THE NFL COMBINE: DOES IT PREDICT PERFORMANCE IN THE NATIONAL FOOTBALL LEAGUE?


Abstract (Summary)
The authors investigate the correlation between National Football League (NFL) combine test results and NFL success for players drafted at three different offensive positions (quarterback, running back, and wide receiver) during a recent 6-year period, 1999-2004. The combine consists of series of drills, exercises, interviews, aptitude tests, and physical exams designed to assess the skills of promising college football players and to predict their performance in the NFL. Combine measures examined in this study include 10-, 20-, and 40-yard dashes, bench press, vertical jump, broad jump, 20- and 60-yard shuttles, three-cone drill, and the Wonderlic Personnel Test. Performance criteria include 10 variables: draft order; 3 years each of salary received and games played; and position-specific data. Using correlation analysis, we find no consistent statistical relationship between combine tests and professional football performance, with the notable exception of sprint tests for running backs. We put forth possible explanations for the general lack of statistical relations detected, and, consequently, we question the overall usefulness of the combine. We also offer suggestions for improving the prediction of success in the NFL, primarily the use of more rigorous psychological tests and the examination of collegiate performance as a job sample test. Finally, from a practical standpoint, the results of the study should encourage NFL team personnel to reevaluate the usefulness of the combine's physical tests and exercises as predictors of player performance. This study should encourage team personnel to consider the weighting and importance of various combine measures and the potential benefits of overhauling the combine process, with the goal of creating a more valid system for predicting player success.

INTRODUCTION

Each spring, the National Football League (NFL) conducts an event known as the combine, a series of tests designed to assess the skills of promising college football players. The participants, primarily seniors, attend by invitation only to showcase their athletic talents with the goal of being drafted and consequently offered a coveted contract with a
professional football team. Although each participant has demonstrated his skills throughout his collegiate playing career, the combine requires the players to display their talents by performing standardized tests under controlled conditions, thereby creating fair and unbiased assessment conditions. The combine invitees are among the elite college football players in the United States. Out of approximately 10,000 college football players representing 119 National Collegiate Athletic Association (NCAA) Division I teams, about 330, or slightly more than 3%, are invited to Indianapolis, Indiana, for 5 days to participate in drills, exercises, interviews, aptitude tests, and physical exams.

Although the combine clearly possesses face validity as a measure of NFL performance, serious questions remain regarding its predictive validity. Few studies have examined scientifically whether a relationship exists between a graduating collegiate football player's combine performance and subsequent performance as a professional football player. One component of the combine, a test of cognitive ability known as Wonderlic Personnel Test (WPT), and its relationship with NFL performance, has been examined in three studies (1,6,8). None of the studies have shown any relationship between test score and success in the NFL. One study examined the relationship between multiple combine performance measures and NFL draft order (7). The study concluded that the combine most accurately predicts the draft status of running backs, wide receivers, and defensive backs and serves as only a good to fair predictor for other positions.

Beyond the WPT and the study cited above, no studies have examined the relationship between multiple combine measures and performance at the professional football level. However, two studies have examined the relationship between combine-type measures and football skills in collegiate players. In a study involving 40 NCAA Division IA collegiate football players, the only variable among eight running and strength measures that significantly correlated with football-playing ability was the vertical jump, which was significantly correlated with football-playing ability in all position groups (11). In a second study, the results of a series of physical tests similar to combine measures and the Cattell Personality Inventory (CPI) of 59 Division IAA players were correlated with coaches' rankings based on perceptions of overall player performance. Starters were differentiated from nonstarters by the leg squat and two forms of jumping ability. Personality variables were not correlated with coaches' rankings of performance (2).

Unfortunately, the dearth of scientific inquiry into the predictability of the NFL combine does not enable one to draw any conclusions of the validity of any combine measure, with the possible exception of the WPT, which failed to predict various, dependent variables in multiple studies. Interestingly, however, the vertical jump did correlate with playing ability in two studies involving college players, leading one to speculate on the power of the vertical jump to predict performance in the professional ranks.

It is surprising that, as an employment selection device, the NFL combine has not received more scientific analysis and scrutiny. The implications of making false-positive and falsenegative selection errors are huge, given the significant short- and long-term impact of the player-selection process on team success. Presumably, combine performance plays a key role in player selection, and, ultimately, player selection has both significant short- and long-term implications for team performance. Further, the existence and availability of archival combine performance data as well as various dependent variable (NFL success) data makes one wonder why the selection process has been subjected to so little scientific analysis.

The primary purpose of this study is to examine the validity of the combine as a predictor of NFL performance. Specifically, the key research question is: Does a player's performance at the NFL combine in fact correlate with his success as a professional football player?

METHODS

Experimental Approach to the Problem

We will present detailed tables of correlation by position (quarterback, running back, and wide receivers) who attended the combine and were drafted. The data are aggregated for the 6-year study period to yield relatively large sample sizes. Specifically, we will be examining the correlation between the various measures obtained at the combine and the 10 measures of NFL success. Because we are hypothesizing that a zero correlation model describes the data set, we will be looking for sample correlations that are significantly different from zero.

Subjects

Subjects for this study include combine invitees drafted at three positions for the years 1999-2004: quarterback, wide receiver, and running back. As compared with other football positions, relatively objective and measurable NFL performance data are readily available for these three "skill" positions. A total of 306 drafted are included in the 6-year sample: 68 quarterbacks, 152 wide receivers, and 86 running backs. (Nearly half the combine attendees were not drafted; for the years under study, 116 quarterbacks, 272 wide receivers, and 170 running backs-a total of 558 players at these three positions-took part in the combine.) Table 1 shows the numbers of draftees by position for the years (1999-2004)
under study.

Data Collection

Data from the NFL combine were collected from NFLdraftscout.com, a commercial Web site devoted to reporting a wide variety of historical and current data on NFL players, teams, and prospects. Data from the Web site are deemed accurate. For this study, we include the results of all physical exercises as well as the WPT.

Combine results reported in this study include

* 40-yard dash

Starting from a three-point stance, the player runs 40 yards as fast as possible. Times are recorded in 10-, 20-, and 40-yard increments. The 40-yard dash is a test of speed and explosion.

* Bench press

The player's goal in this exercise is to bench press 225 lb as many times as possible. With the exception of quarterbacks and wide receivers, all players participate in this test of upperbody strength.

* Vertical jump

To measure the vertical jump, a player stands flat-footed in front of a pole with a number of plastic flags sticking out. The player then jumps from a standing position and swats as many flags as he can, thus enabling the judge to determine how high the player can jump. This exercise is considered important for wide receivers and defensive backs, where jumping ability is a critical skill.

* Broad jump

The broad jump measures how far a player can jump from a standing position. This drill is most important to positions that use lower-body strength (such as offensive and defensive linemen) to gain an advantage.

* Three-cone drill

In this exercise, three cones are set up in a triangle shape with each cone 5 yards apart. Starting in a three-point stance, the player sprints in a predetermined route among the cones. This exercise tests speed, agility, and cutting ability.

* 20-yard shuttle

In this exercise, the player starts in a three-point stance and runs 5 yards in one direction, 10 yards in the opposite direction, and then sprints back to the starting point. This exercise tests lateral speed and coordination.

* 60-yard shuttle

Similar to the 20-yard shuttle, this exercise requires the player to run 10 and 20 yards instead of 5 and 10 yards before returning to the starting point. This drill is considered to be a test of physical endurance as well as lateral speed and coordination.

* Wonderlic Personnel Test

The WPT is a timed, 12-minute, 50-question test that measures general intelligence. Whereas the test sees its greatest use in industry, it has been given at the combine since the 1970s. Wonderlic Inc. press release information claims that "Every potential draft pick takes the Wonderlic Personnel Test at the combine to prove he does ... or doesn't-have the brains to win the game" (13).

Performance criteria for the positions under study were obtained online from NFL.com, a Web site owned and operated by the NFL. Specific performance criteria and draft order for each position included in this study are listed below.

Quarterback. Quarterback criteria included draft order, salary for years 1, 2, and 3, games played during years 1, 2, and 3, and quarterback rating for years 1, 2, and 3. The quarterback rating based on four factors: percentage of completions per attempt, average yards gained per attempt, percentage of touchdown passes per attempt, and percentage of interceptions per attempt (9).

Running Back. Running back criteria included draft order, salary for years 1, 2, and 3, games played during years 1, 2, and 3, and average yards per carry for years 1, 2, and 3.
Wide Receiver. Wide receiver criteria included draft order, salary for years 1, 2, and 3, games played during years 1, 2, and 3, and average yards gained per reception for years 1, 2, and 3.

For all three positions, the sample size will decrease from year 1 to year 3 because some players' careers lasted for only 1 or 2 years. Further, for the six combine years under study, complete data sets were available for only years 1999-2002. Analysis for combine year 2003 includes performance analysis for years 2003 and 2004; combine year 2004 includes performance analysis for year 2004.

As regards predictor and performance criteria, review by an institutional review board was considered a nonissue because all data reported in this study were retrieved from public access domains. In addition, the names of individuals included in this study are not revealed.

Statistical Analyses

Statistically, we investigate the null hypothesis that each NFL combine measure is unrelated to each NFL performance measure, for each of three offensive positions. Formally, we test

\[ H_{0}: \rho = 0 \]

vs.

\[ H_{a}: \rho \neq 0 \]

multiple times. For instance, one correlation computed is that which quantifies the association of vertical jump (a combine measure) with NFL year 1 salary (a performance measure) for all quarterbacks drafted during the 6-year study period. The null hypothesis assumes that a zero correlation model will describe the data set in the aggregate. We will use a 5% level of significance \( p \leq 0.05 \) for ascertaining statistically significant correlations.

Results

Quarterbacks

Table 2 presents the correlations of the combine exercises with the various aspects of quarterback success, shown in the left column. For instance, the combine exercise with the strongest relation to draft order was vertical jump \((r = -0.36, \text{ in the top line of Table 2})\), an outcome that is statistically significant at \( \alpha = 0.05 \). The negative sign on the correlation indicates that higher vertical jumps are associated with lower (i.e., earlier) draft order picks. The broad jump was also significantly correlated with draft order \((r = -0.32)\), signifying that longer broad jumps are associated with earlier draft order selections. Note that neither vertical jumping nor broad jumping abilities are related to any of the other quarterback success measures. In fact only 3 of the 80 correlations in Table 2 reach statistical significance at the 0.05 level, a result consistent with a random chance model. That is, we would theoretically expect to find about four significant results in a table of this size, when, in fact the population of combine measures has zero correlation with the quarterback success measures.

Further, one of the three correlations (year 1 quarterback rating and 40-yard sprint time) is in the "wrong" direction; that is, higher quarterback ratings are associated with slower times in the 40-yard sprint. (Some correlations in Table 2 are based on fewer than 68 pairs, because all quarterback prospects did not complete each individual combine exercise, and not all drafted players had careers that lasted at least 3 years.)

We therefore conclude that the combine fails to show a consistent significant relationship with the measures of success for quarterbacks chosen for this study: draft order, quarterback rating, games played in, and salary.

Wide Receivers

Table 3 presents the correlations of the combine exercises with the various aspects of wide receiver success, shown in the left column. Note that the bottom three rows of the table have average yards per reception instead of quarterback rating, because we are now considering results for wide receivers, not quarterbacks.

Looking at draft order for wide receivers (row 1 of Table 3), it would seem that sprint times at 10 and 20 yards are directly related to order of selection in the draft. That sprint times for wide receivers should correlate with draft order seems almost intuitive, yet sprint times correlate with only one of the other potential variables, namely, the third-year salary correlation with the 40-yard sprint This correlation, 0.28, is in the "wrong" direction, as slower sprint times are associated with larger third-year salaries. Games played and average yards per reception seem unrelated to all of the combine variables.

As with Table 2 and the quarterback data, only 4 of the 80 correlations generate, values less than 0.05, again a result consistent with a random chance model. We therefore conclude that like the quarterback data set, the various combine measures are unrelated to wide receiver success measures.

We note here that there really aren't three separate and independent measurements for the sprint times. Radier than
Given the findings of this study, however, one should not infer that mental constructs are unrelated to athletic performance. Despite its advocates and some perhaps self-serving statements from Wonderlic Inc., the WPT (Wonderlic Personnel Test) has utility in more general employment situations, the WPT seems to lack validity in this athletic-oriented setting. This conclusion is based on the combination of various measures we examined in this study, the WPT is unrelated to NFL success. Although it may have utility in more general employment situations, the WPT seems to lack validity in this athletic-oriented setting despite its advocates and some perhaps self-serving statements from Wonderlic Inc.

Running Backs

Table 4 gives the correlations of various aspects of success for running backs with the combine exercises. It is immediately apparent that the set of sprint times for the running back prospects is related to the success measures (the bottom three rows of Table 4 now represent average yards per carry in seasons 1, 2, and 3). Except for games played in year 3 and average yards per carry in year 3, the great majority of the sprint measures correlate with the 10 measures of running back success, at a < 0.05. All significant correlations have the "correct" plus or minus sign as well.

The results are quite different for the other seven combine exercises. Only 7 of the 70 correlations in the table (beginning with the vertical jump column) are statistically significant at α = 0.05, and one of these (average yards per carry in year 1 and time for the cone exercise) is in the "wrong" direction. These few significant correlations do not repeat themselves in any year-to-year sense or other consistent pattern, suggesting a random or hit-and-miss pattern for combine measures other than the sprint times.

Summarizing the correlations of the various combine events with the success measures for the three "skill" positions in the NFL (quarterback, wide receiver, and running back), we find that combine exercises are not correlated with success for quarterbacks and wide receivers; for running backs, strong correlations between the three (highly multicolinear) sprint times are found and seem to be predictive of NFL success. However, the other seven measures for running backs bear little or no relation to the success variables.

DISCUSSION

The combine clearly enjoys a significant degree of media hype and the rapt attention of professional football fans. Nonetheless, at least in the sample studied in this research, it clearly lacks any meaningful degree of predictive validity, save for the relationship between sprint times and running back success. One possible explanation for the lack of relationship between overall combine performance and NFL performance lies with the numerous "prep courses" and other learning materials that exist to assist athletes in preparing for the combine. A Google search of "NFL Combine Training Programs" reveals numerous vendors who all claim to improve combine performance, including well-known programs such as the Parisi Speed School, Pro Combine Training, Makeplays.com, Team Hammer, Xplosive Speed, High Intensity Training, Velocity Sports Performance, Mackie Shilstone’s PEP, Speed in Sports, Elite Athlete Training Systems, and many others. As expected, vendors claim impressive gains for those who sign on with them. In a sense, such types of assistance available to the aspiring professional athlete are similar to the materials and courses that exist for graduate school hopefuls preparing for the Law School Admissions Test or Graduate Management Admissions Test for the would-be law or business school student. Like the aspiring law school applicant, the combine hopeful may "burn the midnight oil" by "cramming" at one of several facilities designed specifically to improve combine performance. In the athlete's case, however, the purpose of these programs is on improving physical, not mental, skills, with the notable exception of preparing for the Wonderlic Personnel Test (at least one combine preparation program, Mackie Shilstone’s Performance Enhancement Program, includes Wonderlic test preparation in its program).

Of course, a key question is whether attending combine preparation programs actually does improve an athlete's combine performance. Other than the marketing claims made by the vendors themselves, there is no scientific evidence that their preparation improves combine performance. If, in feet, they do, such programs and materials could serve as "performance equalizers," thus diluting the performance differences among combine attendees. For example, assume that an athlete's 40-yard dash time is only "average" but that he attends a training camp and does, in feet, improve his technique and time for the combine. In the 40-yard dash, he now competes effectively with the athlete whose precombine time was already in that range.

As with previous research investigating the mental assessment of NFL candidates, this research also failed to show a relationship between the WPT and NFL success. Viewing the results as a whole, we conclude that for the years, positions, and performance criteria we examined in this study, the WPT is unrelated to NFL success. Although it may have utility in more general employment situations, the WPT seems to lack validity in this athletic-oriented setting despite its advocates and some perhaps self-serving statements from Wonderlic Inc.

Given the findings of this study, however, one should not infer that mental constructs are unrelated to athletic
performance. Although in this study the WPT failed to show a relationship with NFL performance, other studies have found correlations between psychological attributes and athletic performance. In a research review of personnel selection practices in athletics (4), Humara found a number of psychological constructs related to various athletic endeavors, including aggression, leadership, reachability, and self-confidence. Further, affiliation and conformity may predict athletic performance; for instance, an athlete low in conformity and affiliation may not perform well in team sports or under an autocratic coach. On the basis of his review, Humara concludes that in the selection of athletes, in addition to an assessment of an athlete’s past performance and bio (physiological) data, decision makers should make greater use of psychological assessments, including the Athletic Motivation Inventory, the Test of Attentional and Interpersonal Style, and the Profile of Mood States. Another mental construct that has received considerable attention in the literature is anxiety and its relationship with athletic performance. A meta-analysis of the relationship among multiple forms of anxiety and athletic performance found self-confidence to be the strongest and most consistent predictor of performance (3).

Given the research that has been published on the relationship between various constructs of cognitive ability and athletic performance, and the obvious monetary risks of false-positives and false-negatives when selecting professional football players, it is surprising that the NFL has not used a more sophisticated approach to the measurement of mental constructs for combine participants. Possibly, the use of higher-level personality and cognitive ability assessments—for example, those discussed in this section—would benefit the NFL by improving the fit between the combine hopefuls and the job of professional football.

Perhaps a more plausible explanation for the lack of correlation between combine performance and NFL performance is that combine exercises simply measure athletic skill and not actual football-playing ability. The difference between skill and ability has been cited in human resource management theory. A skill narrowly focuses on a particular task, whereas ability more broadly relates to a multiple set of tasks, or competency (10). For example, a running back may possess raw speed (e.g., running full speed from point A to point B, such as a 40-yard dash) but lack running ability in an actual contest, where he must avoid tacklers by often reacting instantaneously with cuts and jukes and “reading the defense.” Speed does not automatically make an excellent running back, just as command of the English language does not necessarily translate into an effective public speaker. Regardless of a player's position, football is a complex sport, demanding multiple physical and mental skills for successful performance. For example, one study identified 14 performance dimensions for the quarterback position through a job analysis, including coordination, running speed, vision, ability to learn, football sense, physical reactions, stress tolerance, and being a team player (5).

The primary focus of combine exercises is on speed, strength, and agility (athletic skills) in a noncontact environment without the conditions and challenges a player finds in an actual game situation. The combine does not measure other characteristics important for player success, such as motivation, decisiveness, and ability to work effectively in a team (7). Given the substantial gap between the skills assessed at the combine vs. the skills actually required to successfully compete at the professional football level, the lack of relationship between combine performance and NFL performance is, therefore, not surprising.

PRACTICAL APPLICATIONS

Although the NFL combine remains a key element in the player-selection process, the overall findings from this study call into question the usefulness of the majority of exercises and suggest that the significant sums of money spent on the combine are not wisely spent. Conceivably, the predictive power of combine exercises could prove stronger for other draft classes and/or NFL positions other than those examined in this study, but this study suggests otherwise.

The 10-, 20-, and 40-yard dash exercise results do correlate significantly with NFL performance, but only for the running back position. This exercise should continue to be used to draft running backs, whether it is a combine activity or a test given under other conditions, such as Pro Day workouts. The vertical jump, shown to correlate with football player performance among college players in other studies, was not predictive of success at the professional level in this study.

Although this study calls into question the usefulness of the majority of combine exercises, other aspects of the combine may nonetheless prove beneficial for drafting players. Several important activities that take place at the combine are not reported in this study, because their results are generally not made public. They include position-specific drills (e.g., pass routes for wide receivers), physical measurements (height, weight, arm and hand length, and body fat percentage), team interview, Cybex test (a measure of flexibility and joint movement), injury evaluation, and urine test. Absent a scientific analysis, however, one can only speculate as to the usefulness of these exercises as well.

In the broader context of the draft process, it should be noted that a potential player's physical performances at the combine are not the only criteria on which draft decisions are made. Hall of Fame San Francisco 49er coach Bill Walsh, for example, with the aid of assistant coaches and scouts, collected and analyzed game film as well as data gleaned from interviews with the players' coaches and trainers. They also conducted the team’s own interviews, intelligence tests, and personality tests. Walsh developed a comprehensive profile on players that would be given draft consideration by the
49ers (12). Extensive player profiles such as those prepared by Walsh are likely prepared by other NFL teams as well and may provide significant data, beyond combine data, on which a draft decision will be made.

In addition, as mentioned previously, the combine affords an opportunity to assemble and analyze the majority of key professional prospects at the same location, at the same time. Thus, for practical purposes, the combine is an efficient process from the standpoint of facilitating social interactions among players and team personnel.

To conclude, NFL team owners and managers must use caution in determining the value of the combine in drafting players and the weight and importance given to various combine exercises. Consideration should be given to eliminating those exercises and tests that prove useless in predicting player success while restructuring the combine to include only those activities that actually predict success at the professional playing level. This suggests that an overhaul of the combine process may be due and that the inclusion of contemporary human resource selection practices should be an important part of the combine restructuring.

[Reference]

REFERENCES

[Author Affiliation]
FRANK E. KUZMITS AND ARTHUR J. ADAMS
Department of Management, College of Business, University of Louisville, Louisville, Kentucky

[Author Affiliation]
Address correspondence to Frank E. Kuzmits, kuzmits@louisville.edu.
22(6)/1721-1727
Journal of Strength and Conditioning Research
© 2008 National Strength and Conditioning Association

References
References (13) Cited by (7)

Indexing (document details)
Subjects: Coaches & managers, College football, Studies, Athletes, Skills, Intelligence tests, Exercise, Data collection