Spring 2019

The Relationship Between The NFL Scouting Combine And Game Performance Over A Five Year Period

Jordan Cook
Georgia Southern University

THE RELATIONSHIP BETWEEN THE NFL SCOUTING COMBINE AND GAME PERFORMANCE OVER A FIVE YEAR PERIOD

by

JORDAN COOK

(Under the Direction of Greg Ryan)

ABSTRACT

The National Football League (NFL) Scouting Combine offers good face validity however, there has been doubt on the ability of the NFL Combine to predict future success of athletes in relation to in game performance (i.e., predictive validity). This study analyzed the NFL Combine data of 1537 college football players who participated in the Combine between 2013 and 2017 and their subsequent year's performance in the NFL. The measures used from the Combine were the six different measures of athletic performance; 40-yard dash, vertical jump (VJ), broad jump, shuttle run (PRO), 3-cone drill, and bench press (BP). The measure of NFL performance was average snaps played (avgS). AvgS was derived from total (offensive/defensive and special teams) snaps (TS), and games played (GP) for each position per season over the sample timeframe. Individual athletic performance measures were normalized via Z-scores for each event completed at the NFL Combine. Average Z-scores were calculated for every athlete when compared to all other athletes (avgCZ) and athletes who played the same position (avgPZ). Correlational analysis was used to ascertain whether the physical performance tests were associated with subsequent performance. A multiple linear regression (MLR) was performed to examine whether individual event Combine performance could predict the subsequent year's performance in the NFL. Of the 35 correlations found when examining relationships only two correlations were found to be moderately strong, avgCZ - avgS2 (r=0.320), avgPZ - avgS2 (r=0.332), whereas the majority were found to be weak (r<0.3). Furthermore, data analysis suggests that Combine measures can only explain approximately 2.6% of the variance in avgS one year following the Combine when using three (VJ, BP, and PRO) performance tests as predictors. The primary results of this study suggest that the NFL Combine lacks predictive ability when examining first year game performance. Furthermore, it also lacks

correlational strength when examining relationships between performance and subsequent five years performance in the NFL. Caution should be used if coaches, general managers, and other front office staff are considering the use of Combine data as a possible selection for the upcoming NFL Draft.

INDEX WORDS: National Football League, Scouting combine, Predictive analytics, Correlational analysis, Regression analysis, Sports performance, Physical performance testing, Data normalization, American football

THE RELATIONSHIP BETWEEN THE NFL SCOUTING COMBINE AND GAME PERFORMANCE OVER A FIVE YEAR PERIOD

by

JORDAN COOK

B.S., University of Maine Presque Isle, 2017

A Thesis Submitted to the Graduate Faculty of Georgia Southern University

in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

STATESBORO, GEORGIA

THE RELATIONSHIP BETWEEN THE NFL SCOUTING COMBINE AND GAME PERFORMANCE OVER A FIVE YEAR PERIOD

by

JORDAN COOK

.

Major Professor:	Greg Ryan
Committee:	Ronald Snarr
	Stephen Rossi

Electronic Version Approved:
May 2019

TABLE OF CONTENTS

LIST OF TABLES

Page

CHAPTER 1
INTRODUCTION

Every year the National Football League (NFL) invites the top prospects from collegiate football to showcase individual talents in a series of performance and psychological tests, as part of the NFL Combine (Combine). The purpose of the combine is to provide insight into the medical history, athletic abilities, psychological state, as well as demonstrations of skill in positional drills of potential prospects. In addition, the Combine is an avenue used by coaches, scouts, and general managers as additional criteria for potential selection in the upcoming NFL Draft.

The physical examination and medical history collected at the combine is used to grade the athlete on their ability to participate in the NFL. Although research is limited, the grade of athlete has been reported to offer insight into the probability of playing in the NFL with high, low, and fail grade athletes having a 58%, 55%, and 36% chance of playing at least one game in the NFL, respectively. Furthermore, higher grade athletes have an extended career (41.5 games) compared to those of a low (34.3 games) and failing grade (19 games) (Brophy et al., 2008). Athletes are also asked to complete a psychological test called the Wonderlic test. The Wonderlic is used to briefly assess the athlete's mental capability and intelligence (Dodrill & Warner, 1988; Matthews & Lassiter, 2007). For quarterbacks scoring well on this test could be advantageous even though research has suggested no relationship with draft status, a relationship has been reported with future performance in the NFL (i.e. games started, approximate value, NFL wins, NFL pass yards) (Pitts & Evans, 2018). The most publicized aspect of the combine are the performance tests. The physical performance tests are designed to assess athletic abilities and fitness associated with American football. The tests include the 40-yard dash (40yd), pro-agility shuttle (PRO), 225lb bench press repetition max test (BP), 3-cone drill (3C), vertical jump (VJ) and broad jump (BJ).

The important physical attributes of American football players has been well researched and indicates the importance of strength, speed, and power in the game (Berg, Latin & Baechle 1990; Black & Roundy, 1994; Fry & Kraemer, 1991; Garstecki, Latin & Cuppett 2004; Schmidt, 1999). These six

performance tests have been thoroughly researched in their ability to measure these athletic characteristics (Riebe, Ehrman, Liguori & Magal, 2017). Interestingly, research has suggested that these attributes can differentiate between: starters and non-starters (Fry & Kraemer, 1991); level of competition (i.e., division 1 vs division 2) (Garstecki, Latin & Cuppett, 2004); and draft status (i.e., drafted vs non-drafted) (Sierer, Battaglini, Mihalik, Shields & Tomasini, 2008). Therefore, since the physical characteristics being tested at the Combine appear to be in line with the attributes identified as important to American football performance it offers good face validity. However, there has been doubt on the ability of the NFL combine to predict future success of athletes within the NFL (i.e., predictive validity).

Work by Kuzmits and Adams (2008) examined the predictive validity of the combine for quarterbacks, wide receivers and running backs from 1999-2004. It was found that combine exercises were not correlated with NFL success for either quarterbacks or wide receivers. For running backs, strong correlations were observed for sprints times (40, 20, and 10 yard) and measures of success in the NFL. The measures of success used for running backs consisted of draft order, yearly salary, games played per season and average yards per carry. Furthermore, work by Teramoto, Cross and Willick (2016) similarly examined the predictive validity of the NFL Combine for running backs and wide receivers. For running backs they observed, the best predictor of performance (rushing yards per attempt) for the first three years and over the career of the athlete was the 10 yard sprint time. VJ was observed at being significantly associated with performance (receiving yards per reception) over the first three years and career in the NFL when examining wide receivers,. Similarly, McGee and Burkett (2003) examined the 2000 NFL Combine to determine if performance measures could predict draft status. A total of seven positions were analyzed; quarterback, wide receiver, running back, offensive line, defensive line, defensive backs, and linebackers using prediction equations. Results indicated and the study concluded that the combine accurately predict draft status for running backs (r^2= 1.00), wide receivers (r^2 =1.00), defensive backs (r^2=1.00). Furthermore, it can also be used for, offensive linemen (r^2=0.70), defensive linemen (r^2= 0.59), linebackers, (r^2= 0.22, quarterbacks, (r^2=0.84), but with less accuracy.

With previous research being limited in its scope by either number of years or number of positions examined (e.g. WR, QB, and RB) little scientific analysis into the totality of the NFL Combine is available. Furthermore, with all NFL Combine and NFL game data being publicly accessible on web-based domains it is interesting that greater amount of scientific analysis has not been published. The primary intent of this study was to investigate possible relationships between Combine performance and subsequent year(s) performance in the NFL for: Quarterbacks (QB), Wide Receivers (WR), Defensive Backs (DB), Defensive Lines\ (DL), Linebackers (LB), Offensive Line (OL), Running Backs (RB), and Tight Ends (TE) over a 5 year period from 2013-2017. Secondly, it was to examine if performance in individual events is predictive of first years performance in the NFL.

CHAPTER 2
METHODS

EXPERIMENTAL APPROACH TO THE PROBLEM

This study analyzed the NFL Combine data of college football players who participated in the Combine between 2013 and 2017 and their subsequent year's performance in the NFL. The measures used for analysis were the NFL combines six different measures of athletic performance as listed in Table 1: 40yd, VJ, BJ, 3C, PRO, and BP. Measures of NFL performance were total (offensive, defensive, and special teams) snaps (TS) and games played (GP) for each position over the sample timeframe. For example a player who attended the 2013 NFL Combine would have five years of performance data available in comparison a player who attended the 2017 Combine would only have single year of performance available. The number of active players for each respective combine over the five year period is listed in Table 2. Within each year TS was divided by GP to calculate average snap count (AVGS).

Table 1. Performance measures used, associated attributes and protocols used in the National Football League Combine as stated by the National Football League ("Workouts & Drills", 2019).

Test	Attribute	Description/Protocol
40-Yard Dash	Speed and Explosive Power	The athlete runs 40 yards as fast as possible. These athletes are timed at 10, 20 and 40-yard intervals.
Vertical Jump	Lower Body Explosive Power	The athlete's reach and highest jump are measured. The differential between the reach and the highest flag the athlete touches is his vertical jump measurement.
Broad Jump	Lower-body Explosion and Lower-body Strength.	The athlete explodes out as far as he can horizontally whilst remaining stationary on the landing.
Shuttle Run	Lateral Quickness and Explosive Power	The athlete starts in the three-point stance, explodes out 5 yards to his right, touches the line, goes back 10 yards to his left, left hand touches the line, pivot, and he turns 5 more yards and finishes.
3-Cone Drill	Change of direction	The athlete goes 5 yards to the first cone and back. Then, he turns, runs around the second cone, runs a weave around the third cone, changes directions, comes back around that second cone and finishes.

Bench Press	Upper-body Strength and Endurance	The athlete bench presses 225 pounds for as many reps as the athlete can get.

SUBJECTS

This research study included a total of 1568 subjects' data who attended the NFL Combine within a 5-year sample between 2013-2017. For the purpose of this study, positions were grouped together to form nine groups; QB, WR, DB, DL, LB, OL, RB, ST, and TE. A total of; 79 QB, 229 WR, 287 DB, 257 DL, 182 LB, 249 OL, 167 RB, 31 ST, and 87 TE NFL Combine data was examined. Due to the low sample size associated with ST group the position group was removed from the data set. For analysis, subject data was broken down by position in addition to being separated by year, as demonstrated in Table 3. Due to this study only requiring secondary analysis of data which is publicly available on web-based domains, which do not disclose individual's health information, Institutional Review Board approval was not required.

Table 2. Number of participants separated by position and year drafted from the National Football League Combine.

Position Group	**2013**	**2014**	**2015**	**2016**	**2017**	**Total**
QB	14	18	14	18	15	79
RB	37	33	35	29	33	167
TE	18	16	18	15	20	87
OL	52	44	53	53	47	249
DL	42	50	53	63	49	257
LB	34	34	39	39	36	182
DB	57	56	56	59	59	287
WR	35	50	45	43	56	229
Total	289	301	313	319	315	1537

Note: QB= quarterbacks, RB=running backs, TE= tight ends, OL= offensive linemen, DL= defensive linemen, LB = linebackers, DB= defensive backs, WR= wide receivers

DATA COLLECTION

The data obtained for this study was collected from pro-football-reference.com and nflcombineresults.com. Data from each NFL combine year was collected from www.nflcombineresults.com, snaps and games played was collected from www.pro-football-reference.com**.** First, athletes who attended and participated in the NFL Combine between 2013 and 2017 were identified using nflcombineresults.com. Once identified their combine data was then collected, including performance in the: 40yd, PRO, BP, VJ, BJ, and 3C. Secondly, annual career performance statistics of TS and GP was collected for each athletes entire NFL careers up to the 2017 season.

STATISTICAL ANALYSES

Initially, individual athletic performance measures were normalized via Z-scores for each event completed at the NFL combine. Average Z-score were calculated for every athlete when compared to all other athletes (avgCZ) and athletes who played the same position (avgPZ). This method of normalization of various data for analysis has been previously supported (Bunn, Ryan, Button, & Zhang, 2017). Average snaps played (avgS) was then calculated, by dividing TS by GP, for each season for each athlete. To then determine whether the six athletic performance measures were related to future performance in the NFL both avgCZ and avgPZ was correlated against avgS for each season (i.e., avgS1 = avgS one year out from the combine, avgS2 = avgS two years out from the combine). After examination it was found the data violated assumptions of normality thus Spearman's rho correlations were used for analyses. Next, a forward selection multiple linear regression (MLR) was performed to examine whether individual event Combine performance could predict the subsequent years performance in the NFL. Data was excluded pairwise to allow for use of cases that contain some missing data (i.e., individuals who didn't complete every event). Furthermore, additional forward selection MLR were performed to examine whether individual event Combine performance could predict the subsequent years performance in the NFL when categorized by position. Data was analyzed using SPSS (version 25.0, SPSS, Inc., Chicago, IL). An *a priori* 5% level of significance ($p < 0.05$) was used to ascertain statistically significant correlations.

Table 3. Number of participants categorized by years out from National Football League Combine.

	Years out from NFL Combine				
Position Group	**1 year**	**2 years**	**3 years**	**4 years**	**5 years**
QB	79	64	46	32	14
RB	167	134	105	70	37
TE	87	67	52	34	18
OL	249	202	149	96	52
DL	257	208	145	92	42
LB	182	146	107	68	34
DB	287	228	169	113	57
WR	229	173	130	85	35
Total	1537	1222	903	590	289

Note: NFL= National Football League, QB= quarterbacks, RB=running backs, TE= tight ends, OL= offensive linemen, DL= defensive linemen, LB = linebackers, DB= defensive backs, WR= wide receivers

CHAPTER 3
RESULTS

DESCRIPTIVE STATISTICS

Of the 1536 athletes whose data was collected a total of 1504 were used for the correlational and regression analysis. A total of 32 participants were excluded from analysis as they did not perform any of the 6 Combine measures, therefore an avgCZ or avgPZ could not be calculated. Of the 1504 athletes used for analysis 802 (53.5%) completed all 6 events. When broken down by event; 1426 (94.8%) completed the 40yd (4.78 $\pm$ 0.30 seconds), 1185 (78.8%) completed BP (20.00 $\pm$ 6.00 repetitions), 1309 (87.0%) completed VJ (32.55 $\pm$ 4.28 inches), 1300 (86.4%) completed BJ (115.40 $\pm$ 9.31inches), 1092 (72.6%) completed PRO (4.40 $\pm$ 0.26 seconds), and 1056 (70.2%) completed 3C (7.26 $\pm$ 0.41 seconds). Further breakdown of events completed by position group can be seen in Table 4. When examining snaps played, 1096 athletes went on to average at least 1 snap played in the NFL. When examining this more closely by separating the athletes by position group, 221 (78.6%) DBs, 198 (79.5%) DL, 136 (77.7%) LBs, 172 (69.6%) OL, 39 (51.3%) QBs, 116 (70.3%) RBs, 65 (76.5%) TEs, and 149 (65.9%) WRs went onto average at least 1 snap in the NFL.

COMBINE CORRELATION ANALYSIS

The results of the correlation analysis for avgCZ and avgPZ for the combine as a whole are presented in Table 5. A Spearman's rho correlation coefficient was calculated for the relationship between avgCZ, avgPZ and avgS1-5 for the combine as a whole. AvgCZ had significant, weak, positive correlations with avgS1 and avgS2. Furthermore, significant, weak, positive correlations were found between avgPZ and avgS1 and avgS3. Athletes who performed better at the combine on average played more snaps over the first two years of their career than their competitive peers.

Table 4. Number of participants per National Football League Combine event over a five year period from 2013-2017.

Position	40yd	BP	VJ	BJ	PRO	3C
	Count (Avg ± SD) (s)	Count (Avg ± SD) (reps)	Count (Avg ± SD) (in)	Count (Avg ± SD) (in)	Count (Avg ± SD) (s)	Count (Avg ± SD) (s)
DBs	265 (4.53 ± 0.10)	232 (15.06 ± 4.05)	240 (35.56 ± 2.72)	240 (121.98 ± 5.74)	188 (4.18 ± 0.14)	185 (6.95 ± 0.18)
DL	241 (4.97 ± 0.22)	211 (24.82 ± 5.03)	214 (30.96 ± 3.87)	211 (111.36 ± 8.54)	189 (4.55 ± 0.21)	188 (7.48 ± 0.34)
LBs	164 (4.73 ± 0.12)	144 (21.47 ± 4.05)	149 (33.61 ± 3.15)	157 (117.96 ± 5.96)	117 (4.33 ± 0.15)	109 (7.13 ± 0.22)
OL	235 (5.25 ± 0.18)	205 (25.28 ± 4.99)	210 (27.13 ± 3.05)	202 (102.49 ± 6.42)	200 (4.76 ± 0.19)	191 (7.83 ± 0.30)
QBs	73 (4.84 ± 0.16)		71 (30.83 ± 2.83)	71 (111.83 ± 6.58)	67 (4.73 ± 0.15)	63 (7.11 ± 0.21)
RBs	156 (4.60 ± 0.13)	141 (19.50 ± 4.92)	149 (34.05 ± 3.26)	146 (118.64 ± 5.78)	100 (4.30 ± 0.15)	99 (7.08 ± 0.21)
TEs	76 (4.76 ± 0.13)	72 (20.26 ± 4.32)	67 (32.93 ± 2.92)	66 (117.02 ± 5.92)	56 (4.40 ± 0.13)	54 (7.18 ± 0.20)
WRs	216 (4.52 ± 0.10)	180 (13.50 ± 4.04)	209 (34.82 ± 3.15)	207 (121.09 ± 5.88)	175 (4.23 ± 0.15)	167 (6.95 ± 0.20)

Note: Avg = average; SD = standard deviation; s= seconds; reps= repetitions; in= inches; 40yd = 40 yard dash; BP = bench press; VJ = vertical jump; BJ = broad jump; PRO = pro agility; 3C = 3-cone drill; DBs = defensive backs; DL= defensive linemen; LBs = linebackers; OL = offensive linemen; QBs = quarterbacks; RBs= running backs; TEs = tight ends; WRs = wide receivers; No QBs completed the BP over the 5 year period

Table 5. Spearman's rho correlation matrix of average positional and total combine Z-scores and average snaps played in the National Football League over a five year period.

	avgCZ	avgPZ	avgS1	avgS2	avgS3	avgS4	avgS5
avgCZ	1	0.793**	0.133**	0.103**	0.045	0.016	0.006
p-value		0.01	0.01	0.01	0.17	0.69	0.92
n		1504	1504	1202	899	587	289
avgPZ		1	0.148**	0.146**	0.132**	0.044	0.097
p-value			0.01	0.01	0.01	0.29	0.1
n			1504	1202	899	587	289

Note: avgCZ = average combine Z-score; avgPZ = average positional Z-score; avgS1 = average snaps played 1 year out from the combine; avgS2 = average snaps played 2 years out from the combine; avgS3 = average snaps played 3 years out from the combine; avgS4 = average snaps played 4 years out from the combine; avgS5 = average snaps played 5 years out from the combine
** = Correlation is significant at the 0.01 level

POSITIONAL CORRELATION ANALYSIS

The results of the correlation analysis for avgCZ and avgPZ for the combine as a whole are presented in Table 6. A Spearman's rho correlation coefficient was calculated for the relationship between avgCZ, avgPZ and avgS1-5 for each of the eight position groups. No significant relationships were observed for DL and QBs when examining avgCZ or avgPZ and the first 5 years performance.

DEFENSIVE BACKS

Table 6 presents the correlation analysis for DBs. Significant, weak, positive correlations between avgCZ and the first 4 years of performance were found. Similar relationships were found between avgPZ and performance measures where significant, weak, positive correlations were found in relationship to avgS1, avgS2, and avgS3. DBs that performed better at the combine had a greater snap average for the first 4 years of their career. Furthermore, those who performed better than other DBs had a greater snap average for the first 3 years of their career.

LINEBACKERS

Table 6 presents correlations between average Z-score (avgCZ, avgPZ) and avgS over a 5 year period for linebackers. AvgCZ had significant, weak, positive correlations with avgS1 and avgS3 as well as a significant, moderate, positive correlation with avgS2. AvgPZ demonstrated similar relationships,

with significant, weak, positive correlations to avgS1, avgS3 and significant, moderate, positive correlation to avgS2. The positive direction of these relationships indicates LBs who performed better at the combine had a greater snap average for the first 3 years of their career.

OFFENSIVE LINEMEN

The results of the correlational analysis for OL are presented in Table 6. Significant weak, positive correlations were observed between avgCZ and the first 3 years performance in the NFL. An additional significant weak, positive correlation was observed between avgPZ and second year performance (avgS2). OL who had a greater avgCZ played on average a great number of snaps for 3 years out from the combine.

RUNNING BACKS

Table 9 presents the results of the correlational analysis for RBs. Weak positive correlations between avgCZ and the first three years of performance were found. Similarly, again for avgPZ weak positive correlations were found for the first three years of performance indicating a significant relationship between the variables. RBs that had greater avgCZ and avgPZ scores played on average a greater number of snaps for the first three years of the career.

TIGHT ENDS

Table 6 presents correlations found between average Z-score (avgCZ, avgPZ) and avgS over a five year period for TEs. A weak positive correlation was found between avgCZ and avgS1. The positive direction of this relationships indicates TEs who performed better at the combine on average had a greater snap average the following season.

WIDE RECEIVERS

The results of the correlational analysis for WRs are presented in Table 11. Weak positive correlations were observed between avgCZ and the first 2 years performance in the NFL. Additional weak positive correlations were observed between avgPZ and the first 3 years performance in the NFL indicating a significant relationship between the variables. WR who had a greater avgCZ played

on average a great number of snaps for 2 years out from the combine. Similarly, WRs who performed better examining the combine as a whole had a greater snap average for the first 3 years of their career.

Table 6. Spearman's rho correlation matrix of average positional and total combine Z-scores and average snaps played in the National Football League over a five year period for all position groups.

Position Group		avgCZ	avgPZ	avgS1	avgS2	avgS3	avgS4	avgS5
DBs	**avgCZ**	1	0.797**	0.279**	0.292**	0.242**	0.211*	0.175
	p-value		0.01	0.01	0.01	0.01	0.02	0.19
	n		281	281	224	169	113	57
	avgPZ		1	0.266**	0.259**	0.240**	0.173	0.101
	p-value			0.01	0.01	0.01	0.17	0.1
	n			281	224	169	113	57
LBs	**avgCZ**	1	0.957**	0.280**	0.320**	0.254**	0.194	0.239
	p-value		0.01	0.01	0.01	0.01	0.12	0.17
	n		175	175	142	106	67	34
	avgPZ		1	0.289**	0.332**	0.269**	0.203	0.227
	p-value			0.01	0.01	0.01	0.1	0.2
	n			175	142	106	67	34
OL	**avgCZ**	1	0.698**	0.150*	0.269**	0.188*	0.074	0.042
	p-value		0.01	0.02	0.01	0.02	0.47	0.77
	n		247	247	201	149	96	52
	avgPZ		1	0.057	0.167*	0.044	-0.039	0.104
	p-value			0.37	0.02	0.6	0.71	0.46
	n			247	201	149	96	52
RBs	**avgCZ**	1	0.937**	0.240**	0.215*	0.197*	0.102	0.209
	p-value		0.01	0.01	0.01	0.04	0.4	0.22
	n		165	165	132	105	70	37
	avgPZ		1	0.232**	0.237**	0.234*	0.128	0.187
	p-value			0.01	0.01	0.02	0.29	0.27
	n			165	132	105	70	37
TEs	**avgCZ**	1	0.916**	0.225*	-0.087	0.01	0.023	0.208
	p-value		0.01	0.04	0.49	0.95	0.9	0.41
	n		85	85	66	52	34	18
	avgPZ		1	0.163	-0.12	0.046	0.03	0.299
	p-value			0.14	0.34	0.75	0.86	0.23
	n			85	66	52	34	18
WRs	**avgCZ**	1	0.866**	0.146*	0.192*	0.132	0.087	0.013
	p-value		0.01	0.03	0.01	0.13	0.43	0.94
	n		226	226	172	130	85	35
	avgPZ		1	0.223**	0.210**	0.203*	0.157	0.117

p-value	0.01	0.01	0.02	0.15	0.5
n	226	172	130	85	35

Note: avgCZ = average combine Z-score; avgPZ = average positional Z-score; avgS1 = average snaps played 1 year out from the combine; avgS2 = average snaps played 2 years out from the combine; avgS3 = average snaps played 3 years out from the combine; avgS4 = average snaps played 4 years out from the combine; avgS5 = average snaps played 5 years out from the combine; DBs = defensive backs; LBs = linebackers; OL = offensive linemen; RBs = running backs; TEs = tight ends; WRs = wide recievers
* = Correlation is significant at the 0.05 level; ** = Correlation is significant at the 0.01 level

COMBINE REGRESSION ANALYSIS

The data was inspected to check for any violations of assumptions. The normal probability and residual plots were examined to ensure that the assumptions of normality, linearity, and homoscedasticity were not violated. The observed tolerance levels for each of predictor used in each of the models were above 0.1 indicating no multicollinearity (Field, 2009). Durbin watson values were all within recommended values which are accepted as normal and as no values were less than or greater than 1.0 or 3.0, respectively there was no definite cause for concern (Field, 2009). However, after inspection, three models (Total Combine, DL, and LBs) had standard residual values of above 3.0 or below -3.0 which suggests potential outliers (Field, 2009). Although, Cook's distance was not seen to be greater than 1.0 for any of the models produced suggesting that no one case had any excessive influence on the regression coefficients (Tabachnick & Fidell, 2007); thus they were not removed for the regression models.

The results of the multiple linear regression (MLR) analysis on the Combine measures and performance subsequent years performance in the NFL are presented in Table 7. A MLR was calculated to predict participants avgS based upon performance in each of the six combine measures (i.e., 40yd, BJ, VJ, 3C, PRO, BP) using every athlete who participated in the combine over the five year period. A statistically significant regression model was found (F = 8.516, p = 0.01), with an adjusted r_a^2 of 0.026, when using VJ, BP, and PRO as predictors. The regression model explained 2.6% of the variance in avgS played the year following the Combine. The positive regression coefficient for both VJ and BP indicates that a greater jump height and more repetitions, for each test respectively, were associated with a greater avgS count. The negative regression coefficient for PRO indicates that a faster time is associated with a

greater avgS count. However, BJ (*p=0.39)*, 40yd (*p*=0.10), nor 3C (*p*=0.21) were significant to the regression model.

Table 7. Summary of regression analysis of average combine performance in physical performance tests over a five year period on predicting average snaps played one year out from the National Football League Combine.

Significant Predictors	*B (SE)*	*β*	*t*	*p*	r_a^2	*SEE*	95% CI	
							Lower Bound	*Upper Bound*
VJ	0.471 (0.248)	0.093	1.902	0.06	0.017	21.0577	-0.015	0.957
BP	0.373 (0.132)	0.11	2.832	0.01	0.022	21.4587	0.114	0.631
PRO	-9.290 (4.283)	-0.113	-2.169	0.01	0.026	21.411	-17.697	-0.883

Note: VJ = vertical jump; PRO = shuttle run; BP = bench press; dependent variable = average snaps played; SEE = standard error of the estimate; CI = confidence interval

POSITIONAL REGRESSION ANALYSIS

The results of the multiple linear regression (MLR) analysis on the Combine measures and performance subsequent year's performance in the NFL are presented in Table 8. A MLR was calculated to predict participants avgS based upon performance in each of the 6 combine measures (40yd, BJ, VJ, 3C, PRO, BP) for every athlete over the five year period when separated by position. No regression equation could be calculated for the position of QBs due to lack of significant predictors.

DEFENSIVE BACKS

A significant regression equation was found (F = 8.973, p = 0.01), with an r_a^2 of 0.095, when using 40yd and VJ as predictors. The regression model explained 9.5% of the variance in avgS played the year following the Combine for DBs. The positive regression coefficient for VJ indicates that greater jump heights were associated with a greater avgS count. The negative regression coefficient for 40yd indicates that a faster time is associated with a greater avgS count. However, BP (*p*=0.43), BJ (*p*=0.30), PRO (*p*=0.70), nor 3C (*p*=0.89) were significant to the regression model.

DEFENSIVE LINEMEN

A significant regression equation was found (F = 7.056, p = 0.01), with an r_a^2 of 0.043, when using PRO as a predictor. The regression model explained 4.3% of the variance in avgS played the year

following the Combine for DL. The negative regression coefficient for PRO indicates that a faster time is associated with a greater avgS count. However, 40yd (p=0.11), BJ, BP (p=0.05), VJ (p=0.83), BJ (p=0.29), nor 3C (p=0.71) were significant to the regression model.

LINEBACKERS

A significant regression equation was found (F = 12.906, p = 0.01), with an r_a^2 of 0.202, when using PRO and 40yd as predictors. The proposed regression equation was indicated to be:

$$avgS = 423.609 – (46.344*PRO) – (42.240*40yd)$$

where both PRO and 40yd are measured in seconds. The regression model explained 20.2% of the variance in avgS played the year following the Combine for LBs. The negative regression coefficient for 40yd and PRO indicates that a faster time is associated with a greater avgS count. However, BJ (p=0.24), VJ (p=0.86), BP (p=0.39) nor 3C (p=0.83) were significant to the regression model.

OFFENSIVE LINEMEN

A significant regression equation was found (F = 9.026, p = 0.01), with an r_a^2 of 0.053, when using PRO as a predictor. The regression model explained 5.3% of the variance in avgS played the year following the Combine for OL. The negative regression coefficient for PRO indicates that a faster time is associated with a greater avgS count. However, 40yd (p=0.23), BJ (p=0.11), VJ (p=0.24), BP (p=0.33), nor 3C (p=0.43) were significant to the regression model.

RUNNING BACKS

A significant regression equation was found (F = 6.188, p = 0.02), with an r_a^2 of 0.069, when using BJ as a predictor. The regression model explained 6.9% of the variance in avgS played the year following the Combine for RB. The positive regression coefficient for BJ indicates that a greater jump heights were associated with a greater avgS count. However, VJ (p=0.25), BP (p=0.18), PRO (p=0.73), 40yd (p=0.63), nor 3C (p=0.31) were significant to the regression model.

TIGHT ENDS

A significant regression equation was found (F = 5.830, p = 0.02), with an r_a^2 of 0.115 when using 3C as a predictor. The regression model explained 11.5% of the variance in avgS played the year following the Combine for TE. The negative regression coefficient for 3C indicates that a faster time is associated with a greater avgS count. However, BJ (p=0.33), VJ (p=0.27), BP (p=0.58), PRO (p=0.60), nor 40yd (p=0.26) were significant to the regression model.

WIDE RECEIVERS

A significant regression equation was found (F = 5.182, p = 0.02), with an r_a^2 of 0.042, when using 40yd as a predictor. The regression model explained 4.2% of the variance in avgS played the year following the Combine for WR. The negative regression coefficient for 40yd indicates that a faster time is associated with a greater avgS count. However, BJ (p=0.56), VJ (p=0.66), BP (p=0.73), PRO (p=0.59), nor 3C (p=0.37) were significant to the regression model.

Table 8. Summary of regression analysis of average positional performance in physical performance tests over a five year period on predicting average snaps played one year out from the National Football League Combine.

Position	Significant Predictors	*B (SE)*	*β*	*t*	*p*	r_a^2	*SEE*	95% CI	
								Lower Bound	*Upper Bound*
DBs	**40yd**	-54.335 (20.112)	-0.218	-2.701	0.01	0.068	23.319	-94.082	-14.589
	VJ	1.666 (0.713)	0.188	2.337	0.02	0.095	22.983	0.257	3.075
DL	**PRO**	-15.827 (5.958)	-0.207	-2.656	0.01	0.043	16.042	-27.595	-4.059
LBs	**PRO**	-46.344 (13.378)	-0.334	-3.464	0.01	0.164	18.747	-72.915	-19.773
	40yd	-42.240 (16.594)	-0.245	-0.245	0.01	0.219	18.218	-75.197	-9.283
OL	**PRO**	-32.843 (10.932)	-0.231	-3.004	0.01	0.053	27.001	-54.433	-11.253
RBs	**BJ**	0.648(0.261)	0.262	2.488	0.02	0.069	13.925	0.13	1.166
TEs	**3C**	-28.026 (11.608)	-0.339	-2.415	0.02	0.115	15.942	-51.407	-4.649
WRs	**40yd**	-39.366 (16.329)	-0.204	-2.411	0.02	0.042	19.164	-71.661	-7.071

Note: 40yd = 40yard dash; VJ = vertical jump; PRO = shuttle run; BJ = broad jump; 3C = 3-cone drill; dependent variable = average snaps played; DBs = defensive backs; DL = defensive lines; LBs = linebackers; OL = offensive linemen; RBs = running backs; TEs = tight ends; WRs= wide receivers; SEE = standard error of the estimate; CI = confidence interval

CHAPTER 4

DISCUSSION

The NFL hosts the annual Scouting Combine in which coaches, scouts, and general managers attempt to gain greater insight into potential prospects for the upcoming NFL Draft. The primary results of this study suggest that the NFL Combine lacks predictive ability when examining first year game performance. Furthermore, it also lacks strength when examining relationships between performance and subsequent five year performance in the NFL.

The analysis of the physical performance tests used at the Combine revealed that normalized performance (avgPZ or avgCZ) is weakly correlated with performance in the NFL. Of the 35 correlations only two correlations were found to be moderately strong, avgCZ - avgS2 (r=0.320), avgPZ - avgS2 (r=0.332), whereas the majority were found to be weak (r<0.3). Furthermore, our data analysis suggests that Combine measures can only explain approximately 2.6% of the variance in avgS one year following the Combine when using three (VJ, BP, and PRO) performance tests as predictors. Results of this study are in line with the results found in the literature (Kuzmits & Adams, 2008; Teramoto, Cross & Willick, 2016). A possible explanation as to why performance is a precursor could be explained by further examination of the tests used in the Combine. Our results indicated that tests were significantly predictive of game performance for the eight position groups, however, with the positional regression models only explaining up to 21.9% on the variance a significant amount is left unexplained.

Previous literature have attempted to explain why the Combine struggles to offer insight into future performance (Kuzmits & Adams, 2008; Robbins, 2010; Teramoto, Cross & Willick, 2016). An explanation which has previously been conceived is the concept of equalization of performance due to the "rigorous prep courses" which athletes attend prior to the NFL Combine (Kuzmits & Adams, 2008; Robbins, 2010). To further build upon this point, the athletes who attend the prep courses, and ultimately attend the Combine practice, become better and improve the skill(s) tested at the Combine (i.e., physical performance tests). The improvement of non-sport specific performance (i.e., speed, strength, and power) does not necessarily equate to improvement in football playing ability, although it has been suggested to

be a precursor (Garstecki, Latin & Cuppett, 2004). Therefore, athletes who have improved their ability to perform well in the tests used at the combine may not necessarily be high performers on the field. Thus the equalization of performance would ultimately affect the ability of the combine to differentiate between the high and low in game performers.

When examining each individual performance test used at the combine, further explanation could be found. It has been reported that real game situations are not predetermined and is most often an unpredictable visual stimuli (i.e., agility) (Gabbett et al., 2008; Henry et al., 2011). Both "agility" tests used at the combine, the 3C and PRO tests, are predetermined routes which do not require a reaction to a stimulus. Therefore, although both tests are accurate and reliable, they are measures of change of direction (COD) speed (i.e., physical and technical factors) (Stewart, Turner & Miller, 2012) not agility. COD speed is a central component of multidirectional sports (Jones, Bampouras, & Marrin, 2009) but the tests fail to measure the cognitive factors associated with agility (i.e. decision making speed and accuracy) (Young, Dawson & Henry, 2015). Therefore, if the combine only reveals insight into two of the three aspects associated with agility, true insight into possible game performance would be hindered.

The type of running surface has been called into question. As of 2015, 17 and 14 NFL stadiums are equipped with artificial turf or natural grass, respectively (Egan, 2015). While research has indicated that on newer forms of artificial turf linear sprint speed is similar to natural grass; however, COD speed was reported to be significantly faster (-3.0% ±2.8%, $p < 0.001$) (Gains, Swedenhjelm, Mayhew, Bird & Houser, 2010). The Combine is held at Lucas Oil Stadium, IN which is equipped with artificial, more specifically FieldTurf (FieldTurf, Calhoun, GA). Although there is an absence of an absolute difference, nearly half of the stadiums do use natural grass which could ultimately affect COD speed and therefore in game performance. In addition, the Combine is performed in minimalist attire whereas in games athletes perform with pads and helmets. Works by Brechue, Mayhew and Piper (2005) reported that football equipment has seen to impair sprint performance by -2.9% +- 1.8% when compared to wearing shorts and t-shirt. An in game reduction in sprint performance could explain the lack of predictive strength for the 40yd in this study.

A secondary notion, as evidence builds against the current predictive validity of the NFL Combine, it opens discussion as to why the Combine is still held if it offers little relationship or predictability to future game performance. One possible reason is that it offers an already lucrative franchise, an additional chance for monetary gain. It was reported that in 2010 approximately 5.2 million people watched the NFL Combine, in 2012 the viewing base had grown to 6.51 million (Heitner, 2013). The removal of such a highly publicized event, which continues to engross more and more views, could be deemed as a costly move. Secondly, while outside the scope of this study, some research has investigated the ability for the Combine to predict draft status (Hartman, 2011; McGee and Burkett, 2003; Robbins, 2010). Although research is conflicting, works by McGee and Burkett (2003) reported that performance at the Combine is highly predictive of eventual draft status for RBs, WRs, and DBs it could be financially advantageous for athletes to perform well at the Combine. According to "2017 NFL Draft Tracker" (2017), the average contract value for a first round draft pick was approximately $14,947,470 with $8,854,246 being guaranteed as a signing bonus. In comparison, the average contract value for a round two draft pick was $5,518,265 with $2,153,283 being guaranteed as a signing bonus. With the potential to possibly be selected earlier in the draft and increase possible earnings it could offer reason as to why athletes proceed to partake and train in order to perform well.

Nevertheless, sports performance is a complex multifaceted concept, demanding multiple physical and mental skills (Mackenzie & Cushion, 2013). It is important to note that the six physical performance tests examined in this study are only one aspect of the Combine. Although very little research has examined the predictive ability of the other aspects medical history, Wonderlic, and sport specific drills it has suggested its tentative usefulness (Brophy et al., 2008; Pitts & Evans, 2018). While outside the scope of this study to comment any further on the other aspects of the Combine, this could suggest that NFL teams may weight performance in other aspects higher than that of the physical performance tests. Alternatively, it could suggest that the Combine as a whole does not bare much usefulness and thus receives little acknowledgement from NFL scouts or general managers.

There are limitations associated with this study. Firstly, the data used for this study was public access data obtained from third party online sources. If possible these sources manually check across a variety of reliable web sources to attempt to ensure accuracy. Secondly, not every athlete at the Combine completes every event. As stated in the results, only 802 (53.5%) athletes completed all six events between 2013-2017. Within our total sample, no QB participated in the BP and only 70.2% and 72.6% of athletes completed the 3C and PRO, respectively. Thus, avgCZ and avgPZ of athletes who completed more events would offer a better representation of their athletic performance than an athlete who only completed one event. Furthermore, due to the large number of correlations performed, spurious findings could be present. Thus, some significant correlations found could be attributable to the random chance model. Lastly, this study only analyzed the physical performance tests which is one part of the NFL Combine. Unfortunately, the results in the Wonderlic, medical history, and position specific drills these tests are not publicly accessible, therefore, we were unable to include them for analysis. The addition of these tests for analysis could offer a greater insight into the relationship between the NFL Combine and future NFL performance.

CHAPTER 5

PRACTICAL APPLICATIONS

Of the 1537 athletes whose data was collected this study found the Combine was able to explain 2.6% of the variance in average snaps played in the NFL. The results of this study add to the growing body of literature which suggests the lack of usefulness of the Combine in predicting future NFL performance. Although, examining Combine data comparison by position group could offer greater insight to future performance. Even so, caution should be used if coaches, general managers, and other front office staff are considering the use of Combine data as an aid for possible selection in the upcoming draft. The revision of the physical performance tests used to ones that offer a greater transfer to real game performance could aid in improvement to the predictability of the Combine. Future studies should direct research to other aspects of the Combine which could be advantageous in providing insight into future performance (i.e. Wonderlic, medical history). In addition, an all-encompassing study of every aspect of the Combine could provide valuable insight to the predictability of the Combine as a whole and not as individual sections.

REFERENCES

2017 NFL Draft Tracker. (2017). Retrieved from https://www.spotrac.com/nfl/draft/2017/

Berg, K., Latin, R.W., & Baechle, T. (1990). Physical and performance characteristics of NCAA division I football players. *Res. Quart. 61*:395-401.

Black, W., & Roundy, E. (1994). Comparisons of size, strength, speed, and power in ncaa division 1-a football players. *Journal of Strength and Conditioning Research, 8*(2), 80-85. doi: 10.1519/00124278-199405000-00004

Brophy, R., Chehab, E., Barnes, R., Lyman, S., Rodeo, S., & Warren, R. (2008). Predictive value of orthopedic evaluation and injury history at the nfl combine. *Medicine & Science in Sports & Exercise, 40*(8), 1368-1372. doi: 10.1249/mss.0b013e31816f1c28

Bunn, J.A., Ryan G.A., Button, G.R., & Zhang, S. (2017). Evaluation of strength and conditioning measures with game success in Division I collegiate volleyball. *Journal of Strength and Conditioning Research*. Advance online publication. doi: 10.1519/JSC.0000000000002181

Clark, K., Rieger, R., Bruno, R., & Stearne, D. (2017). The NFL Combine 40-Yard Dash. *Journal of Strength and Conditioning Research*. doi: 10.1519/jsc.0000000000002081

Coburn, J. (2012). Measuring Power. Strength and Conditioning Journal, 34(6), 25-28. doi: 10.1519/SSC.0b013e3182708a7d

Dodrill, C. B., & Warner, M. H. (1988). Further studies of the Wonderlic Personnel Test as a brief measure of intelligence. *Journal of Consulting and Clinical Psychology, 56*(1), 145-147.http://dx.doi.org/10.1037/0022-006X.56.1.145

Egan, J. (2015). Turf Battle in the NFL: Natural vs. Artificial - Lawnstarter. Retrieved from https://www.lawnstarter.com/blog/lawn-care-2/turf-at-nfl-stadiums/

Field, A. (2009). *Discovering statistics using IBM SPSS statistics*. Los Angeles, CA: SAGE.

Fry, A.C., & Kraemer, W.J. (1991). Physical performance characteristics of American collegiate football players. *J. Appl. Sport Sci. Res. 5*:126-138.

Gabbett T.J., Kelly J.N., Sheppard J.M. (2008) Speed, change of direction speed, and reactive agility of rugby league players. *Journal of Strength and Conditioning Research 22*(1), 174-181.

Gaines, C. (2016). The 32 NFL teams are worth as much as every MLB and NBA team combined. Retrieved from https://www.businessinsider.com/nfl-teams-value-2016-9

Gains, G., Swedenhjelm, A., Mayhew, J., Bird, H., & Houser, J. (2010). Comparison of Speed and Agility Performance of College Football Players on Field Turf and Natural Grass. *Journal of Strength and Conditioning Research*, *24*(10), 2613-2617. doi: 10.1519/jsc.0b013e3181eccdf8

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

Haff, G., & Triplett, N. (2015). Essentials of strength training and conditioning (4th ed.). Champaign, IL: Human Kinetics.

Hartman, M. (2011). Competitive performance compared to combine performance as a valid predictor of nfl draft status. *Journal of Strength and Conditioning Research, 25*, S105-S106. doi: 10.1097/01.jsc.0000395746.03546.e8

Hartman, M. (2011). Competitive Performance Compared to Combine Performance as a Valid Predictor of NFL Draft Status. *Journal of Strength and Conditioning Research, 25*, S105-S106. doi: 10.1097/01.jsc.0000395746.03546.e8

Heitner, D. (2013). The Economics Of The NFL Combine. Retrieved from https://www.forbes.com/sites/darrenheitner/2013/02/21/the-economics-of-the-nfl-combine/#708d9ffd385e

Henry G., Dawson B., Lay B., Young W. (2011) Validity of a reactive agility test for Australian football. *International Journal of Sports Physiology and Performance*, *6*(4), 534-545.

History | NFL Combine. Retrieved from http://www.nflcombine.net/history/

Jones, P., Bampouras, T., & Marrin, K. (2009). An investigation into the physical determinants of change of direction speed. *The Journal of Sports Medicine and Physical Fitness*, *49*(1), 97-104.

Kuzmits, F., & Adams, A. (2008). The NFL combine: does it predict performance in the national football league?. *Journal of Strength and Conditioning Research, 22*(6), 1721-1727. doi: 10.1519/jsc.0b013e318185f09d

Leard, J., Cirillo, M., Katsnelson, E., Kimiatek, D., Miller, T., Trebincevic, K., & Garbalosa, J. (2007). Validity of two alternative systems for measuring vertical jump height. *Journal of Strength and Conditioning Research*, 21(4), 1296-1299.

Levinger, I., Goodman, C., Hare, D., Jerums, G., Toia, D., & Selig, S. (2009). The reliability of the 1RM strength test for untrained middle-aged individuals. *Journal of Science and Medicine In Sport, 12*(2), 310-31

Mackenzie, R., & Cushion, C. (2013). Performance analysis in football: A critical review and implications for future research. *Journal of Sports Sciences, 31*(6), 639-676. doi: 10.1080/02640414.2012.746720

Mann, J., Ivey, P., Brechue, W., & Mayhew, J. (2015). Validity and reliability of hand and electronic timing for 40-yd sprint in college football players. *Journal of Strength and Conditioning Research, 29*(6), 1509-1514. doi: 10.1519/jsc.0000000000000941

Matthews, T., & Lassiter, K. (2007). What does the wonderlic personnel test measure?. *Psychological Reports, 100*(3), 707-712. doi: 10.2466/pr0.100.3.707-712

Mayhew, J., Gaines, G., Swedenhjelm, A., Cooper, J., Bird, M., & Houser, J. (2010). Difference in 40-yd dash and pro-agility times on artificial turf and natural grass. *Journal of Strength and Conditioning Research, 24*, 1. doi: 10.1097/01.jsc.0000367166.10606.0a

Mayhew, J., Ware, J., Bemben, M., Wilt, B., Ward, T., & Farris, B. et al. (1999). The nfl-225 test as a measure of bench press strength in college football players. *Journal of Strength and Conditioning Research, 13*(2), 130-134. doi: 10.1519/00124278-199905000-000066. doi: 10.1016/j.jsams.2007.10.007

Mayhew, J., Ware, J., Cannon, K., Corbett, S., Chapman, P., & Bemben, M. et al. (2002). Validation of the NFL-225 test for predicting 1-RM bench press performance in college football players. *Journal of Sports Medicine and Physical Fitness, 43*(2), 304-208.

Mcgee, K., & Burkett, L. (2003). The national football league combine: a reliable predictor of draft status?. *The Journal of Strength and Conditioning Research, 17*(1), 6. doi: 10.1519/1533-4287(2003)017<0006:tnflca>2.0.co;2

Mcgee, K., & Burkett, L. (2003). The national football league combine: a reliable predictor of draft status?. *The Journal of Strength and Conditioning Research, 17*(1), 6. doi: 10.1519/1533-4287(2003)017<0006:tnflca>2.0.co;2

Menzel, H., Chagas, M., Szmuchrowski, L., Araujo, S., Campos, C., & Giannetti, M. (2010). Usefulness of the jump-and-reach test in assessment of vertical jump performance. *Perceptual and Motor Skills, 110*(1), 150-158. http://dx.doi.org/10.2466/pms.110.1.150-158

New England Patriots franchise value 2002-2018 | Statistic. (2018). Retrieved from https://www.statista.com/statistics/194482/franchise-value-of-the-new-england-patriots-since-2006/

Nimphius S. Increasing agility. In: Joyce D., Lewindon D., editors. High-Performance Training for Sports. Human Kinetics; Champaign, IL, USA: 2014. pp. 185–198.

Pitts, J. D., & Evans, B. (2018). Evidence on the importance of cognitive ability tests for NFL quarterbacks: what are the relationships among Wonderlic scores, draft positions and NFL performance outcomes? *Applied Economics, 50*(27), 2957. https://doi.org/10.1080/00036846.2017.1412081

Retrieved from https://www.forbes.com/nfl-valuations/list/#tab:overall

Riebe, D., Ehrman, J., Liguori, G., & Magal, M. (2017). ACSM's guidelines for exercise testing and prescription (10th ed.). Philadelphia: Wolters Kluwer.

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*, *24*(11), 2888-2899. doi: 10.1519/jsc.0b013e3181f927cc

Rodriguez-Rosell, D., Mora-Custodio, R., Franco-Márquez, F., Yáñez-García, J., & González-Badillo, J. (2016). Traditional vs. sport-specific vertical jump tests: reliability, validity and relationship with the legs strength and sprint performance in adult and teen soccer and basketball players. *Journal of Strength and Conditioning Research,* 31(1):196–206. https://doi.org/10.1519/JSC.0000000000001476

Schmidt, W. (1999). Strength and physiological characteristics of NCAA division III american football players. *Journal of Strength and Conditioning Research, 13*(3), 210-213. doi: 10.1519/00124278-199908000-00005

Sierer, S., Battaglini, C., Mihalik, J., Shields, E., & Tomasini, N. (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. doi: 10.1519/jsc.0b013e31815ef90c

Spasic, M., Krolo, A., Zenic, N., Delextrat, A., & Sekulic, D. (2015). Reactive Agility Performance in Handball; Development and Evaluation of a Sport-Specific Measurement Protocol. *Journal of Sports Science & Medicine*, *14*(3), 501-6.

Stewart, P., Turner, A., & Miller, S. (2012). Reliability, factorial validity, and interrelationships of five commonly used change of direction speed tests. *Scandinavian Journal of Medicine & Science in Sports, 24*(3), 500-506. doi: 10.1111/sms.12019

Sue, R., Harris, C., Adams, K., Berning, J., & DeBeliso, M. (2017). Determination of trials needed for measurement consistency of standing long jump in collegiate volleyball players. *Medicine & Science in Sports & Exercise, 49*, 756. doi: 10.1249/01.mss.0000519011.64616.e3

Tabachnick, B., & Fidell, L. (2012). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson Education.

Tarter, B., Kirisci, L., Tarter, R., Weatherbee, S., Jamnik, V., McGuire, E., & Gledhill, N. (2009). Use of aggregate fitness indicators to predict transition into the national hockey league. *Journal of Strength and Conditioning Research*, *23*(6), 1828-1832. doi: 10.1519/jsc.0b013e3181b4372b

Teramoto, M., Cross, C., & Willick, S. (2016). Predictive value of national football league scouting combine on future performance of running backs and wide receivers. *Journal of Strength and Conditioning Research*, *30*(5), 1379-1390. doi: 10.1519/jsc.0000000000001202

Teramoto, M., Cross, C., Rieger, R., Maak, T., & Willick, S. (2017). Predictive Validity of National Basketball Association Draft Combine on Future Performance. *Journal of Strength and Conditioning Research*. doi: 10.1519/jsc.0000000000001798

Workouts & Drills. (2019). Retrieved from http://www.nfl.com/combine/workouts

Young, W., Dawson, B., & Henry, G. (2015). Agility and Change-of-Direction Speed are Independent Skills: Implications for Training for Agility in Invasion Sports. *International Journal of Sports Science & Coaching*, *10*(1), 159-169. doi: 10.1260/1747-9541.10.1.159\

Young, W., Russell, A., Burge, P., Clarke, A., Cormack, S., & Stewart, G. (2008). The use of sprint tests for assessment of speed qualities of elite australian rules footballers. *International Journal of Sports Physiology and Performance*, *3*(2), 199-206. doi: 10.1123/ijspp.3.2.199

APPENDIX A

LITERATURE REVIEW

The National Football League (NFL) is a multi-billion dollar business. According to Forbes in 2016 the total worth of the NFL was estimated at $74.8 billion in 2018 it was estimated to be $82.2 billion (https://www.forbes.com/nfl-valuations/list/#tab:overall). In 2016 the average NFL franchise was valued at $2.3 billion compared to $2.5 billion in the 2018 valuation (Gaines, 2016). Increases in franchise worth can be associated with achieving post season seasons and super bowl appearances, as with all sports it pays to win. For example in 2013 the New England Patriots were worth $1.8 billion in the following season, 2014/15, they went on to win Super Bowl XLIX. In 2015 they were re-estimated at being worth $3.2 billion. Again in 2017 they won another Super Bowl, Super Bowl LI, and they are now estimated to be worth $3.8 billion ("New England Patriots franchise value 2002-2018", 2018). As the value increases so does the pressure for franchises to put together winning seasons. Thus, the need to gain a potential competitive edge has unlocked the exploration of alternate avenues to possibly offer greater insight into a player's future game performance. One of those avenues is the NFL Combine.

The NFL Combine started in 1982 with a total of 163 players attending and originally was called National Invitational Camp ("History | NFL Combine", n.d.). Currently every year the NFL invites the top prospects from collegiate football to showcase individual talents in a series of performance and psychological tests. The purpose of the combine is to provide insight into the medical history, athletic abilities, psychological state, as well as demonstrations of skill in positional drills of potential prospects. At the NFL combine the physical examination and medical history is used to grade the athlete on their ability to participate in the NFL. Although research is limited, the orthopedic screen and injury history has been seen to offer insight into the probability of playing in the NFL with higher grade athletes having an extended career compared to those of a low grade (Brophy et al., 2008).

The athletes are also asked to complete a psychological test called the Wonderlic test. The Wonderlic is used to briefly assess the athlete's mental capability and intelligence (Dodrill & Warner,

1988; Matthews & Lassiter, 2007). For quarterbacks scoring well in this test could be advantageous as research has suggested it could offer insight to future performance in the NFL (Pitts & Evans, 2018).The most publicized aspect of the combine are the physical performance tests. The physical performance tests are designed to assess athletic abilities associated with american football and include; the 40-yard dash (40yd), pro-agility shuttle (PRO), 225lb bench press repetition max test (BP), 3-cone drill (3C), vertical jump (VJ) and broad jump (BJ).

The topic of physical attributes of American football has been well researched and indicates the importance that strength, speed, and power play in the game (Berg, Latin & Baechle 1990; Black & Roundy, 1994; Fry & Kraemer, 1991; Garstecki, Latin & Cuppett 2004; Schmidt, 1999). The battery of physical tests used in the NFL Combine have been arranged to allow for testing of each of the physical attributes associated with the game of American football. Interestingly, research has suggested that these attributes can differentiate between; starters and nonstarters (Fry & Kraemer, 1991), level of competition (Garstecki, Latin & Cuppett 2004), and draft status (i.e. drafted vs non-drafted) (Sierer, Battaglini, Mihalik, Shields & Tomasini, 2008). Sierer, Battaglini, Mihalik, Shields and Tomasini (2008) found that when American football players were seperated into 3 groups of skill, big skill, and linemen, and the performance in the individual performance tests was used, there was a significant difference between drafted and non-drafted players. Furthermore it has could be suggested that these attributes are becoming more valuable as players are becoming bigger, stronger, faster. Works by Fitzgerald & Jensen (2018) investigated how performances differed between athletes invited to 1999-2000 and 2015-2016 NFL Combine. They observed that results were significantly faster for in 40 yard dash and jumped significantly higher in the vertical jump. Furthermore, significant increases in performance were suggested in the broad jump and 3-Cone drill.

Other sports also offer a combine where prospects can demonstrate their potential. The National Basketball Association (NBA) also hold an annual NBA Combine in which top prospects are invited to. At the NBA Combine athletes are tested through 8 measures of anthropometrics and 6 measures of; speed, strength, and agility. The 8 measures of anthropometrics consist of; body fat percentage, hand

length, hand width, height without shoes, height with shoes, standing reach, weight, wingspan. The 6 measures of speed, strength, and agility are; lane agility, shuttle run, three-quarter court sprint, standing vertical jump, maximum vertical jump, bench press. Although the NBA Combine is well established research assessing the predictive validity of the tests is limited. One study analyzed the relationship between NBA Combine performance in the measures of anthropometrics as well as those of speed, strength, and agility and on-court performance for players in their first 3 years in the NBA. On-court performance was measured through a multiple of different statistics including but not limited to; player efficiency rating, offensive and defensive plus/minus. It was found that in conclusion the NBA Combine had value in ability to predict performance in the NBA (Teramoto, Cross, Rieger, Maak & Willick, 2017).

Another major sport which offers a combine is the National Hockey League (NHL) who also annually hold a combine. At the NHL scouting combine the athletes are placed through 12 fitness tests; height, weight, Y-balance, Functional Movement Screen, grip strength, VO2max test, standing long jump, jump station, bench press, pro agility, max repetition pull ups, Wingate test. Similarly to the NBA research into the predictive validity of the NHL scouting combine is limited. Tarter et al (2009) investigated the NHL scouting combine over a 4 year period. Fitness tests were group into; upper body strength, lower body power, body composition, and energy systems and was analyzed using a factor analysis in which an overall composite index was derived. The results suggested that players scoring in the 90th percentile on this index had either a 72% or 60%, dependent on position (defensemen/forwards), chance of playing in the NHL within 4 years.

The physical performance tests used in the NFL Combine are used because previous research has found that these tests are valid and reliable ways to test the physical attributes deemed important to American football. The physical performance tests can be separated into three distinct categories; speed and agility, lower body power, and upper body strength or endurance.

The speed and agility tests consists of the 40-yard dash, pro agility, and 3-cone drill. The 40-yard dash (36.6m) is a test used to measure the athletes sprinting ability which literature has suggested includes both acceleration (Young et al., 2008) and maximum velocity components (Clark, Rieger, Bruno &

Stearne, 2017). In addition the 40-yard dash has been seen to be a valid and reliable measure of speed (r=0.99) (Mayhew et al., 2010), (r=0.988) (Mann, Ivey, Brechue & Mayhew, 2015). The two agility tests seen are the pro-agility shuttle and the 3-cone drill. The pro-agility shuttle has been found to be a highly reliable and valid measure of change of direction (r=0.90, 95%CI=0.84-0.94) (Stewart, Turner & Miller, 2012). Although findings have differed within the literature with Mayhew et al. (2010) finding a lower reliability (r=0.80), it should be noted timings were measured using hand timers instead of electronic timing gates. Thee 3-cone drill, often referred to as the L-Run again has been shown to be a highly reliable and valid measure of change of direction ability (r=0.96, 95%CI=0.90-0.96) (Stewart, Turner & Miller, 2012).

The only upper body based test performed at the combine is the 225lb bench press repetition max test. The test is described by the NFL as being an upper body strength and endurance test. The current gold standard for field based strength tests is the one repetition maximum (1RM) (Levinger et al., 2009). The 225lb bench press repetition max test differs from the 1RM due it being a fixed weight and requires you to lift the weight for a max number of repetitions. At the NFL Combine up to 49 repetitions have been recorded which is why it is common belief that the test doesn't test pure strength but that of muscular endurance. Research has looked into the ability to predict 1RM from the 225lb bench press repetition max test, results indicate that with the use of specific 225lb bench press repetition max test equations it can be used to predict 1RM but as repetitions increase it becomes more inaccurate (Mayhew et al., 2002; Mayhew et al., 1999).

The two lower body power tests used are the vertical jump and the broad jump. The NFL employs the use of the Vertec device to measure jump height. Although vertical jump testing has been seen as a valid and reliable measure of lower body explosive power (Rodriguez-Rosell, Mora-Custodio, Franco-Márquez, Yáñez-García, & González-Badillo, 2016), the validity and reliability of the Vertec have been questioned. Research by Leard et al (2007) found that the Vertec system had a high correlation with the 3-camera system but the mean significantly differed. When comparing the Vertec to a contact mat research has shown that the Vertec device tends to overestimate jump height (Menzel et al., 2010). In summary of

previous research it seems that the accuracy of the Vertec device depends on which method it is being compared to. Although the Vertec is not the gold standard, it can still offer an insight into lower body explosive power, more specifically vertical power. In comparison the second power test used, the broad jump, is used to again test muscular power but horizontal power (Coburn, 2012; Haff & Triplett, 2015). Literature has found to be a reliable and valid measure of lower body power (r=0.88, 90%IC= 0.61-0.91) (Sue, Harris, Adams, Berning & DeBeliso, 2017).

The combine has been thoroughly researched in its ability to measure athletic characteristics (Riebe, Ehrman, Liguori & Magal, 2017) and appears to be in line with the attributes which have been indicated to be important to American football performance (i.e. face validity), there has been doubt on the ability of the NFL combine to predict draft status as well as future success of athletes within the NFL (i.e. predictive validity).

Several studies have been conducted to investigate the link between NFL Combine performance and then eventual draft status, each with varying results and conclusions. A study conducted by Robbins (2010) investigated whether the NFL Combine physical performance tests could predict draft order between the years of 2005-2009. The correlation between physical performance and draft order was analyzed through Pearson r correlations, looking and both direction and strength of the relationship. The study concluded that the physical performance tests had little to no relationship with draft success, whether data was normalized or not. Sprint time and jumping ability appeared to be the best predictors.

Another study conducted by Hartman (2011) similarly investigated the use of performance indicators on draft status. Hartman used 15 running backs from NCAA FBS schools who were selected in the 2009 NFL Draft. In contrast to works by Robbins (2010) the performance indicators used by Hartman (2011) were a combination of both combine measures (40-yard dash and predicted sprint power) and statistics from the players final year of eligibility (total yards, yards per carry, and touchdowns). The use of a Pearson product moment correlations employed to examine the relationships between draft status and performance indicators. The results found a significant moderate relationship between total yards and

draft status ($r= -0.66$, $p<0.05$). It was concluded that total yards was the best predictor for running backs in the 2009 NFL draft.

McGee and Burkett (2003) also examined whether physical performance tests could determine draft status. The study used eight performance measures from the 2000 NFL Combine to examine the relationship, the 40-yard dash was split into three separate tests (10, 20, and 40 yard dash). The relationship between combine data and draft status was analyzed using multiple linear stepwise regressions. Prediction equations were generated for seven positional groups; Quarterback, Wide receivers, Running backs, Offensive line, Defensive line, Defensive backs, and Linebackers. Results indicated and the study concluded that the combine accurately predict draft status for running backs $r^2= 1.00$, wide receivers ($r^2= 1.00$, defensive backs ($r^2=1.00$. Furthermore, it can also be used for other positions but with less accuracy (Offensive Linemen ($r^2=0.70$), defensive linemen ($r^2= 0.59$), Linebackers, ($r^2= 0.22$, Quarterbacks, ($r^2=0.84$).

Similarly research has also then investigated the predictive validity of NFL Combine data to subsequent in game performance in the NFL. Works by Kuzmits and Adams (2008) examined the predictive validity of the combine for; quarterbacks, wide receivers and running backs from 1999 to 2004. Four performance measures were used for each position, draft order, yearly salary, and games played per season were used for every position. Each position then had an additional performance measure, for quarterbacks player ranking was used, for running backs average yards per carry was used, and for wide receivers average yards gained per reception was used. They found that combine exercises were not correlated with NFL success for either quarterbacks or wide receivers. For running backs strong correlations between three highly multicolinear sprints times (40, 20, and 10 yard) and the measures of success used was observed.

Works by examined the predictive validity of the NFL Combine from 2000-2009 for running backs and wide receivers. For both position performance measures were number of games played in the first 3 years of the NFL careers and in the entire NFL careers, draft position and career statistics (running backs - rushing yards per attempt, wide receivers - receiving yards per reception). The relationships

between combine data and the performance measures was tested first by the use of a Pearson's correlation coefficient. The predictive ability was then tested through the use of multiple linear regressions analysis. For running backs they observed, the best predictor of performance was the 10 yard sprint time for both the first three years ($p=0.002$, $r^2=0.154$) and over the career of the athlete ($p<0.001$, $r^2=0.229$). Vertical jump was observed at being significantly associated with performance in the NFL when examining wide receivers, for both the first three years ($p=0.001$, $r^2= 0.077$) and career performance ($p=0.004$, $r^2=0.086$). Both observations were after accounting for the covariates of; height, weight, number of games played, and draft position.