RELATIONSHIP BETWEEN PERFORMANCE TESTS AND YARDS PER CARRY AVERAGE IN COLLEGIATE RUNNING BACKS

A Thesis

Presented to

the Faculty of the College of Science

Morehead State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Jason Kyle Newsom

April 24, 2017

Accepted by the faculty of the College of Science, Morehead State University, in partial fulfillment of the requirements for the Master of Arts degree.

____
Manual Michael Probst, Ed. D.
Director of Thesis

Master's Committee: ____________________________, Chair
Dr. Manual Michael Probst

____
Dr. Dana Seelig

____
Dr. Gina Blunt

____
Date

RELATIONSHIP BETWEEN PERFORMANCE TESTS AND YARDS PER CARRY AVERAGE IN COLLEGIATE RUNNING BACKS

Jason Kyle Newsom
Morehead State University, 2017

Director of Thesis: ____
Manuel Michael Probst, Ed. D.

The purpose of the study was to determine the relationship between a forty-yard sprint test, twenty-yard shuttle, vertical jump test and yards per carry average in collegiate running backs. Archival data of, 40-yard sprint times, twenty-yard shuttle times, vertical jump measurements and yards per carry on male football running backs between the ages of 18-24 for the 2010 season from 4 universities (1 NCAA-1, 1 NCAA division 1-AA and 2 NCAA division 2) was requested from the respective strength and conditioning coaches. Yards per carry were taken from the official statistics reported by the respective teams over the entire 2010 season. In order to prevent any way of identifying a subject, all personal information other than age, weight, performance test results and yards per carry was removed prior to data access. Because of the relatively small number of subjects from each school (6, 6, 3 and 5), a Kruskal-Wallis

nonparametric ANOVA was initially used to determine differences in the performance test results across the different schools in order to determine if it could be possible to aggregate the data. No significant differences were found among the data from the four schools and the data was subsequently aggregated for average yards per carry, 40-yard sprint times, 20-yard shuttle times and vertical leap ($p > 0.05$). Aggregate data for each independent variable (40-yard sprint times, 20 yard shuttle and vertical leap), and the independent variable, average yards per carry, were entered into a step-wise regression to determine significant relationships. An alpha-level $p< .05$ was used to determine significance. The correlation with average yards per carry and 40- yard sprint times was weak, negative and not significant ($r = -0.11$, $p \geq 0.05$), the correlation for average yards per carry for the twenty-yard shuttle was moderately strong, negative and significant ($r = -0.43$, $p < 0.05$) and the correlation for average yards per carry and vertical leap was positive, moderately strong, and also significant ($r = 0.49$, $p < 0.05$). A stepwise multiple- correlation of average yards per carry and the independent variables showed a significant, strong positive correlation ($r = 0.72$, $p < 0.05$, table 3). The regression model with all three predictors produced $R^2 = 0.52$, $F(3, 21) = 7.59$, $p < .05$. Independently the results showed weak relationships but when combine showed a stronger relationship to the independent variable. The findings as well as literature reviewed in preparation for the current study suggest that there is a need to explore ways of predicting on field success, such as yards per carry average.

Accepted by: ________________________________, Chair
Dr. Manual Michael Probst

____
Dr. Dana Seelig

____
Dr. Gina Blunt

Table of Contents

CHAPTER 1

Introduction

The ability to predict success in football is of major concern at the high school, college and professional levels. It is assumed that by testing football players, coaches can determine who will be more successful on the field. If this is true, these test results can allow coaches to more accurately predict on-the-field performance thus, leading to better overall team success. In addition, predicting performance may impact recruiting, players' salaries, predictions of team success and off-season training programs. Successful and accurate recruiting is one of the most important aspects to a successful college football program (Daus, Wilson, & Freeman, 1989). Accurate recruiting can greatly assist the coach regarding a player's position and rank as well as enhance the player's performance (Daus, Wilson, & Freeman, 1989).

Teams with top ten recruiting classes will have better seasons in the future than those that did not have a highly touted recruiting class (Langlett, 2003). If these tests do not predict performance on the field, then coaches may be wasting valuable time and money and costing the team wins by playing athletes that simply perform well on tests rather than perform well in the game.

Coaches and recruiters assume that test performance predicts on the field performance; however, the relationship may only be that these tests can predict how a coach perceives a football player's ability. If coaches are wrong about how well these tests predict on the field performance, they may be under the incorrect assumption that one player is better than another and may inadvertently give more playing time to the wrong players. This could result in a misuse of time and resources in developing high levels of performance on these tests, leading to overall poorer team performance. In addition, if these tests do not predict on-the-field performance then, off-season programs built around improving scores on these performance tests are inherently flawed. If these assumptions are wrong, coaches could be wasting a tremendous amount of money and time on athletes based on faulty information. Players with better test scores usually end up with statistically better seasons, this could be due to the fact that a coach assumes they are the better football player but this is merely an assumption. This assumption could leave the best football players on the side-line, while players with better performance tests are allowed to play (Barker et al., 1993).

Components of success typically revolve around the ability to run with the ball, block or being able to get to a player with the ball. One measure of success is the average yards per carry of a running back. Running with a football involves explosive speed, power and agility, attributes that are vital to success in football (Hoffman, et al., 2004). Because football is an anaerobic sport, typical tests to measure these attributes include the 40-yard sprint for speed, the vertical leap for power and the shuttle run for agility (Sawyer, Ostarello, Suess, & Dempsey, 2000). Although these tests are typically performed at the beginning and end of the season of many of sports, and whose data is often supposedly used to recruit football players from every position, including the NFL combines, there is conflicting data to support the apparent connection between performance on these speed and power tests and on- field success of running backs.

Therefore, the purpose of this study was to determine the relationship between a forty-yard sprint test, twenty-yard shuttle, vertical jump test and yards per carry average in collegiate running backs.

Independent variables: 40-yard sprint, 20-yard shuttle, vertical jump

Dependent variable: Yards per carry average

Hypotheses

H_o: The null hypothesis is that there will be no statistically significant relationship between the dependent variable, average yards per carry, and the independent variables, performance on the 45-yard sprint, vertical leap and 20-yard agility shuttle.

H_a: The alternate or research hypothesis of this study is that there will be a statistically significant relationship between the dependent variable and the independent variables.

Limitations of the Study

There are a number of limitations. The subject pool will be small. The schools ran different types of offenses some pass heavy and some were primary run teams.

The schools played in different conferences and different levels (Division I and Division II). Relying on archival data also limits the study because the data was collected by many individuals. The tests were also run on different surfaces. The players also may not have played the same amount of games. It is also assumed that the players gave their best effort in all the tests and the coaches performed the assessments properly.

Operational Definitions

Average yards per carry: Calculated by dividing total yards gained by the number of carries a player has throughout the season.

Carries: A football term describing anytime the running back is handed the ball in the back field.

Forty Yard Sprint: A test in which the participant starts from a four-point stance and sprints forty yards. Time begins on the first movement of the runner and stops when the participant reaches the forty-yard mark.

Twenty Yard Shuttle: A test in which the participant starts in a three point stance mid way between two lines that are ten yards apart, the participant sprints to touch one line then turns and touches the other line then turns and finished through the starting point. The time begins on the runner’s first movement and stops when the runner passes through the starting point after having touched both lines.

Delimitations of the study

The subject pool was comprised of 25 collegiate running backs between the ages of 18 and 24. Their average height was 5’9 with a range from 5’5- 6’1 and their average weight was 193 lbs with a range from 180-230.

Significance of Study

Results of this study may provide very useful information to football programs. Establishing a relationship between performance testing and on the field performance can take some of the guess work out of recruiting the “right” player for a team, it can help strength and conditioning coaches to develop more appropriate programs for their athletes, it can help coaches better predict player on the field performance. Conversely, if there is no relationship between performance testing and on the field performance, then strength coaches can stop spending time teaching their athletes how to perform well on these tests. This study will help everyone responsible for the success of a football team to put players on the field that will produce more wins.

CHAPTER 2

Review of Literature

Strength, size, power, and speed are fundamental to success in football (Hoffman, Cooper, Wendell, & Kang, 2004). Football is basically an anaerobic sport in that the energy needed to compete at all positions involves a rapid production of ATP from anaerobic energy sources, namely the phosphogen and glycolytic systems (Pincervo & Bompa, 1997). These systems provide levels of ATP production at a fast enough rate to allow the athlete to perform with speed, strength and power while playing the game. Since most football plays last on average less than 6 seconds (Rhea, Hunter, & Hunter, 2006), the ATP needed must be provided rapidly, which is what the anaerobic bioenergetics systems are designed to do. The abilities of these systems to provide that energy is typically measured by performance testing such as strength/power tests, speed tests, jumping tests or agility tests to measure the ability to rapidly change direction and therefore, only last a few seconds. On the surface, it would appear that all football players' physical performances should be measured using the same tests because they all need speed, strength, power and agility. These tests should also be reliable and possess external validity to some actual measure of on-field performance. For instance, 40-yard sprint times should correlate with some measure of performance in running backs, such as yards-per-carry average. Performance on these tests can therefore also lead to the perception of an athlete's abilities, playing time, division play, recruiting and monetary earnings.

Tests such as the bench press, forty-yard sprint, power clean, vertical jump and agility shuttle have been used to measure strength, power and speed in football players (Burke, Winslow, & Strube, 1980) and are conducted in order to predict player performance (yards per carry average, touchdowns, tackles). The tests go a long way in establishing player playing time (Barker et. al., 1993). Coaches assume there is a positive relationship between test results and on-the-field production, an assumption that shapes how they will recruit, train, and rank their athletes (Barker et. al., 1993). In a broader sense, performance testing to predict performance or even draft status is often used, albeit, with mixed results**.**

Sawyer, Ostarello, Suess, and Dempsey (2000) used a Division I football team as participants ($n = 40$) to investigate the relationship between performance testing and a player's perceived football playing ability. Prior to the season, researchers measured the participant's height and weight and administered the following performance tests: vertical jump, 10 and 20-yard sprints, the 20-yard shuttle, squat, power clean, Olympic snatch and bench press. After two thirds of the season, players were independently ranked by two offensive and two defensive coaches for their football playing ability on a scale from one to nineteen. The coaching staff's rankings was averaged and the scores were used as the players' final ranking. The coaches' rankings were based solely on their perceptions of the players' football playing abilities. Using a Spearman's rank-order correlation, significant correlations ($p < 0.05$) were found in both the bench press ($r = -0.48$) power cleans ($r = -0.58$) for the defensive group, but not in the offensive group (bench $r = -0.03$ and power clean $r = - 0.35$) for rankings and performance.

The running backs group showed significant ($p < 0.05$) correlations between rankings with the 20-yard sprint ($r = 0.63$) and the 20-yard shuttle ($r = 0.74$). The highest correlating test was the vertical jump ($r = -0.5$) for offensive players and ($r = -0.64$) for defensive players (Sawyer et al., 2000). This study appeared to show some relationship between performance tests and a player's ranking, but it did not indicate that this leads to playing a more productive player in the game.

Barker et al. (1993) investigated the relationship between performance tests and a player's ranking on the team. Fifty-nine Division IAA football players were asked to perform the following tests: body composition (using skin fold measurements at seven sites on the body), five yard sprint, ten yard sprint, vertical jump, 1.5 mile run, 300-yard shuttle, and a personality survey (Catell 16 PF Questionnaire). Players were ranked based on their perceived athletic ability regardless of their position. The offensive and defensive starters and second-string players were ranked independently by the offensive coordinator, defensive coordinator, and strength and conditioning coach, respectively; then, the entire coaching staff ranked the entire team. The players were ranked from 1 to 42 based on the coaches' perception of the player's football playing ability. The coaches' rankings were averaged. A Spearman's rank-order correlation was used to establish the extent of the relationship between each variable and the coach's ranking. Body composition ($r_s = 0.65$) and vertical jump ($r_s = -0.72$) were the only statistically significant indicators of a players rank ($p < 0.05$). Ten-yard sprint time ($r_s = 0.71$) and 5-yard sprint time ($r_s = 0.69$) were both significant predictors of a players ranking. The researchers suggested that vertical jump (r_s $= -0.72$ and other measures of power should be emphasized as the main way of evaluating football playing ability.

Burke, Winslow, and Strube (1980) analyzed the relationship between performance tests and a player's classification in a Division I football team through an eight-week spring training program. The coaching staff put the players into the categories of starter, player, and non-player based on the coach's perception of the player's football ability. The athletes ($n = 67$) were then tested before and after the spring training program using six criteria: lean body mass from skin fold measurements, bench press, squat, 40-yard dash, one mile run and skin fold measurements. A Chi-Square test for correct or incorrect placement was used to predict a player's classification. The research showed that these criteria could significantly predict a player's classification ($p < 0.003$). The battery of tests correctly predicted the right category for 58.5% of the players. Forty-yard sprint ($f = 1.00$) ($df = 4$) (Chi Square test = 15.95) was one of the strongest contributors to player classification, leading the authors to the conclusion that the best predictor of success in football is speed (Burke et al., 1980).

Predicting how well an athlete will do on various performance tests may also be important to the football coaches. Davis, et al. (2004) compared anthropometric measures of 46 Division I football players to various performance tests to find possible relationships. Anthropometric measurements included height, weight and hamstring length. The performance measurements included, hang clean, bench press, 40-yard sprint, 20-yard shuttle, and vertical jump. Positive correlations were found for body weight and 40-yard sprint times ($r = 0.018$, $p < 0.0001$).

Ranking a player during the recruiting process can sometimes determine which division recruits them. The divisions are Division I, Division II and Division III. One would assume that better players would play at the highest level (Division I). Fry & Kraemer (1991) gathered data on performance tests to explore the relationship between the tests and the division and classification of players. Information from strength coaches was gathered on each team's bench press, squat, power clean, 40- yard sprint and vertical jump. Six Division I, seven Division II and six Division III schools responded, with a total of 776 athletes. A Spearman-Rho was used to measure the correlation between the performance tests and division played. Division I athletes performed significantly ($p < 0.05$) better in all tests than the other division's squat ($rs = 0.18$), clean ($rs = 0.23$), 40-yard ($rs = 0.12$), vertical jump ($rs = 0.24$) than players in the other two divisions. Of particular interest to the current study was the running backs group. Division I and Division II backs had significantly lower forty-yard sprint times ($rs = 0.15$) and the bench press ($rs = 0.18$). There was a significant difference between Division I and Division III running back in the vertical jump and power clean ($rs = 0.15$ and $rs = 0.24$, respectfully, $p < 0.05$) The sample showed no significant difference between any of the divisions in the back squat ($rs = 0.28$, $p > 0.05$) among the running back group. The authors suggested their results showed that coaches can use performance tests, especially the bench press, 40-yard sprint, and power clean, to predict at which division level a player might play.

In an effort to explore the relationship between performance tests and whether a player plays Division 1 or Division II football, Garstecki, Latin and Cuppet (2004) measured relative bench press (bench/body weight), relative squat (squat/ body weight), relative power clean (power clean/ body weight), vertical leap and body composition on players from Division 1 schools (n = 26 schools n = 112 players) and Division II schools (n = 23 schools and n = 152 players). The data showed that mean vertical jump for Division 1 players was 80.1 $\pm$ 10.2 cm and 70.1 $\pm$ 12.1 cm for Division II. The mean 40-yard sprint time for Division 1 players was 4.74 $\pm$ 0.3 seconds and for Division 2 players was 4.88 $\pm$0.3, both vertical jump and 40-yard sprint times were significantly better for the Division 1 players ($p \leq 0.01$). Garstecki, Latin and Cuppet (2004) also found significant differences ($p \leq 0.01$) between the two levels within the positions groups, running backs. Division 1 running backs showed significantly higher vertical jump (87.4 $\pm$ 7.0 cm) than division II running backs (77.8 $\pm$ 12.1 cm) as well as faster 40-yard sprint times (4.48 $\pm$ 0.1 seconds and 4.59 $\pm$ 0.2 seconds, respectively, $p < 0.05$). These differences may have been due to better training facilities and lower percentage of body fat in the Division I players (Garstecki, Latin, & Cuppet, 2004).

In an effort to understand how players are evaluated in the recruiting process (Ghigiarelli, 2011) gathered archival data on 2560 high school seniors who were highly recruited out of high school between the years of 2001 and 2009. The measurements used in this study included: height, weight, 40-yard sprint, 20-yard shuttle, vertical jump and broad jump. The recruits were split into groups based on their star ranking; a ranking scale of 1-5 where 5 stars being the best or highest ranking: 2 Star (n = 191), 3 Star (n = 1057), 4 Star (n = 625) and 5 Star (n = 142).

To analyze the data the researchers created two groups, five and four star recruits were paired together in a group referred to as highly recruited and three and two star recruits were paired into a group labeled recruited. A 2-way ANOVA was used to test for differences between the groups and their performance on the tests. The Highly Recruited group was significantly heavier, faster and jumped higher than the Recruited group ($p < 0.05$). The running backs group mean 40-yard sprint time was 4.56 ± 0.17 seconds and for the recruited group the mean for the 40-yard sprint test was 4.60 ± 0.16 seconds. The difference did not reach statistical significance ($p > 0.05$). The mean 20-yard shuttle time for the running backs in the highly recruited group was 4.31 ± 0.26 seconds and for the recruited group the mean was 4.37 ± 0.20 seconds, also not achieving statistical significance ($p > 0\ .05$). The mean vertical jump test in meters for the highly recruited running backs was 0.80 ± 0.09 and the recruited group mean was 0.82 ± 0.07. There was no significant between the two groups of running backs within the variables used in the current study ($p > 0.05$). The researchers concluded the information shows that college recruiters are more likely to recruit bigger and faster athletes (Ghigiarelli, 2011).

Rating football players on performance tests is convenient because they are easily quantifiable. The University of Nebraska was the first college program to develop a system to score players on performance tests, the Performance Index, as early as the 1970s (Rigoni, 2013). The physical tests include lean body mass, hang clean, back squat, bench press, 10-yard dash, 20-yard shuttle, and vertical jump. The researcher wanted to compare the results of the physical tests to a measure of toughness to see which one better predicted success of a player. In this study, success was defined by asking the football coaching staff to categorize the 47 subjects into two categories, players that contributed during the season and players who did not contribute. The researchers defined toughness as the ability to handle physically and mentally stressful situations. In order to measure toughness the researcher measured cortisol levels via an oral swab pre and post training sessions. A lower increase in cortisol levels indicate a better tolerance for stressful situation and in this study was defined as toughness. All testing was done prior to the season and directly after the season the coaches were asked to classify the 47 subjects into two groups contributors and non-contributors. An ANOVA was used to compare the two groups' cortisol reactivity. In the non-contributor group the cortisol increased 66 $\pm$ 0.91% and the contributors experienced an increase in cortisol levels of 24 $\pm$ 58%. The results almost reached statistical significance (F (1, 45) = 3.099, p = 0.085) suggesting that there may be a way to determine who could be a contributor based on cortisol reactivity. A nested model comparison using r2-change F-test was used to determine if using the physical tests with the cortisol levels would lead to better predictability of classification. The physical test (r2 =.28, F (7, 39) = 2.209, p=0.54) was weaker than the model that included physical tests and cortisol reactivity (r2 = 0.39, F (8, 38) =2.970, p = 0.11). Players who had better cortisol reactivity and had better vertical jump scores were more likely to be characterized as contributors (Rigoni, 2013).

Presumably the most stringent evaluation process for a football player is the NFL Combine, which involves testing athletes on the 225 lb. bench press, 40-yard dash with 10-yard and 20-yard splits, vertical jump, pro-agility shuttle, broad jump and 3-cone drill. An athlete's abilities based on Combine performance is supposed to influence draft status, however, these relationships have shown mixed results. McGee and Burkett (2003) analyzed data from the NFL Combines class of 2000 (n=326) to determine if draft status was related to performance on the NFL Combines. Independent variables included in the analysis varied depending on position and included 225 lb. bench press, 40-yard dash with 10 yard and 20-yard splits, 20-yard shuttle pro-agility run, 60-yard shuttle, 3-cone drill, vertical jump, standing broad jump, with draft order as the dependent variable. The results of the study showed quite a bit of variability with performance and draft status depending on the player's position. For instance, the highest correlation was for running backs, wide receivers, and defensive backs ($r^2 = 1.0$) with quarterbacks r^2=0.84, offensive lineman r=0.70 and defensive lineman r^2=0.59 with linebackers showing the lowest correlation r^2=0.22. The authors suggest the reason for such a strong relationship for the running back, wide receiver and defensive backs group was that they are positions that most rely on speed and agility, with the most significant predictors of success in these positions being 3-cone drill, height, weight, 10-yard dash and vertical jump. In addition, a repeated measures ANOVA found a significant difference in draft status between the first 2 round picks compared to rounds 6 and 7 and their performances on the different combinations of dependent variables used in the regression analysis. The main point of the study is that not all performance tests are weighted the same for all positions and that performance on the 40-yard and 20-yard agility run were not strong predictors of draft status. This study only analyzed players who were drafted.

Sierer, Battaglini, Mihalik, Shields & Tomasini, (2008) examined performances on the NFL Combine between drafted and non-drafted players categorized as “skilled players” (wide receivers, cornerbacks, free safeties, strong safeties, and running backs), “big skilled players” (fullbacks, linebackers, tight ends and defensive ends) and “lineman”. Using a student's t-test, significant differences were found between drafted and non-drafted skilled players in all player categories and tests ($p < 0.01$). Performance on the usual Combine performance tests were analyzed using independent t-tests between the players drafted and those not drafted. The results showed "skilled players" drafted players had significantly better 40-yard dash times, better 3-cone drill times, better pro-agility shuttle and vertical leap ($p < 0.01$). All drafted players performed better on the 40-yard dash, 3-cone drill.

Robbins (2010) analyzed draft order by player position and NFL Combine test results for the classes 2005 to 2009 (n=1155) using raw, normalized and ratio data and found low predictability to draft order. Specifically, for the running back group, predictability for the 40-yard sprint time was $r = 0.45$, for the vertical jump was $r = 0.47$ and the 20 yard shuttle was $r = 0.02$ for all sets $p < 0.05$. Among all groups the best predictors of draft order for all groups was the 40-yard sprint and vertical jump. The author suggests that other performance tests should be analyzed in order to evaluate players due to the low correlation between draft order and the physical tests.

In an attempt to determine a relationship between on-field performance and NFL Combine performance data, Kuzmitz and Adams (2008) analyzed archival data from 1999 to 2004. The Combine data included in the analysis were the 10, 20 and 40 yards dashes, bench press, vertical jump, broad jump, 20 and 60-yard shuttles, 3-cone drill as well as the Wonderlic Personnel Test (a test of intelligence). A number of positions were analyzed using combine performance data, draft order, first 3 years salary and games played during the first three years. A significant ($p < 0.05$) correlation was found only for a quarterback's draft order and vertical leap ($r = -0.36$) and broad jump ($r=-0.32$). However, this meant that higher vertical jumps were associated with lower draft pick orders and longer broad jumps were related to earlier draft picks. Only 3 of 80 correlations achieved statistical significance. The authors concluded that the Combine tests “fail to show a consistent significant relationship with measures of success for quarterbacks in areas of draft order, quarterback rating, games played and salary”. For wide receivers there were significant correlations for times for the 10-yard, 20-yard split and 40-yard sprint ($r=0.22$, $r=0.24$, respectively, $p< 0.05$), but 40-yard sprint is only correlated to 3rd year salary ($r=0.28$). This shows slower sprint times are actually correlated to higher 3rd year salaries. Only 4 of 80 possible correlations achieved statistical significance. The authors concluded that the receivers group could not be reliably evaluated by the Combine tests. The running backs group draft order did correlate well to the sprint times for 40-yard sprint ($r = 0.31$), 20-yard sprint ($r = .20$) and 10-yard sprint ($r = 0.28$). Variables with a significant relationship ($p < 0.05$) with yards per carry average were: 40-yard sprint ($r= -0.34$) in year one and ($r = -0.27$) in year two, 20-yard sprint ($r = -0.32$) in year 2, 10-yard sprint ($r = -0.37$) in year 1 and ($r = -0.32$) in year 2. Yards-per-carry average did not have a strong relationship with vertical jump ($r = -0.18$) in the first year, ($r = 0.07$) in the second year, and ($r = 0.17$) in the third year. The twenty-yard shuttle also did not meet the threshold of significance ($r = 0.21$) in the first year, ($r = -0.17$) in the second year, and ($r = -0.04$) in the third year. Sprint times for the 40-yard, 10-yard and 20-yard splits were highly correlated with each other, calling into the question the need to obtain all of these values. Nevertheless, only 7 of 70 possible correlations showed statistical significance ($p < 0.05$). The author’s interpretation of the data suggests that, with the possible exception in running backs, combined test performances are not very useful when trying to predict general success in football.

The studies reviewed for the current study revealed my things about what we know about measuring athletic ability but also revealed how little the transfer of athletic ability to the playing field has actually been studied.

Summary

The ability to predict success in football is of utmost importance in classifying, testing, training and recruiting football players for certain positions. Forty-yard sprint times, twenty-yard shuttles and vertical jump test appear to have some validity in predicting draft status, ranking on a team and classification as a starter or non-starter. Vertical jump appears to be the strongest predictor across the literature, however, literature is not clear on how each individual criterion or predictor or which combinations predicts actual on the field performance. Therefore, it becomes very important to determine which criteria or which combinations of criteria are the best predictors of on the field success so that a more efficient and accurate way can be found. Coaches and recruiters need to know if there truly is a transfer from the tests that measure the speed and power of their athletes, to the playing field in the. Determining the relationship between performance tests and game performance will help coaches make better player selections and engage in more efficient recruiting. There needs to be more research to determine how the performance tests transfer from the training programs to on the field performance.

This carry over from training to the playing field is so important because, training adaptations are specific to the type of activity, the volume, and intensity of the exercise program. The transfer from training to competition only occurs when the skills that are practiced are actually those that are performed in the game (Rhea, Hunter, and Hunter 2006).

CHAPTER 3

Methods

Subjects

Archival data of average height and weight, 40-yard sprint times, twenty-yard shuttle times, vertical jump measurements and yards per carry on male football running backs between the ages of 18-24 for the 2010 season from 4 universities (1 NCAA-1, 1 NCAA division 1-AA and 2 NCAA division 2) was requested from the respective strength and conditioning coaches. The average height of the subjects was 69 inches (65-73 inches) and the average weight was 193.1 lbs (180-230 lbs).

Procedures

40-yard sprint times, twenty-yard shuttle times and vertical jump data were collected at the end of summer prior to the 2010 season during 1 day by the respective teams' strength coaches. Sprint times were recorded using hand-held stopwatches and recorded in seconds. Vertical jump was measured using a Vertec Jump Measurement System®. Yards per carry were taken from the official statistics reported by the respective teams and averaged over the entire 2010 season. In order to prevent any way of identifying a subject, all personal information other than age, weight, performance test results and yards per carry was removed prior to data access.

Statistical Analysis

Because of the relatively small number of subjects from each school (6, 6, 3 and 5), a Kruskal-Wallis nonparametric ANOVA was initially used to determine differences in the performance test results across the different schools in order to determine if it could be possible to aggregate the data. No significant differences were found among the data from the four schools and the data was subsequently aggregated for average yards per carry, 40-yard sprint times, 20-yard shuttle times and vertical leap ($p > 0.05$). Descriptive data for the performance tests is presented as means $\pm$ SD (Table 1). Aggregate data for each independent variable (40-yard sprint times, 20 yard shuttle and vertical leap), and the independent variable, average yards per carry, were entered into a step-wise regression to determine significant relationships. An alpha-level $p < .05$ was used to determine significance.

Table 1. Descriptive data of the performance variables

Variables	Means	SD	N
Average yards per carry	4.83	1.35	25
40-yard sprint (sec)	4.67	0.16	25
20-yard shuttle (sec)	4.58	0.16	25
Vertical leap (inches)	32.08	3.33	25

CHAPTER 4

Results

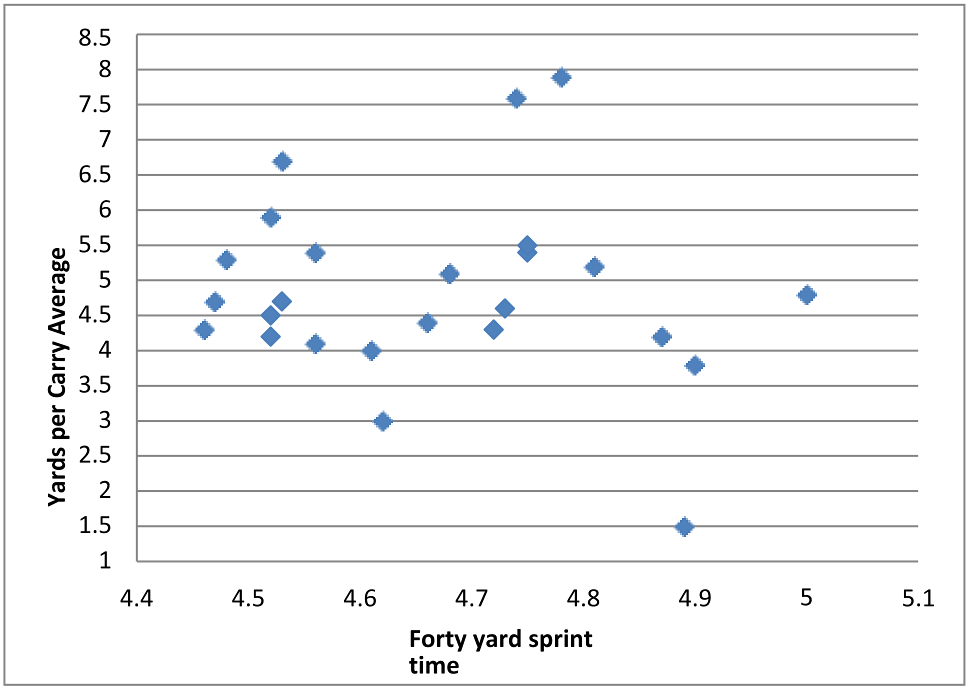
The results of the Pearson product-moment correlation analysis between the independent variables 40-yard sprint time, vertical jump and 20-yard shuttle run, and the dependent variable average yards per carry (Table 2). The correlation with average yards per carry and 40-yard sprint times was weak, negative and not significant ($r = -0.11$, $p \geq 0.05$), the correlation for average yards per carry for the twenty-yard shuttle was moderately strong, negative and significant ($r = -0.43$, $p < 0.05$) and the correlation for average yards per carry and vertical leap was positive, moderately strong, and also significant ($r = 0.49$, $p < 0.05$). Graphs 1 - 3 show the scatter plots of each independent variable and average yards per carry. A stepwise multiple-correlation of average yards per carry and the independent variables showed a significant, strong positive correlation ($r = 0.72$, $p < 0.05$, table 3). The regression model with all three predictors produced $R^2 = 0.52$, $F(3, 21) = 7.59$, $p < .05$.

Table 2.

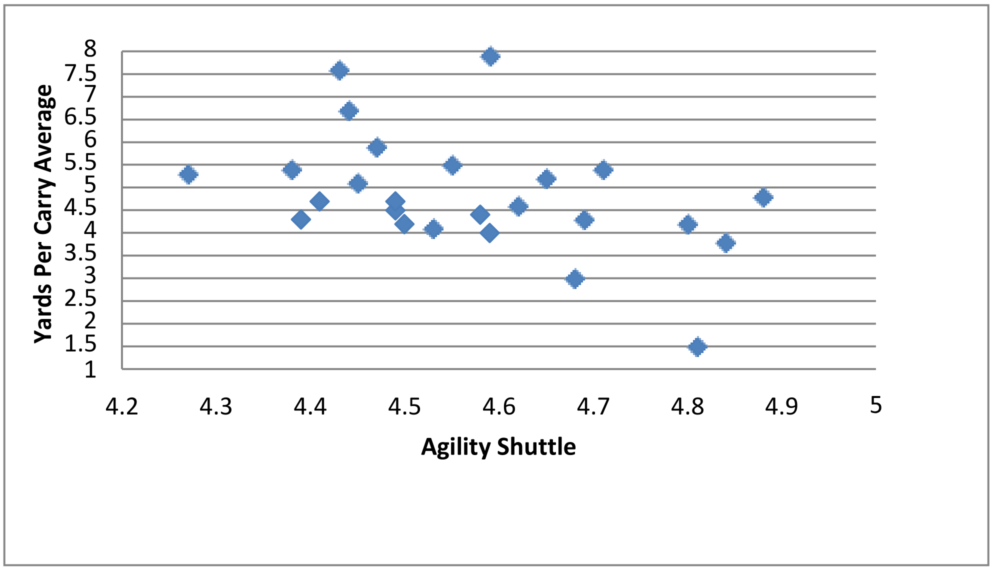
Correlation and Significance of Independent Variables and Dependent Variable

N = 25	Yards per Carry
Pearson Correlation 40 yards	- 0.11
Agility Shuttle Vertical Jump	- 0.43
	0.49
40 yards Agility	0.31
Shuttle Vertical Jump	0.02
	0.01

Graph 1. Correlation of Average Yards per Carry and 40-Yard Sprint Time



Graph 2. Correlation of Average Yards per Carry and 20-Yard Shuttle Times



Graph 3. Correlation of Average Yards per Carry and Vertical Jump

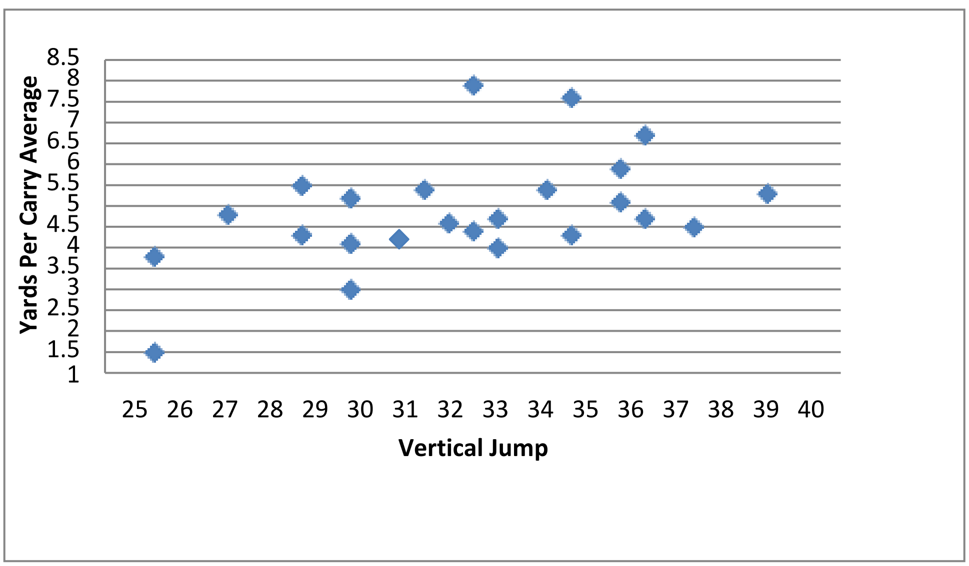


Table 3. Correlations for the Independent Variables and Average Yards per Carry

Model	R	R^2	Adjusted R^2	SEM	Df
1	0.49[a]	0.24	0.20	1.18	1
2	0.61	0.38	0.32	1.09	1
3	0.72	0.52	0.45	0.98	1

a. Predictors: (constant), vertical leap

b. Predictors: (constant), vertical leap, 40-yard sprint

c. Predictors: (constant), vertical leap, 40-yard sprint, yard agility shuttle

Table 3 shows a significant correlation for vertical leap and average yards carried. The level of significance increased as 40-yard sprint and agility shuttle were added to the model (table 4).

Table 4. Analysis of Variance Summary

Model		SS	Df	MS	F	Sig
1	Regression Residual Total	9.90 31.83 41.72	1 23 24	9.90 1.38	7.15	0.01
2	Regression Residual Total	15.66 26.06 41.72	2 22	7.83 1.12	6.61	0.00
3	Regression Residual Total	21.57 20.15 41.72	3 21 24	7.20 0.96	7.50	0.00

1 Predictors: (constant) vertical jump

2 Predictors: (constant) vertical jump, 40-yard sprint

3 Predictors: (constant), vertical jump, 40-yard sprint, agility shuttle

Table 5. Table of Coefficients for Combined Independent Variables and Average Yards per Carry

	B	Std. Error	Beta	T	Significance
Constant Vertical	-8.058	13.925		-0.579	0.569
Leap 40-yard	0.192	0.109	0.485	1.175	0.093
Sprint Agility	8.748	2.517	1.1016	3.475	0.002
Shuttle	-7.464	3.008	3.008	-2.482	0.022

Dependent Variable: Yards per Carry

The prediction equation of the dependent variable from the independent variables from table 5 is as follows: Y = -8.058 + .192 (X1) + 8.748 (X2) - 7.464 (X3), where X1 is the vertical leap in inches, X2 is the 40-yard sprint times in seconds and X3 is the agility shuttle in seconds.

CHAPTER 5

Conclusion

The results of the analysis seem to show that none of the performance measures were an effective tool for predicting yards per carry average. Vertical jump was the best single predictor of average yards per carry. This could indicate that the most important physical attribute of those evaluated in this study for running backs would be lower body power. The next strongest correlation alone was the 20-yard shuttle. This could indicate that the second most important attribute in running backs could be agility. The weakest correlation out of the three performance tests was the 40-yard sprint. This would seem to show that straight-line speed is not as important as the other two variables for producing yards per carry for a running back. As we added variables to the model, the correlations got stronger, to the point where when all three of the variables were added the model could predict 52% of the change in yards per carry average. This could explain why the NFL scouting combine uses multiple tests for every position to predict on the field performance (Robbins, 2012). As variables are added, the level of prediction should be more consistent. This study raises an interesting question: If three of the most common performance test only account for half of yards per carry average how can we predict the other half more accurately.

Even at the highest level of football, the NFL uses testing data inconsistently at best even the data that they gathered from an event they created in order to gather that data (Robbins, 2010). There is a major need for more research on this topic.

As shown in Barker et al. (1993) personality and performance tests do in fact affect a coaches perception of an athlete’s football ability without taking into account any real on the field production. The present study attempted to establish a link between performance tests and actual game performance as measured by yards-per-carry average in running backs. While these performance measures do have a weak relationship to yards per carry average there are likely other variable that are hard to measure, the ability to read defenses, ability to use blockers and patience to let plays develop. Finding a way to measure these attributes will give coaches even more tools to use in deciding who will play and who to recruit.

The relationship between performance test and actual on the field measures of success must be further explored in order to justify using these tests to set depths charts. It is only through research that coaches can verify the specificity of these tests. Verifying the test's validity will ensure that a coach is accurately assessing a player's ability in the preseason as well as improving a football player's ability in the off- season. Until these tests are validated to transfer to on the field success, coaches are merely assuming these tests lead to on the field success. In fact, if these performance tests do not transfer to measurable on the field performance then coaches may be playing and recruiting athletes that are good at tests but not the best football players. Further research is needed to identify ways of measuring intelligence and understanding of the game of football and other factors and their relationship between test results and on the field performance.

Other factors that contribute to on-field success such as psychological traits, intelligence and understanding of the game have been largely left out of most studies but were included in Barker's (1993) study. The researchers used a personality test to see if it was correlated to a players ranking on a team and it was found that there was a significant relationship ($p<0.05$) between a player's classification (starter, non starter) and their personality test (Barker et al., 1993). This highlights the need for further research into alternative ways predict on the field success. Further research is also needed to verify that traditional tests such as the ones used in this study actually transfer to on the field measurable success at every position. For instance: does a 40-yard sprint time lead to more tackles for defensive players or more catches for receivers, does a vertical jump translate to better blocking by an offensive line men? Does a 20-yard shuttle relate to sacks for a defensive end? Research verifying or disputing these performance tests transfers to on the field success will change everything from evaluating recruits, off-season strength and conditioning programs and depth chart decisions.

Coaches have been taught that performance tests indicate on the field success. Whether this is actually true or not has not been questioned thoroughly, because coaches do what was done in the past, assuming it is the most effective. An athlete that scores high on specific performance tests will end the season with a better statistical performance than an athlete that tested worse. The reason for that may be the coach's perception will make the coach play the athlete with the better test scores (Barker et al., 1993). Research has shown a positive correlation between performance tests and players' ranking on particular teams. The motor skills tested are those typically associated with being a good football player are tested in these performance tests such as, strength, power and speed (Barker et al., 1993).

Theoretically performing well on these tests should transfer onto the field. In other words, it is assumed that a player who performs well on the field should also demonstrate high levels of performance on the various performance tests. Strength and conditioning coaches use this assumption to develop programs with the thought that by improving a player’s performance on a performance test, the player will demonstrate better performance on the field (Sawyer, Ostarello, Suess, & Dempsey, 2000). One problem these causes, especially at the NFL combine. Programs designed to improve test scores bring every ones test scores closer together thereby making the test invalid and the tests are not leading to on the field success anyway (Kuzmits, 2008) Sawyer et al. (2000) and many other researchers have compared the results of the performance tests to the coaches’ ranking of players, based on their perceived playing ability. Other factors that may contribute to on field success such as psychological traits, intelligence and understanding of the game.

Future studies

To improve the accuracy of the results, future researchers should gather the data rather than using archival data. As always a bigger pool of subject will always lead to a better picture of the total population. All the tests should be run on the same surface if at all possible. The biggest thing that will help further answer the question underlying all of this which is, how do coaches get the best players on the field, is add more variables. Traditional tests such as power clean squat, broad jump. To take it a step further the researcher could begin using intelligence measurements such as grade point average, SAT scores, football knowledge surveys. The key is to tie these tests to on the field measurements because on the field performance is what actually matters.

What This Means for Coaches

The relationship between performance tests and actual on the field measures of success must be further explored in order to justify using those performance tests to set depths charts and making recruiting decisions. It is only through research that coaches can verify the specificity of these tests. Verifying the test's validity will ensure a coach is accurately assessing a player's ability in the preseason as well as improving a football player's ability in the off-season. Until these tests are validated to transfer to on-the-field success, coaches are merely assuming these tests lead to success. In fact, if these performance tests do not transfer to measurable on-the-field performance, then coaches may be playing and recruiting athletes that are good at tests, but not the best football players. Further research is needed to identify ways of measuring intelligence and understanding of the game of football and its relationship between test results and on the field performance. This highlights the need for further research into alternative ways predict on the field success. For instance: does a 40- yard sprint time lead to more tackles for defensive players or more catches for receivers, does a vertical jump translate to better blocking by an offensive line men? Does a 20-yard shuttle relate to sacks for a defensive end? Research verifying or disputing these performance tests transfers to on-the-field success will change everything from evaluating recruits, off-season strength and conditioning programs and depth chart decisions. Coaches have been taught that performance tests indicate on-the-field success. Whether this is actually true or not has not been questioned thoroughly, because coaches do what was done in the past, assuming it is the most effective way. Research has shown a positive correlation between performance tests and players' ranking on particular teams. Theoretically, performing well on these tests should transfer onto the field. In other words, it is assumed a player who performs well on the field should also demonstrate high levels of performance on the various performance tests, something that is frequently done by coaches (Sawyer, Ostarello, Suess & Dempsey, 2002).

Further research into other variables that might be contributing factors such as the ability to read a defense, intelligence and aggression is needed. Position specific training and testing should also be studied. Positional training should reflect the demands that are placed on each separate position. For example a linebacker and wide receiver will have very different demands during competition. Position specific drills have been shown in studies to translate in better training outcomes (Salvo & Pigozzi, 1998). Drills that directly related to skills used during a volley ball match also revealed an elevated level of proficiency in those skills during competition over non-specific skill training (Gabbet 2008). Coaches should consider specificity before doing anything in their training program, because if it does not translate into on the field performance then it is a waste of time and energy. The idea of specificity must also be considered when deciding which tests to use to predict on the field performance. In determining if a performance test is worth the extra time and money its costs to administer the test a coach must consider the ability of the test to discriminate between players that will have a better game performance and those that will not perform well in a game (Burke et al., 1980).

Additional Study

In addition running back archival data from two NFL draft classes (24 players) were evaluated. Their ten-yard split twenty-time and their forty-yard sprint along with their twenty-yard shuttle and vertical jump test were gathered. From the NFL's combine result tracking website (http://nflcombineresults.com/nflcombinedata.php?). Their yards per carry average were then gathered from the NFL's website (http://nflcombineresults.com/nflcombinedata.php?). Analysis of the data showed that all of the independent variables were weakly correlated with the dependent variable, performance on the field (Table 6).

Table 6. Correlations for Average Yards per Carry and Performance Measures

Pearson Correlation	Yards per Carry	Significance
10 yard split (sec)	- 0.19	0.19
20 yard split (sec)	- 0.10	0.32
40 yard sprint (sec)	- 0.11	0.30
Vertical jump (inches)	0.03	0.44
20 yard shuttle	- 0.17	0.22

A stepwise multiple correlation was performed, which found that after all of the variables were entered into the model, the combined R was 0.26, with an adjusted R^2 of -0.19. This means that only 19% of the change in average yards carried can be explained by these performance factors, therefore the largest amount of variability in average yards carried by a running back during that year must be explained by other factors than what were collected during the combines. This small side study showed that a ten-yard time does correlate much better to the performance on the field for a running back. However, none of the tests had a statistically significant correlation. The findings seem to match those done by other researchers who found very little correlation between performance tests and selected success measures. Kuzmits and Adams, 2008 found that both 40 yard sprint times ($r = -0.34$) and 10 yard sprint times ($r = -0.37$) had significant relationships with yards per carry average in the first years of a running backs professional career. No other combine test reached statistical significance ($p < 0.05$) (Kuzmits, 2008). In preparing for this research there was very little found on the relationship between performance at the combine and on the field production which is confusing because of the availability of data and the enormous financial implications based on drafting players that will be successful on the field. Further study is needed to explore the relationship between combine testing and on the field production (Kumitz, 2008).

References

Barker, M., Wyatt, T., Johnson, T.L., Stone, M.H., O'Bryant, H.S., Poe, C., & Kent, M. (1993). Performance factors, psychological assessment, physical characteristics, and football playing ability. *Journal of Strength and Conditioning Research, 7*(4*)*, 224-233.

Burke, E.J., Winslow E., & Strube, M. (1980). Measures of body composition and performance in major college football players. *Journal of Sports Medicine, 20,* 173-180.

Daus, A.T., Wilson, J., & Freeman, W.M. (1989). Predicting success in football. *The Journal of Sports Medicine and Physical Fitness*, *29*(2) 209-212.

Davis, D., BARNETTE, B., Kiger, J., Mirasola, J., &Young, S. (2004). Physical characteristics that predict functional performance in division I college football players*. Journal of Strength and Conditioning Research, 18*(1), 115-120.

Fry, A., & Kraemer, W., (1991) Physical performance characteristics of American collegiate football players. *Journal of Applied Sports Science Research, 5*(3), 126-138.

Gabbett, T. (2008) Do skill-based conditioning games offer a specific training stimulus for junior elite volleyball players*? Journal of Strength & Conditioning Research, 22*(2) 509-517*, doi: 10.1519/JSC.0b013e3181634550.*

Garstecki, M., Latin, R., & Cuppett, M. (2004) Comparison of selected fitness and performance variables between NCAA division I and Division II football players. *Journal of Strength & Conditioning Research, 18*(2) 292-297.

Ghigiarelli, J. (2011) Combine performance descriptors and predictors of recruit ranking for the top high school football recruits from 2001 to 2009: Differences between Position Groups. *Journal of Strength & Conditioning Research, 25*(5), 1193- 1203.

Hoffman, J., Cooper, J., Wendell, M., & Kang, J. (2004). Comparison of Olympic vs. traditional power lifting training programs in football players. *Journal of Strength and Conditioning Research, 18*(1), 129-135.

Kuzmits, E. F. & Adams, A. J. (2008). The NFL combine: Does it predict performance in the national football league? *Journal of Strength and Conditioning Research, 22,* 1721-1727.

Langlett, G. (2003). The relationship between recruiting and team performance in Division 1A college football. *Journal of Sports Economics, 4*(3), 240-245.

Little, T., &Williams, A. (2005). Specificity of acceleration, maximum speed, and agility in professional soccer players. *Journal of Strength and Conditioning Research, 19*(1), 76-78.

McGee, K., & Burkett, L. (2003). The national football league combine: a reliable predictor of draft status? Journal of Strength and Conditioning Research, 17(1), 6-11.

Pincervo DM & Bompa TO (1997). A physiological review of American Football. Sports Medicine, 23(4):247-260.

Rhea, M., Hunter, R., and Hunter, T, (2006). Competition modeling of American football, observational data and implications for high school, collegiate, and professional player conditioning. *Journal of Strength and Conditioning Research, 20*(1), 58-61.

Robbins, D. (2012). Relationships between National Football League combine performance measures. *Journal of Strength & Conditioning Research, 26*(1), 226- 231.

Robbins, D. (2010). The National Football League (NFL) combine: does normalized data better predict performance in the NFL draft? *Journal of Strength and Conditioning Research*, 23(11), 2888-2899.

Rigoni, Joseph B. (2013). Toughness Predicts Performance in College Football" (2013). *Theses, Dissertations, and Student Research: Department of Psychology*. http://digitalcommons.unl.edu/psychdiss/51

Salvos, V., & Pigozzi, F. (1998). Physical training of football players based on their positional roles in the team: effects on performance related factors. *Journal of Sports Medicine and Physical Fitness, 38,* 294-297.

Sawyer, D.T., Ostarello, J.Z., Suess, E.A., & Dempsey, M. (2000). Relationship between football playing ability and selected performance measures. *Journal of Strength and Conditioning Research*, *16*, 611-616.

Sierer, S.P., Battaglini, C.L., Mihalik, J.P., Shields, E.W. & Tomasini, N.T. (2008). The national football league combine: performance differences between drafted and nondrafted players entering the 2004 and 2005 drafts. *Journal of Strength and Conditioning Research*, 22(1), 6-12.

Wilmore, J., & Costill, D. (2004). An introduction to exercise and sport physiology. In Bahrke, M., Garret, L., & Rhonda, J. (Eds.), *Physiology of Exercise and Sport third edition* (pp.19). Human Kinetics.

Young, W., McDowell, M., & Scarlett B. (2001). Specificity of sprint training and agility training methods. *Journal of Strength and Conditioning.*

Appendix A

40 yard	yards per c arry	agility shuttle	vertic al jump	player		
4.68	5.1	4.45	35.5	a	School 1	
4.73	4.6	4.62	32	b		
4.87	4.2	4.8	31	c		
4.74	7.6	4.43	34.5	d		
4.52	4.5	4.49	37	e		
4.62	3	4.68	30	f		
4.75	5.4	4.71	31.5	a	School 2	
4.56	5.4	4.38	34	b		
4.47	4.7	4.41	33	c		
4.53	4.7	4.49	36	d		
5	4.8	4.88	27.5	e		
4.81	5.2	4.65	30	e		
4.75	5.5	4.55	29	f		
4.78	7.9	4.59	32.5	g		
4.48	5.3	4.27	38.5	a	School 3	
4.9	3.8	4.84	26	b		
4.61	4	4.59	33	c		
4.46	4.3	4.39	34.5	d		
4.56	4.1	4.53	30	e		
4.52	4.2	4.5	31	f		
4.52	5.9	4.47	35.5	a	School 4	
4.53	6.7	4.44	36	b		
4.72	4.3	4.69	29	c		
4.66	4.4	4.58	32.5	d		
4.89	1.5	4.81	26	e		
					Average weight	193.0
					Average Height	5'9

Appendix B

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.033[a]	.001	-.044	.853
2	.168[b]	.028	-.064	.861
3	.183[c]	.033	-.112	.880
4	.256[d]	.066	-.131	.888
5	.257[e]	.066	-.193	.912

a. Predictors: (Constant), vertical jump

b. Predictors: (Constant), vertical jump , 20 ydshuttle

c. Predictors: (Constant), vertical jump , 20 ydshuttle, 40 yd

d. Predictors: (Constant), vertical jump , 20 ydshuttle, 40 yd, 10 yd

e. Predictors: (Constant), vertical jump , 20 ydshuttle, 40 yd, 10 yd , 20 yd

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.257[a]	.066	-.193	.912

a. Predictors: (Constant), 20 ydshuttle, 10 yd , 20 yd , vertical jump , 40 yd

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.062	5	.212	.256	.931[a]
	Residual	14.958	18	.831		
	Total	16.020	23			

a. Predictors: (Constant), 20 ydshuttle, 10 yd , 20 yd , vertical jump , 40 yd

b. Dependent Variable: yds per carry