrieved May 25, 2003 from http://www.cooperinst.org/shopping/Web Format/FITNESSGRAM Reference Guide TOC.htm.

Jackson, A.W., Fromme, C., Plitt, H., & Mercer, J. (1994). Reliability and validity of a 1-minute push-up test for young adults. *Research Quarterly for Exercise and Sport*, 65, A-57 (Abstract).

Mathews, D. K. (1978). Measurement in physical education (5th ed.). Philadelphia, PA: W.B. Saunders Company.

McManis, B.G., Baumgartner, T.A, & Wuest, D.A. (2000). Objectivity and reliability of the 90-degree push-up test. *Measurement in Physical Education and Exercise Science*, 4(1), 57-67.

Pate, R.R., Burgess, M.L., Woods, J.A., Ross, J.G., & Baumgartner, T. (1993). Validity of field tests of upper body strength. *Research Quarterly for Exercise and Sport*, 64, 17-24.

President's Council on Fitness and Sport (2001). President's Challenge, Physical Activity and Fitness Awards Program. Bloomington, IN: President's Council on Fitness and Sport.

Safrit, M.J., & Wood, T.M. (1995). Introduction to measurement in physical education and exercise science (3rd ed.). St. Louis, MO: Mosby.

Tritschler, K. (2000). Barrow & McGee's Practical measurement and assessment (5th ed.). Baltimore, MD: Lippincott, Williams & Wilkins.

University of Massachusetts (2002). Health/Fitness assessment tutorial: Muscular strength and endurance. Boston, MA: University of Massachusetts.

Table 1

Descriptive Statistics for the Data of the Participants in the Study

Variable	n	Mean	Standard Deviation	Minimum	Maximum
Bent-knee Overall	77	21	9	0	46
Bent-knee Rater A	37	21	11	2	41
Bent-knee Rater B	37	21	10	2	40
Revised Push-Up	45	9	7	0	31
Bench Press	77	14	9	1	45

CHAPTER 4

SUMMARY AND CONCLUSIONS

There were a total of 87 participants in the study. The ages ranged from 18 to 25 with a mean age of 20. The minimum and maximum scores for the bent-knee push-up were 0 and 46, respectively, with a mean of 21. The minimum and maximum scores for the revised push-up were 0 and 31, respectively, with a mean of 9. The scores for the bench press ranged from 1 to 45 with a mean score of 14. The score of zero occurred once for the bent-knee push-up and twice for the revised push-up. Still, the number of zero scores is so small that both tests seem appropriate for use in fitness testing for college-aged females. The participant who scored the zero on the bent-knee push-up was also one of the two participants to score a zero on the revised push-up. The bent-knee push-up test seems to be slightly easier than the revised push-up test since, on the average, more bent-knee push-up were executed than were revised push-ups.

There were 37 females who completed bent-knee push-ups on the first day when two raters where present. The intraclass correlation coefficient for one rater was .997. This correlation score suggests high objectivity between raters. Rater A and rater B had a mean of 21. The scores for each person varied only by one push-up when variations occurred between the two raters. Baumgartner, Jackson, Mahar, and Rowe (2003) state that the interscorer objectivity coefficient should be at least .80. The bent-knee push-up appears to be easily administered by different raters without the raters having different

scores. The protocol for the bent-knee push-up test seems to be simple enough for the raters and the participants to follow without any difficulty.

Thirty-two participants had bent-knee push-up scores for both day 1 and day 2 bent-knee push-ups. Of the 37 females tested on day one by the two raters, five did not show up to be tested on the second day. The intraclass correlation coefficient for one day was .83. This is an acceptable value for the stability reliability evidence. The mean scores for day 1 and day 2 were 21 and 23, respectively. The same rater administered and scored the bent-knee push-up both days. Baumgartner, et al. (2003) suggests that the stability reliability coefficient should be at least .80. Baumgartner, Oh, Chung, and Hales (2002) found higher stability reliability than we did for the revised push-up (.90, .93, .95, .95). The moderate stability reliability suggests that the scores may change slightly between days for each participant even if the rater remains the same. On the average, participants performed better the second day of testing than they did on the first day. This improvement could be due again to motivation or to a slight learning effect.

Seventy-seven females performed the bent-knee push-up and bench press. Of the 87 participants who originally performed the bent-knee push-up on the first day of testing, 10 did not show up on the day the bench press test was administered. The correlation between the bent-knee push-up and bench press scores was .67. This is an acceptable correlation for criterion validity evidence for the bent-knee push-up scores (Baumgartner, et al., 2003). The same rater administered and scored both tests.

Forty-one participants completed both the revised push-up and the bench press. Of the 45 participants who performed the revised push-up, four were not present on the day the bench press test was administered. The Pearson Correlation coefficient between

the revised push-up and bench press was .68. This value shows acceptable criterion validity for the revised push-up. The mean scores for the revised push-up and bench press were 9 and 13, respectively.

The criterion validity coefficient for the bent-knee push-up test and the revised push-up test was .67 and .68, respectively. These similar correlations suggest that interpretations of arm and shoulder girdle strength and endurance based on the bent-knee push-up scores may be equally as valid as interpretations based on the revised push-up scores for college aged women. The means for the bent-knee push-up scores and the revised push-up scores were almost at equal distances from the mean for the bench press scores, yet in opposite directions. The mean for the bent-knee scores was greater than the mean for the bench press scores, while the mean for the revised push-up scores was less than the mean for the bench press scores. Baumgartner, et al. (2002) found a higher correlation than we did (.80) between the revised push-up scores and bench press scores for females.

The correlation coefficient between the bent-knee push-up and revised push-up scores was .75. There is a moderate relationship between the two versions of the push-up. Forty-five participants had complete scores for both of these tests. The mean of the scores for the bent-knee push-up and revised push-up were 21 and 9, respectively.

In conclusion, the interscorer objectivity for the bent-knee push-up scores in this study is very good for college-aged females. The stability reliability of these scores is acceptable. The validity obtained in this study for the bent-knee push-up test is acceptable. Since the validity coefficients for both the bent-knee and the revised push-up are basically the same, either test can be used. Females appear to be capable of

performing both types of push-ups with minimal zero scores. Thus, either test rather than the 90-degree push-up should be used in fitness testing for college-aged females. However, one test should be chosen at the time of testing and used on all females in the group tested. The use of the bent-knee push-up rather than the revised push-up would allow for better measurement of the lower strength females, while the revised push-up would be an easier measure of the higher strength females.

REFERENCES

Baumgartner, T.A., Jackson, A.S., Mahar, M.T, & Rowe, D.A. (2003). Measurement for evaluation in physical education and exercise science (7th ed.). New York, NY: McGraw-Hill.

Baumgartner, T.A., Strong, C.H., & Hensley, L.D. (2002). Conducting and reading research in health and human performance (3rd ed.). New York, NY: McGraw-Hill.

Baumgartner, T.A., Oh, S., Chung, H., & Hales, D. (2002). Objectivity, reliability, and validity for a revised push-up test protocol. *Measurement in Physical Education and Exercise Science*, 6(4), 225-242.

Clark, H. H. (1967). Application of measurement to health and physical education (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Cooper Institute for Aerobics Research. (2001). *The FITNESSGRAM Reference guide*. Retrieved May 25, 2003 from http://www.cooperinst.org/shopping/Web Format/FITNESSGRAM Reference Guide TOC.htm.

Jackson, A.W., Fromme, C., Plitt, H., & Mercer, J. (1994). Reliability and validity of a 1-minute push-up test for young adults. *Research Quarterly for Exercise and Sport*, 65, A-57 (Abstract).

Mathews, D. K. (1978). Measurement in physical education (5th ed.). Philadelphia, PA: W.B. Saunders Company.

McManis, B.G., Baumgartner, T.A, & Wuest, D.A. (2000). Objectivity and reliability of the 90-degree push-up test. *Measurement in Physical Education and Exercise Science*, 4(1), 57-67.

Miller, D.K. (2002). Measurement By the Physical Educator: Why and How (4th ed.). New York, NY: McGraw-Hill.

Pate, R.R., Burgess, M.L., Woods, J.A., Ross, J.G., & Baumgartner, T. (1993). Validity of field tests of upper body strength. *Research Quarterly for Exercise and Sport*, 64, 17-24.

Payne, N., Gledhill, N., Katzmarzyk, P.T., Jamnik, V., & Ferguson, S. (2000). Health implications of musculoskeletal fitness. *Canadian Journal of Applied Physiology*, 25(2), 114-126.

President's Council on Fitness and Sport (2001). President's Challenge, Physical Activity and Fitness Awards Program. Bloomington, IN: President's Council on Fitness and Sport.

Safrit, M.J., & Wood, T.M. (1995). Introduction to measurement in physical education and exercise science (3rd ed.). St. Louis, MO: Mosby.

Tremblay, A., & Chiasson, L. (2002). Physical Fitness in Young College Men and Women. *Canadian Journal of Applied Physiology*, 27(5), 563-575.

Tritschler, K. (2000). Barrow & McGee's Practical measurement and assessment (5th ed.). Baltimore, MD: Lippincott, Williams & Wilkins.

University of Massachusetts (2002). Health/Fitness assessment tutorial: Muscular strength and endurance. Boston, MA: University of Massachusetts.