

Running head: WONDERLIC AND NFL PERFORMANCE

Not much more than *g*? An examination of the impact of intelligence on NFL performance

Brian D. Lyons

California State University, Fresno

Brian J. Hoffman*

University of Georgia

John W. Michel*

Towson University

[In press, *Human Performance*]

Footnote and Acknowledgements

*Second authorship was determined alphabetically as both authors contributed equally to the delivery of this manuscript. We would like to thank Kevin J. Williams, Walter C. Borman, and two anonymous reviewers for providing helpful comments in improving the overall quality of the manuscript. Portions of this manuscript were presented at the 20th and 21st Annual Meetings of the Society for Industrial and Organizational Psychology.

Correspondence concerning this manuscript should be addressed to Brian D. Lyons, California State University, Fresno, Craig School of Business, 5245 N. Backer Avenue, MS PB 7, Fresno, CA 93740-8001; e-mail: blyons@csufresno.edu.

Abstract

The purpose of this study was to determine the efficiency and equity of general mental ability (GMA) in a non-traditional employment setting, professional football. The National Football League (NFL) uses a measure of GMA, the Wonderlic Personnel Test, to evaluate potential draftees in an assessment-style environment. A total of 762 NFL players, represented from three draft classes, were included in our sample. In terms of efficiency, results indicated that GMA was unrelated to: (a) future NFL performance; (b) selection decisions during the NFL Draft; and (c) the number of games started in the NFL. In regards to equity, differential prediction analyses by race suggested only the existence of intercept bias. The implications of these findings to the NFL and the selection literature are further discussed.

KEY WORDS: intelligence; general mental ability; job performance; selection.

Not much more than *g*? An examination of the impact of intelligence on NFL performance

Within a selection system, a predictor can be evaluated in terms of its efficiency and equity (Murphy, 2002). Efficiency refers to how well a measure is related to performance-related criteria such as productivity or profit. On the other hand, equity emphasizes the extent to which a measure engenders subgroup bias or discrimination against minority groups. Consequently, one can evaluate a measure according to these criteria and establish trade-offs to facilitate decisions for inclusion (Murphy, 2002).

Voluminous empirical research supports the validity of general mental ability (GMA) as a predictor of job performance (e.g., Schmidt & Hunter, 1998; Viswesvaran & Ones, 2002). To that end, GMA consistently demonstrates the strongest criterion-related validity of existing predictors (Murphy, 2002; Ree, Earles, & Teachout, 1994; Schmidt, 2002; Schmidt & Hunter, 1998). Specifically, meta-analytic estimates of this relationship typically result in an uncorrected validity coefficient of approximately .30 (Bobko, Roth, & Potosky, 1999) and a corrected validity coefficient of .51 (Schmidt & Hunter, 1998). However, ethnic group differences associated with measures of GMA have garnered substantial attention (e.g., Herrnstein & Murray, 1994). Given the legal landscape and a demographic shift to a more diverse workforce (Offermann & Gowing, 1993; Outtz, 2002), the potential for adverse consequences to minorities is a probable deterrent against the use of GMA in selection contexts (e.g., Outtz, 2002; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997). Despite these concerns, GMA is still a frequently assessed predictor in a wide range of employment settings.

For example, the National Football League (NFL) administers a GMA instrument to collegiate prospects at its annual NFL Combine, held approximately two months before the NFL Draft. Consistent with more traditional employment settings, the incorporation of GMA in this

context is likely predicated on the assumption that more intelligent players will acquire positional-related knowledge more quickly and be able to use this knowledge to perform better during a game (Schmidt, 2002). Through its annual use, GMA is *assumed* to be a valid predictor in this context. Similar to traditional employment settings, demonstrating the validity of predictors should also be important to the NFL, especially when one considers the cost associated with signing high draft picks (e.g., 1st or 2nd round). From the perspective of selection utility, a single Type 1 error could cost a team millions of dollars and be a detriment to overall team performance.

Despite the existing criterion-related validity evidence for GMA, it is unclear the extent to which these findings will generalize to a less traditional employment setting such as the NFL. The failure of existing research to examine the relevance of GMA in an occupation dominated by roles requiring employees to possess high levels of muscular strength, cardiovascular endurance, and movement quality (cf., Hogan, 1991) represents an important omission from the literature. Accordingly, the purpose of this study is to determine if GMA is (a) *efficient* in predicting future performance and (b) *equitable* in regards to subgroup differences in the prediction of performance.

In terms of efficiency, Murphy (2002) stated that GMA is “likely to be correlated with performance in virtually any job, in part because all jobs call for some learning, judgment, and active information processing” (p. 176). The positive relationship between GMA and decision-making (Gully, Payne, Koles, & Whiteman, 2002), problem solving (Stevens & Campion, 1999), and the acquisition of job knowledge (Schmidt, 2002) has been empirically demonstrated in the literature. Although success in the NFL is likely a function of physical ability, GMA may also be an important determinant of performance. Specifically, NFL performance requires that

players learn complex schemes and playbooks, understand the tendencies of the different teams they play each week, and quickly process information and adjust their play multiple times during the course of a single game. Accordingly, consistent with prior research indicating that GMA is related to job performance regardless of profession, setting, task composition, and level of job complexity (Schmidt & Hunter, 1998; Schmidt, Hunter, & Pearlman, 1981), there is reason to believe that GMA is an important predictor of performance in the NFL.

To assess GMA, the NFL has been administering the Wonderlic Personnel Test (WPT) at the Combine since the 1970s (Wonderlic, Inc., 2004). An on-line publication produced by Wonderlic, Inc. (2004), *HR Measurements*, stated that the WPT is an essential assessment during the Combine because “smarter people make better teammates and deliver more wins to the team.” Given the import of “on-the fly” processing of information, the complex schemes associated with present day professional football, and the research suggesting that GMA is related to performance across job settings (Schmidt & Hunter, 1998; Schmidt et al., 1981), it is expected that GMA will possess a positive, nonzero relationship with NFL performance. Thus, the following hypothesis is offered:

Hypothesis 1: GMA will be positively related to NFL performance.

A meta-analysis by Schmidt and Hunter (1998) demonstrated that although GMA is related to performance across levels of job complexity, the relationship is stronger in more complex jobs characterized by greater cognitive demands. In relation to this study, certain positions may require more problem solving and decision-making ability (e.g., quarterbacks) than other positions that primarily rely on physical attributes and instinct (e.g., running back). For example, quarterbacks must digest an offensive playbook and recall the assignments and routes of other positions during game situations. Additionally, they should be prepared to read defensive

alignments and react to coverage within split seconds of the play. These tasks seem to entail a higher level of learning comprehension, problem solving, and decision-making than other offensive and defensive positions. As a result, to the degree that some positions engender less complexity than others, the relationship between GMA and performance may vary by position. Based on these assertions, we offer the following hypothesis:

Hypothesis 2: Position type will moderate the relationship between GMA and NFL performance, such that the relationship will be stronger for quarterbacks than other positions.

Supplemental Research Questions

In addition to the two hypotheses, three research questions are presented to further ascertain GMA's efficiency and equity in the NFL. The first question pertains to determining GMA's equity through a differential prediction analysis by race. The last two questions attempt to determine GMA's efficiency in predicting selection in the NFL Draft and the number of games a player starts during a given NFL season. Each is further explained below.

Since we are drawing on the same population as traditional employment studies (e.g., Roth, BeVier, Bobko, Switzer, & Tyler, 2001), subgroup discrepancies in GMA should exist and be demonstrated in this study. Therefore, we refrain from formally hypothesizing racial differences in GMA – instead, we would like to determine the extent to which GMA predicts NFL performance differently for African-Americans and Caucasians. The presence of bias can be determined by examining the differences in intercepts and slopes between two subgroups within a moderated regression model (Bartlett, Bobko, Mosier, & Hannon, 1978). Intercept bias exists when there are significant differences in subgroup performance. Conversely, slope bias is present when a test is predicting better for one subgroup than the other (Schmitt & Chan, 1998).

In general, few studies have detected racial differences in intercepts or slopes (Bartlett et al., 1978). When differences do occur, however, intercept bias is found more often than slope bias (Bartlett et al., 1978), most likely due to a lack of statistical power when constructing an interaction term to test for slope differences (Aguinis, Beaty, Boik, & Pierce, 2005). Although intercept differences are more often noted, they usually lead to the overprediction of minority group performance (Schmitt & Chan, 1998). Despite this evidence, examining differential prediction in this study is warranted because of sample and context differences. Similar to traditional settings, slope bias may exist due to existing subgroup differences in GMA (Bobko et al., 1999). On the other hand, since the NFL is primarily composed of African-American athletes (Bivens & Leonard, 1994), intercept bias may be more prevalent because African-Americans are the dominant subgroup. Another difference is that the criterion in this context entails a high degree of physical ability that traditional occupations may not require for successful performance. Therefore, studies done in traditional settings where Caucasians are the dominant subgroup and the criterion is not largely manifested in physical ability, the results found for predictive bias (or lack thereof) may not generalize to this setting. Consequently, a research question is offered to determine such results.

Research Question 1: Does GMA differentially predict NFL performance by race?

It has been reported that some NFL teams question the validity of the WPT (Mulligan, 2004) while other teams consider the test results a vital part of their selection processes (Merron, 2002). Given the apparent disagreement among league decision makers over the utility of the WPT, this study sought to examine the influence WPT scores have during the draft selection process. If NFL team decision makers largely weigh the WPT when making selection decisions and subgroup differences manifest, adverse impact may be present within the selection process.

On the other hand, if the WPT is not utilized during the draft process but subgroup differences exist, one should question the utility of this instrument especially if the measure is unrelated to future performance. To resolve such an issue, we propose the following research question to determine the relationship between GMA and selection in the NFL Draft.

Research Question 2: Does GMA affect selection in the NFL Draft?

A Wonderlic, Inc. press release (2005) suggested that a positive relationship exists between the WPT and the number of games a prospect starts during an NFL season.

Unfortunately, we were unable to locate any empirical research substantiating this assertion. It is reasonable to believe that in the competitive environment of the NFL, playing time will be a direct function of a player's performance. In other words, those draftees who play in more games have a greater probability of attaining performance-related statistics. In turn, those draftees who elicit immediate playing time may be those who digested the playbook quicker and more efficiently. After all, GMA has been shown to be causally related to the acquisition of job knowledge (Schmidt, 2002). Conversely, the number of games started may be due to other facets aside from GMA. For example, where the prospect was selected in the draft may influence how many games he starts – that is, those selected higher (i.e., 1st or 2nd rounds) may play more to justify their draft selection. Moreover, the degree of physical ability may affect how many games a prospect starts during his NFL career. Because of these conflicting sentiments, we offer a final research question to determine if GMA is related to the number of games started.

Research Question 3: Does GMA influence the number of games started in the NFL?

In sum, the purpose of this study is to investigate GMA's efficiency and equity in the NFL. Although voluminous research supports the consistency of the GMA-job performance

relationship across job settings, it is unclear as to whether these findings will generalize to this context. As a result, we will examine: (1) the relationship between GMA and NFL performance; (2) whether the GMA-performance relationship is stronger for quarterbacks than other positions; (3) if GMA differentially predicts NFL performance by race; and (4) the impact of GMA on draft selection and number of games started.

Method

Participants

A total of 762 football players, 256 selected in the 2002 NFL Draft, 257 in the 2003 NFL Draft, and 249 in the 2004 NFL Draft, were included in this study. All traditional offensive and defensive positional players were represented in the sample. However, due to low sample sizes, kickers and punters were excluded. As depicted in Table 1, draftees consisted of 528 African-Americans (69.3%), 215 Caucasians (28.2%), and 19 Other (2.5%).

 Insert Table 1 about here

Measures

General mental ability (GMA). Prior to the annual NFL Draft, the NFL Combine provides owners and coaches an opportunity to evaluate prospects' physical and mental ability. In this setting, GMA is measured with the WPT. Designed as a speeded test, the WPT is a 12 minute timed test consisting of 50 multiple choice and short answer items that purport to measure verbal, numerical, general knowledge, analytical, and spatial relations. Test scores range from 1 to 50. Adult working class norms of the WPT indicate a mean score of 21.75 and a standard deviation of 7.6 (Wonderlic, Inc., 2002). In terms of reliability, internal consistency estimates range from .88 to .94, test-retest values range from .82 to .94, and alternate form

estimates range from .73 to .95 (Wonderlic, Inc., 2002). WPT data for all draftees were collected from secondary sources, *CBS.sportsline.com* and *NFLdraftscout.com*.

NFL performance. To serve as criteria, statistical data from the first three years of performance in the NFL were collected from *NFL.com* and *Stats, Inc.* Collecting three years of performance data closely approximates a player's average tenure in the NFL, which is about three and a half seasons (NFL Players Association, 2007). In addition, examining multiple years of performance enables us to depict the GMA-performance relationship longitudinally and apply relevant GMA-performance theories, such as Murphy's (1989) transitional/maintenance stage theory, to explain the results.

The following NFL seasons served as data for each appropriate draft class: 2002, 2003, and 2004 seasons for the 2002 draft class; 2003, 2004, and 2005 seasons for the 2003 draft class; and 2004, 2005, and 2006 seasons for the 2004 draft class. Depending on the draft class, each year of statistical data represented Year 1, Year 2, or Year 3 of NFL performance. Before gathering the performance data, a priori decision rules were imposed. First, offensive linemen were excluded in the establishment of future NFL performance because we could not locate positive performance statistics such as pancakes or takedowns. After this decision was made, our goals were to include those statistics that were not redundant within position (e.g., we included total tackles instead of including solo and assisted tackles) and accurately portrayed performance (e.g., running back performance is typically evaluated by rushing and receiving production). Performance criteria collected by position are summarized in the Appendix.

Procedure

Data were collected from nationally recognized sports Web sites such as *CBS.sportsline.com*, *NFL.com*, and *NFLdraftscout.com*. This information included each player's

name, position, WPT score, draft selection number in either the 2002, 2003, or 2004 NFL Draft, and performance in the NFL.

To ascertain and compare the relationship between the WPT and performance across positions, all performance criteria within each position were standardized. We negatively coded raw scores for adverse performance criteria such as fumbles and interceptions for offensive positions (e.g., a value of 5 for fumbles was changed to -5). Subsequently, the raw scores for each draft class were transformed into z-scores within each position per year. We then summed all of the representative z-scores and divided this value by the number of performance criteria that position encompasses to create an overall averaged estimation of their performance per year (i.e., Year 1, Year 2, and Year 3). In general, each z-score represents a player's performance relative to other players at the same position. Player injuries were accounted for within the z-score formation as missed time. That is, if a player missed games for reasons related to issues such as health or suspension, we gathered objective data up to that time of injury. Such rationale is not uncommon in the NFL. For example, at the end of a season, missed time due to injury or suspension is not usually taken into account when determining individual awards such as Pro Bowl invitations or Most-Valuable-Player type of awards. If a player is frequently injured from year-to-year (indicating a trend), the player will most likely gain the title of "injury prone", which will affect both his performance statistics and career viability in the NFL. Conversely, if a player was injured only once throughout the three year period, the missed time associated with the injury would not have likely influenced his z-score in the other two years or his overall z-score across the three years (which will be explained below). Number of games started was not included in this performance metric because of the decision to control for this variable in one of the hypothesis testing.

To create an averaged performance value across all three years, we summed the z-score totals for each year and divided this value by the total number of performance criteria. In sum, four z-score performance values were created for each player: Year 1, Year 2, Year 3, and overall averaged performance. Similarly, z-score values for games started were created for Year 1, Year 2, Year 3, and an overall averaged value across all three years. One undesirable result of using z-scores, however, is that half of the scores in the distribution will be negative (Murphy & Davidshofer, 2001). Since we did not want to use or interpret negative performance values, the four z-score performance composites and games started values were transformed into T-scores to produce the final performance estimates.

A final note pertains to the construction of our performance measures. Before creating unidimensional performance criteria, we examined the correlation matrices between each statistical indicator by position and performance year. All inter-item correlation means within position across performance years were moderately high: quarterbacks, .48 - .61; running backs, .58 - .59; wide receivers, .62 - .69; tight ends, .59 - .73; defensive linemen, .41 - .51; linebackers, .38 - .51; and defensive backs, .42 - .46. In very few occasions, however, some of the correlation coefficients within position and performance year were low, with the weakest being in the quarterback position [Year 3, yards per attempt and completions ($r = .01$)], defensive linemen [Year 2, forced fumbles and interceptions ($r = .01$)], and linebackers [Year 2, forced fumbles and interceptions ($r = .04$); Year 3, interceptions and sacks ($r = .02$)]. Despite these few instances of weak correlations, we feel that our unidimensional approach was preferable to a differential weighting approach given the need for parsimony and efficiency (Bobko, Roth, & Buster, 2007; Schmidt & Kaplan, 1971).

Data Analysis

Bivariate correlations were calculated to determine the GMA-NFL performance relationship. Also, partial correlations were estimated to control for the number of games started. Next, moderated regressions were computed to determine whether the strength of the relationship between GMA and performance varied by position or race. Finally, bivariate correlations were computed to ascertain whether GMA was related to selection in the NFL Draft and the number of games started.

Results

Hypotheses Testing

Descriptive statistics and bivariate correlations among study variables are presented in Table 2. Before venturing into the hypotheses testing, it was interesting to find that the WPT adult working class norms and the results from our sample were quite similar (population norms, $M = 21.75$, $SD = 7.6$; sample, $M = 21.04$, $SD = 7.15$). When offensive linemen were included, the two results were nearly identical (sample, $M = 21.61$, $SD = 7.13$). These results lend confidence to the notion that this work context contains individuals that are relatively similar, in terms of GMA, to the traditional employment population.

 Insert Table 2 about here

The first hypothesis stated that GMA would be positively related to NFL performance. Results indicated that the WPT was unrelated to any of the NFL performance criteria: Year 1 ($r = -.02$), Year 2 ($r = -.02$), Year 3 ($r = -.06$), or overall ($r = -.04$).

Since three years of performance data were gathered, it was possible to examine this relationship over time with varying levels of the predictor (cf., Ployhart & Hakel, 1998). As

depicted in Figure 1, NFL performance was regressed on three levels of WPT scores: at the group mean and one standard deviation above and below the mean. Consistent with the results presented above, the beta weights for each level of the WPT were not significantly related to performance. The figure is interesting from the perspective that lower levels of GMA were related to higher performance, on average, and less of a decrement over time than mean or higher levels of GMA. In addition, these results contrast Murphy's (1989) theory, which asserts that GMA is more important during the transitional stages of performance than in the maintenance stages.

Insert Figure 1 about here

In order to rule out possible attenuation to the GMA-performance relationship due to playing time, partial correlations were calculated to control for the number of games started for each year of performance. Similar to the previous results, the partial correlations were not significant for any performance year: Year 1 ($pr = .03$), Year 2 ($pr = -.01$), Year 3 ($pr = .01$), and overall ($pr = .01$). Thus, the number of games started has little impact on the relationship between GMA and NFL performance. Based on these results, the first hypothesis was not supported.

The second hypothesis stated that the strength of the relationship between GMA and NFL performance would be stronger for quarterbacks than other positions. To test this hypothesis, we performed two statistical analyses. First, bivariate correlations between WPT scores and NFL performance were conducted by position type and performance year. The results generated were quite similar across all performance years; consequently, we only report those with the overall performance criterion. As shown in Table 3, no significant positive correlations were detected

between the WPT and NFL performance for any position. In particular, no significant relationships were found within the quarterback position, although the strongest coefficient was found in Year 3 performance ($r = .23, p = .32$). Significant negative correlations were detected, however, within the tight end and defensive back positions. For the tight end position, significant negative correlations were found for two of the four performance criteria: Year 3 ($r = -.35, p < .05$) and overall ($r = -.28, p < .10$). For the defensive back position, significant negative relationships were found across all criteria: Year 1 ($r = -.18, p < .10$), Year 2 ($r = -.19, p < .05$), Year 3 ($r = -.19, p < .10$), and overall ($r = -.21, p < .05$). Together, these results initially suggest that the GMA-NFL performance relationship was not stronger for quarterbacks than other positions.

 Insert Table 3 about here

Next, a moderated regression was conducted using the quarterback position as the referent group. Race was included as a covariate to negate any subgroup differences in GMA. Following Jaccard and Turrisi (2003), each regression model was constructed by first entering the covariate, race, and then the main effects, group centered WPT score and the dummy-coded positions (without quarterbacks), followed by the interaction terms. As illustrated in Table 4, only the results gathered for overall NFL performance are displayed because these results were identical to those found across performance years (i.e., Year 1 through 3). As can be seen, neither the main effects nor product terms were significant, leading to the conclusion that the quarterback position did not engender a significantly different slope than the other positions. Therefore, the second hypothesis was not supported.

 Insert Table 4 about here

Analyzing the Research Questions

The first research question inquired about whether GMA differentially predicts NFL performance by race. Because the first hypothesis was not supported, which found that GMA was unrelated to performance, differential prediction was less likely to exist. However, since the objective of this study was to evaluate GMA's efficiency and equity in this context, it was necessary to examine differences in intercepts and slopes by race to achieve the latter goal. Before conducting this analysis, we examined subgroup differences on WPT scores, which included offensive linemen, to substantiate the belief that such differences exist in this context. Consistent with past GMA research (e.g., Bobko et al., 1999), Caucasians ($M = 27.34$, $SD = 6.22$) scored significantly higher than African-Americans ($M = 19.32$, $SD = 6.02$) on the WPT, $t(649) = -15.23$, $p < .001$. After this was established, a moderated regression was performed that regressed each performance year on WPT scores, race, and the interaction term. As recommended by Bartlett et al. (1978), each regression model was constructed by first entering the predictor, group centered WPT score, then race (dummy-coded), followed by the interaction term.

Similar to the other moderated regression, Table 5 only depicts the results gathered for overall NFL performance because these results were nearly identical to those found across performance years. In general, the WPT did not differentially predict performance by race. However, race did significantly predict performance, which indicates the existence of intercept bias. Specifically, when included in step 2, race significantly predicted performance in Year 1 ($\beta = -.14$, $p < .01$), Year 2 ($\beta = -.10$, $p < .05$), and overall ($\beta = -.12$, $p < .05$). This result was originally reported in Tables 2 and 4. The correlation matrix revealed a significant relationship

between race and performance in Year 1 ($r = -.12, p < .01$), Year 2 ($r = -.08, p < .10$), and overall ($r = -.10, p < .05$). Similarly in Table 4, when race was placed first in the hierarchical order, this predictor was significantly related to Year 1 ($\beta = -.14, p < .01$), Year 2 ($\beta = -.10, p < .05$), and overall performance ($\beta = -.12, p < .01$). In all, these results suggest that African-Americans performed significantly better than Caucasians during these years. However, it should be noted that this effect diminished over time such that no significant differences were detected in Year 3 of performance.

 Insert Table 5 about here

The second research question speculated whether a relationship existed between GMA and selection in the NFL Draft. As depicted in Table 2, the bivariate correlation between WPT scores and draft selection was not significant ($r = .05$). Therefore, these data suggest a null relationship between how well a prospect scores on the WPT and where he is selected in the NFL Draft.

The final research question pertained to whether GMA is related to the number of games started within the NFL. Similar to the previous results, the bivariate correlations between WPT scores and the number of games started across the performance domain were not significant:

Year 1 ($r = -.06$), Year 2 ($r = -.02$), Year 3 ($r = -.08$), or overall ($r = -.06$).

Supplemental Analyses: Supporting the Use of Objective Performance

Additional analyses were performed to support the use of objective statistics in defining performance. More specifically, the ultimate criterion that represents job performance in the NFL may encompass objective (e.g., statistical) and contextual performance behaviors (e.g., Borman & Motowidlo, 1993). In this employment context, however, objective performance is

most likely the primary antecedent to future compensation; thus, the presence of convergent evidence between objective performance and other career-related variables of interest would validate our measurement of performance within this setting.

Since we gathered data from the first three years of performance, we examined relevant career-related variables during a player's third and fourth year in the NFL. The variables selected, NFL employment status (i.e., retention) and salary, closely approximate an individual's career-related viability in this context. Such data were collected from *nfl.com* or *footballsalaries.usatoday.com*.

It was postulated that if a prospect performed well in the NFL, he would engender a higher probability of being retained within the NFL and compensated appropriately for his performance. However, variance due to where the prospect was chosen in the NFL Draft may influence these relationships. As a result, the decision to control for draft selection was made because of the probability that those that were drafted higher (e.g., 1st or 2nd rounds) would possess a higher probability of being retained in the NFL and collect a larger salary due to the original draft selection-based contract. Moreover, front office personnel may be more willing to provide another chance to a prospect who was drafted higher, inferring that these individuals encompass greater physical abilities and potential. Therefore, examining relationships between the objective criteria and these two variables, absent to any influence of where the individual was chosen in the draft, may permit confidence in our measurement of performance.

Because retention status represents a dichotomous outcome, hierarchical logistic regressions were performed to determine if the performance criteria predicted whether the individual was employed in the NFL at Year 3 and Year 4 above and beyond that of where the individual was chosen in the draft. The results for each of the performance criteria are presented

in Table 6. After controlling for draft selection, the performance criteria accounted for a significant amount of additional variance in NFL status at Year 3 and Year 4: Year 1 performance ($\Delta\chi^2 = 40.15, p < .01$; $\Delta\chi^2 = 50.92, p < .01$), Year 2 performance ($\Delta\chi^2 = 14.94, p < .01$; $\Delta\chi^2 = 48.40, p < .01$), Year 3 performance (Year 4 only, $\Delta\chi^2 = 33.36, p < .01$), and overall performance ($\Delta\chi^2 = 45.18, p < .01$; $\Delta\chi^2 = 96.49, p < .01$).

 Insert Table 6 about here

Partial correlations with listwise deletion were utilized to determine the relationship between future salary and the performance criteria, controlling for draft selection. For Year 3 NFL salary, the partial correlations were significant and in the positive direction: Year 1 performance ($pr = .13, p < .01$), Year 2 performance ($pr = .24, p < .01$), and overall performance ($pr = .27, p < .01$). Similar results were also found for Year 4 NFL salary: Year 1 performance ($pr = .17, p < .01$), Year 2 performance ($pr = .21, p < .01$), Year 3 performance ($pr = .30, p < .01$), and overall performance ($pr = .30, p < .01$).

Finally, the significant correlations between the performance criteria in Table 2 (r_s from .53 to .88, $p < .01$) also supported the construct validity of our measurement of performance. In general, all of these results suggest that our use of objective performance criteria was empirically justifiable to represent performance in this context. Further, performance was found to be related to retention decisions and compensation, thus reflecting similarities to traditional employment settings.

Discussion

The primary purpose of this study was to determine the efficiency and equity of GMA in the NFL. Specifically, this study delineated the criterion validity of GMA overall and by

position and whether racial differences existed in the prediction of performance. The results provide implications and future research directions germane to the GMA-job performance literature.

Review of the Hypotheses

In terms of efficiency, this study examined the relationship between GMA and future NFL performance across and within positions. As indicated by the results, GMA had a nonsignificant relationship across three separate years of performance as well as an overall index of performance. A notable difference between our findings and the extant GMA-performance literature may be attributed to the types of tasks that are performed in this setting and the manner in which we defined performance via statistical indicators. The tasks performed in the NFL are physically challenging with a primary emphasis on strength, speed, endurance, and agility. In order to perform well during a game (e.g., make a tackle, rush for a touchdown), a player must remember the specific play but more importantly, possess the physical ability to successfully perform and accumulate performance statistics.

Another interesting finding was that the relationship between GMA and performance was not influenced by position complexity. Specifically, the validity coefficient of GMA was not stronger for a more cognitively complex position (e.g., quarterback) than other routine or skill-based positions. For example, the running back position, where physical attributes and instinct are thought to be more dominant, demonstrated a more positive relationship with performance than the quarterback position. This could be explained, however, by the probability that quarterbacks will typically start playing regularly sometime after their rookie season in the league. During that rookie season, coaches often give quarterbacks time to adjust to the speed of the game and master the offensive playbook. Accordingly, we found that the relationship

between GMA and performance for quarterbacks was strongest (although not statistically significant) during Year 3 ($r = .23$). Based on this finding, it is plausible to suggest that the relationship between GMA and performance may become stronger for quarterbacks in future years. It is also possible that the physical requirements in this context simply attenuate the impact of GMA on performance.

Other within position correlations yielded notable results. For tight ends and defensive backs, GMA had a significant negative relationship with performance. Although the magnitude of these relationships was modest, they could be explained by the notion that performance for these positions entails more of an emphasis on physical ability and instinct than GMA. A defensive back high in GMA but lacking the requisite physical abilities may take too much time thinking about the play instead of reacting to the play. Based on these results, it may be worthwhile for NFL teams to deemphasize WPT score performance for these positions (especially defensive backs) when making selection decisions.

Coupled with evaluating efficiency, this study sought to determine GMA's equity within this context. Consistent with previous GMA literature (e.g., Roth et al., 2001), Caucasians scored significantly higher on the WPT than African-Americans. Although racial differences existed, the results from first research question indicated that the WPT does not differentially predict performance by race. This finding was not surprising as GMA possessed a near zero relationship with performance. However, intercept differences were found such that African-Americans performed significantly better than Caucasians in three of the four performance criteria. These results are atypical to most GMA-performance studies. For example, a recent meta-analysis (McKay & McDaniel, 2006) that studied racial differences in job performance found that Caucasians typically performed better than African-Americans regardless of whether

the criterion was measured subjectively or objectively. This employment context is unique from the perspective that the dominant racial subgroup is African-Americans and not Caucasians (Bivens & Leonard, 1994). As a result, the skewed distribution of race may have contributed to this finding. Another plausible explanation could be that those who display greater physical ability during the Combine are African-American athletes, which may then lead to the skewed distribution of race in the NFL and the higher probability of detecting intercept differences.

The results from the second research question indicated that GMA was unrelated to selection decisions during the NFL Draft. Consequently, NFL teams, on average, do not utilize WPT scores systematically in their evaluation decisions. On the same note, adverse impact is not of concern in this setting because teams do not utilize these scores in their selection decisions. However, teams may decide to use the WPT to screen for outliers – that is, to examine players who score extremely low on the measure and consider these results on a case-by-case basis. If using this case-by-case selection strategy with the WPT, teams should be cognizant of the fact that the WPT was found to be unrelated to future NFL performance, thus limiting its utility within this context.

The last research question found that GMA was unrelated to the number of games started within the NFL. As a result, high GMA does not enable a prospect to achieve more playing time, which conflicts with Wonderlic, Inc.'s (2005) assertion that smarter players receive more playing time. As stated before, other individual difference variables such as physical ability may possess a stronger relationship with the number of games started than pure GMA. Also, situational variables can be considered as another determinant of playing time. For example, where a prospect was chosen in the NFL Draft was found to influence the number of games started across all three years: Year 1 ($r = -.45, p < .01$), Year 2 ($r = -.48, p < .01$), and Year 3 ($r =$

-47, $p < .01$). These results supplement all of our previous findings that GMA is a questionable predictor of performance in this setting.

In sum, the results from this study suggested that GMA has questionable utility in this context. Specifically, GMA did not possess a significant, positive relationship with performance across or within positions. Furthermore, GMA produced significant subgroup differences, was unrelated to where a prospect was selected in the NFL Draft, and possessed a null relationship with number of games started. Evident by these findings, one should examine feasible alternative measures that could be utilized in this setting.

Recommendations and Future Research Directions

Although GMA possessed a null relationship with performance, several concerns about the WPT's administration, item content, and test-taking motivation/perceptions toward the instrument may have affected its predictive efficiency. In terms of administration, learning disabilities can affect WPT scores if not provided extra time to complete the test. It is not certain that prospects with learning disabilities will explicitly state they need extra time before its administration, most likely due to anxiety or fear of being downgraded in the NFL Draft. Consequently, the presence of learning disabilities may lower the probability of attaining prospects' true scores on the WPT, especially considering that this measure is a speeded test of GMA. Similarly, WPT coaching by specific organizations (e.g., Parisi Speed School, 2007) and agent representatives may also adversely affect the probability of attaining a prospect's true score (cf., Sackett, Burris, & Ryan, 1989) and in turn, attenuate its relationship with performance.

A second concern entails whether the WPT's item content genuinely reflects the GMA construct that would be needed to succeed in this context. For example, the WPT is purported to

measure verbal, numerical, general, analytical, and spatial relations ability. Instead of a primary emphasis on crystallized intelligence (Horn, 1976; e.g., verbal, numerical, general knowledge), perhaps a GMA assessment that emphasizes fluid intelligence (e.g., reasoning, picture arrangements/completion, and short/long-term memory) would be more pertinent to the cognitive requirements of the NFL. For instance, recognizing offensive or defensive schemes during a game may reflect a fluid intelligence construct rather than a crystallized construct of GMA. Aside from aptitude tests, another alternative would be to indirectly measure GMA through job knowledge tests or situational judgment tests (SJTs) by each position. For example, an SJT could examine responses to various situations that players would encounter during a game. These tests usually elicit smaller subgroup differences than pure GMA tests (e.g., Weekley & Jones, 1999).

Finally, prospects' perceptions toward the WPT may also affect their score. In general, research has indicated that face validity perceptions toward components of a selection system influence test performance (e.g., Chan, 1997; Chan, Schmitt, DeShon, Clause, & Delbridge, 1997; Ryan & Ployhart, 2000). Related to this study, Chan (1997) found that African-Americans engendered more negative perceptions about the criterion-related validity of GMA tests than Caucasians. Thus, prospects' perceptions towards the test and their subsequent motivation to take the test seriously may be questionable. In turn, negative perceptions toward the WPT likely influence the probability of attaining true scores.

Rather than administering the WPT, other types of selection instruments could be utilized for assessment during the NFL Combine. As stated before, SJTs and job knowledge tests could be applied to this context. Also, achievement motivation may be an important construct to assess as it may relate to a player's competitiveness and willingness to succeed, regardless of failure.

Another strategy would be to assess the fit between the prospect's ability and personality with certain team climates (e.g., Bowen, Ledford, Nathan, 1991). For example, some prospects could be more suited for the West-Coast offense (i.e., emphasizes short, quick passes and ball control) than the Run-and-Gun (i.e., emphasizes a strong running game and passing the ball downfield).

In addition to performance, the NFL could assess another criterion, the propensity and susceptibility to engage in counterproductive behavior. Recently, the NFL instituted a more stringent policy that induces fines and suspensions to those players that engage in counterproductive behaviors off-the-field. One method to screen for such acts would be to administer an integrity test or a personality inventory during the NFL Combine. For example, integrity tests have been shown to adequately predict counterproductive behavior (Ones, Viswesvaran, & Schmidt, 1993) while maintaining equity (Sackett, Burris, & Callahan, 1989). Accordingly, future research should examine the generalizability and psychometric properties of these tests in the NFL, with the intent of achieving efficiency and minimizing equity concerns.

Limitations

Several limitations of this study should be noted. First, predictor and criteria data were gathered from nationally recognized secondary data sources, *CBS.sportsline.com*, *NFL.com*, and *NFLdraftscout.com*. Although these Web sites are secondary data sources, we believe that these are reputable sources for sporting-related information. Second, despite the convergent evidence supporting the use of objective performance, it still may not be entirely indicative of the ultimate performance domain within this context. As a result, our measure of performance may be contaminated due to situational factors outside of a player's control. In particular, the following situational factors may enhance or inhibit a player's statistical performance – degree of depth chart quality, opponent quality (i.e., strength of schedule), and overall talent level of the entire

team. Furthermore, the draft is ordered in such a way that teams with the worst records from the previous season receive higher draft picks (e.g., first overall, second overall, etc.) than better teams.¹ Therefore, if a talented rookie running back gets drafted by a team of poor quality (e.g., lack of offensive line depth and quality, consecutive losing seasons), his performance will most likely have a higher probability of being attenuated than an average running back that gets drafted by a team of higher quality. Despite these shortcomings, the observed impact of our performance criteria on retention decisions and compensation underscore the appropriateness of statistical performance as an indicator of performance in the NFL. A final limitation is sample size within position. Some positions, such as quarterbacks, had less statistical power than others. Such a lack of statistical power most likely inhibited our ability to detect meaningful statistical differences in the moderated regression analyses.

Conclusion

Empirical research has supported the validity of GMA as a predictor of job performance in traditional employment settings. However, the results from this study suggest that in the context of professional football, GMA: (a) possessed a near zero relationship with performance across positions and had an occasional significant negative relationship with performance by position; (b) did not differently predict performance by race; and (c) was unrelated to selection in the NFL Draft or the number of games started during an NFL season. Therefore, its use in the NFL Combine is, at best, questionable in nature. Accordingly, more research is needed to determine if GMA is related to proximal criteria of interest such as trainability ratings, playbook knowledge acquisition, play recall, and assignment recognition.² It may be that the effect of GMA on performance may attenuate to a null relationship through a casual chain of acquiring

¹ We thank an anonymous reviewer for this suggestion.

² We thank an anonymous reviewer for this suggestion.

playbook knowledge, play recall, assignment recognition, and then game performance.

Nevertheless, as this study demonstrated, GMA was unrelated to future performance, which is an important distal criterion for effectiveness in this context. As a result, future research examining the efficiency and equity of other selection instruments is clearly warranted.

References

- Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology, 90*, 94-107.
- Bartlett, C. J., Bobko, P., Mosier, S. B., & Hannan, R. (1978). Testing for fairness with a moderated multiple regression strategy: An alternative to differential analysis. *Personnel Psychology, 31*, 233-241.
- Bivens, S., & Leonard II, W. M. (1994). Race, centrality, and educational attainment: An NFL perspective. *Journal of Sport Behavior, 17*, 1-8.
- Bobko, P., Roth, P. L., & Buster, M. A. (2007). The usefulness of unit weights in creating composite scores: A literature review, application to content validity, and meta-analysis. *Organizational Research Methods, 10*, 689-709.
- Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implication of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology, 52*, 561-589.
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt, W. C. Borman, & Associates (Eds.), *Personnel selection in organizations* (pp. 71-98). San Francisco, CA: Jossey-Bass.
- Bowen, D. E., Ledford, G. E., & Nathan, B. R. (1991). Hiring for the organization, not the job. *Academy of Management Executive, 5*, 35-51.
- Chan, D. (1997). Racial subgroup differences in predictive validity perceptions on personality and cognitive ability tests. *Journal of Applied Psychology, 82*, 311-320.

- Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Reactions to cognitive ability tests: The relationships between race, test performance, face validity perceptions and test-taking motivation. *Journal of Applied Psychology, 82*, 300-310.
- Gully, S. M., Payne, S. C., Koles, K. L. K., & Whiteman, J. A. K. (2002). The impact of error training and individual differences on training outcomes: An attribute-treatment interaction perspective. *Journal of Applied Psychology, 87*, 143-155.
- Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.
- Hogan, J. (1991). Structure of physical performance in occupational tasks. *Journal of Applied Psychology, 76*, 495-507.
- Horn, J. L. (1976). Human abilities: A review of research and theory in the early 1970s. *Annual Review of Psychology, 27*, 437-485.
- Jaccard, J., & Turrisi, R. 2003. *Interaction effects in multiple regression* (2nd ed.). Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-072. Thousand Oaks, CA: Sage.
- McKay, P. F., & McDaniel, M. A. (2006). A reexamination of Black-White mean differences in work performance: More data, more moderators. *Journal of Applied Psychology, 91*, 538-554.
- Merron, J. (2002, February 28). *Taking your Wonderlics*. Retrieved May 6, 2007, from <http://espn.go.com/page2/s/closer/020228.html>
- Mulligan, M. (2004, April 22). Wonderlic scores have NFL teams wondering. *The Chicago Sun-Times*, pp. S4.

- Murphy, K. R. (1989). Is the relationship between cognitive ability and job performance stable over time? *Human Performance*, 2, 183-200.
- Murphy, K. R. (2002). Can conflicting perspectives on the role of *g* in personnel selection be resolved? *Human Performance*, 15, 173-186.
- Murphy, K. R., & Davidshofer, C. O. (2001). *Psychological testing: Principles and applications* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- NFL Players Association. (2007). *FAQs: NFL hopefuls FAQ*. Retrieved June 5, 2007, from http://www.nflpa.org/Faqs/NFL_HopefulsFaq.aspx
- Offermann, L. R., & Gowing, M. K. (1993). Personnel selection in the future: The impact of changing demographics and the nature of work. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 385-417). San Francisco, CA: Jossey-Bass Publishers.
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78, 679-703.
- Ottz, J. L. (2002). The role of cognitive ability tests in employment selection. *Human Performance*, 15, 161-171.
- Parisi Speed School. (2007). *NFL combine: State-of-the-art training for improving NFL combine performance*. Retrieved June 1, 2007, from <http://www.parischool.com/programs/combine.asp>
- Ployhart, R. E., & Hakel, M. D. (1998). The substantive nature of performance variability: Predicting interindividual differences in intraindividual performance. *Personnel Psychology*, 51, 859-901.

- Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology, 79*, 518-524.
- Roth, P. L., BeVier, C. A., Bobko, P., Switzer, F. S., III, & Tyler, P. (2001). Ethnic group difference in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54*, 297-330.
- Ryan, A. M., & Ployhart, R. E. (2000). Applicants' perceptions of selection procedures and decisions: A critical review and agenda for the future. *Journal of Management, 26*, 565-606.
- Sackett, P. R., Burris, L. R., & Callahan, C. (1989). Integrity testing for personnel selection: An update. *Personnel Psychology, 42*, 491-529.
- Sackett, P. R., Burris, L. R., & Ryan, A. M. (1989). Coaching and practice effects in personnel selection. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology*. Chichester: John Wiley and Sons.
- Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance, 15*, 187-210.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.
- Schmidt, F. L., & Kaplan, L. B. (1971). Composite vs. multiple criteria: A review and resolution of the controversy. *Personnel Psychology, 24*, 419-434.
- Schmidt, F. L., Hunter, J. E., & Pearlman, K. (1981). Task differences as moderators of aptitude test validity in selection: A red herring. *Journal of Applied Psychology, 66*, 166-185.

- Schmitt, N., & Chan, D. (1998). *Personnel selection: A theoretical approach*. Thousand Oaks, CA: Sage.
- Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency of various predictor combinations. *Journal of Applied Psychology*, 82, 719-730.
- Stevens, M. J., & Campion, M. A. (1999). Staffing work teams: Development and validation of a selection test for teamwork settings. *Journal of Management*, 25, 207-228.
- Viswesvaran, C., & Ones, D. S. (2002). Agreements and disagreements on the role of general mental ability (GMA) in industrial, work, and organizational psychology. *Human Performance*, 15, 211-231.
- Weekley, J. A., & Jones, C. (1999). Further studies of situational tests. *Personnel Psychology*, 52, 679-700.
- Wonderlic, Inc. (2002). *Wonderlic personnel test & scholastic level exam user's manual*. Libertyville, IL: Wonderlic, Inc.
- Wonderlic, Inc. (2004). How smart is your first round draft pick? *HR Measurements*. Retrieved May 4, 2007, from <http://www.wonderlic.com/resources/newsletters/2004Summer/smartDraftPick.asp>
- Wonderlic, Inc. (2005, March 1). NFL testing provides valuable lesson for all employers. Retrieved May 22, 2007, from <http://www.wonderlic.com/resources/press/2005-03-01.asp>

Table 1

Race Distribution by Draft Year

Race	2002	2003	2004	Sub-total
African-American	187 (73%)	178 (69.3%)	163 (65.5%)	528 (69.3%)
Caucasian	64 (25%)	70 (27.2%)	81 (32.5%)	215 (28.2%)
Other	5 (2%)	9 (3.5%)	5 (2%)	19 (2.5%)
Total	256	257	249	762

Table 2

Means, Standard Deviations, and Correlations

Variables	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8
1. Wonderlic Personnel Test	21.04	7.15	-							
2. Race (1= Caucasian)	0.23	0.42	.51**	-						
3. Draft Selection	128.16	74.63	.05	.14**	-					
4. Year 1 NFL Performance	50.00	7.04	-.02	-.12**	-.49**	-				
5. Year 2 NFL Performance	50.00	6.87	-.02	-.08 [†]	-.49**	.62**	-			
6. Year 3 NFL Performance	50.00	7.09	-.06	-.08	-.38**	.53**	.61**	-		
7. Overall NFL Performance	49.27	5.93	-.04	-.10*	-.54**	.86**	.88**	.85**	-	
8. Year 1 Games Started	50.00	9.80	-.06	-.04	-.45**	.73**	.48**	.33**	.62**	-
9. Year 2 Games Started	50.00	9.79	-.02	-.08 [†]	-.48**	.58**	.76**	.44**	.71**	.57**
10. Year 3 Games Started	50.00	9.77	-.08	-.07	-.47**	.49**	.61**	.74**	.73**	.43**
11. Overall Games Started	49.05	8.12	-.06	-.08 [†]	-.56**	.73**	.75**	.61**	.81**	.83**
12. Year 3 NFL Status	0.76	0.43	-.02	-.07 [†]	-.40**	.28**	.21**	-	.30**	.23**
13. Year 4 NFL Status	0.67	0.47	.02	-.03	-.42**	.33**	.34**	.26**	.41**	.29**
14. Year 3 Salary	595638.61	1074759.81	-.01	-.04	-.17**	.20**	.29**	.29**	.31**	.18**
15. Year 4 Salary	1404483.54	2093745.39	.04	.04	-.15**	.22**	.26**	.33**	.33**	.16**

Note. Offensive linemen were excluded from this analysis. Salary measured in U.S. dollars. Sample sizes ranged from 354 to 640. [†] $p < .10$. * $p < .05$. ** $p < .01$.

Table 2 (cont.)

Means, Standard Deviations, and Correlations

Variables	<i>M</i>	<i>SD</i>	9	10	11	12	13	14
1. Wonderlic Personnel Test	21.04	7.15						
2. Race (1 = Caucasian)	0.23	0.42						
3. Draft Selection	128.16	74.63						
4. Year 1 NFL Performance	50.00	7.04						
5. Year 2 NFL Performance	50.00	6.87						
6. Year 3 NFL Performance	50.00	7.09						
7. Overall NFL Performance	49.27	5.93						
8. Year 1 Games Started	50.00	9.80						
9. Year 2 Games Started	50.00	9.79	-					
10. Year 3 Games Started	50.00	9.77	.61**	-				
11. Overall Games Started	49.05	8.12	.88**	.83**	-			
12. Year 3 NFL Status	0.76	0.43	.20**	-	.26**	-		
13. Year 4 NFL Status ^a	0.67	0.47	.32**	.26**	.38**	.73**	-	
14. Year 3 Salary ^a	595638.61	1074759.81	.22**	.28**	.28**	-	.09 [†]	-
15. Year 4 Salary ^b	1404483.54	2093745.39	.23**	.27**	.27**	.04	-	.13*

Note. Offensive linemen were excluded from this analysis. Salary measured in U.S. dollars. Sample sizes ranged from 354 to 640. [†] $p < .10$. * $p < .05$. ** $p < .01$.

Table 3

Correlations between the Wonderlic Personnel Test and Overall NFL Performance by Position

Position	Overall NFL Performance
QB (<i>n</i>)	.03 (32)
RB (<i>n</i>)	.12 (51)
WR (<i>n</i>)	.04 (74)
TE (<i>n</i>)	-.28 [†] (41)
DL (<i>n</i>)	.03 (104)
LB (<i>n</i>)	.04 (75)
DB (<i>n</i>)	-.21* (120)

Note. QB = quarterbacks; RB = running backs; WR = wide receivers; TE = tight ends; DL = defensive linemen; LB = linebackers; and DB = defensive backs. [†] $p < .10$. * $p < .05$.

Table 4

Moderated Regression Results Examining the Influence of Position Complexity

Variables	Overall NFL Performance		
	Step 1	Step 2	Step 3
Race	-.12**	-.17**	-.17**
WPT		.01	.10
RB		-.07	-.06
WR		-.10	-.09
TE		-.004	-.01
DL		-.08	-.08
LB		-.07	-.07
DB		-.12	-.12
WPT x RB			.01
WPT x WR			-.02
WPT x TE			-.08
WPT x DL			-.002
WPT x LB			-.01
WPT x DB			-.14
R^2	.02**	.02	.04
Adjusted R^2	.01	.00	.01
F	7.33**	1.24	1.32
ΔR^2	.02**	.00	.02

Note. WPT = Wonderlic Personnel Test. Quarterbacks were the referent subgroup. Standardized coefficients are reported. $N = 490$. ** $p < .01$.

Table 5

Moderated Regression Results Examining Differential Prediction by Race

Variables	Overall NFL Performance		
	Step 1	Step 2	Step 3
WPT	-.05	-.01	.02
Race		-.12*	-.09 [†]
WPT x Race			-.08
R^2	.003	.02*	.02*
Adjusted R^2	.00	.01	.01
F	1.33	3.70*	3.03*
ΔR^2	.003	.01*	.003

Note. WPT = Wonderlic Personnel Test. African-Americans were the referent subgroup. Standardized coefficients are reported. $N = 490$. [†] $p < .10$. * $p < .05$.

Table 6

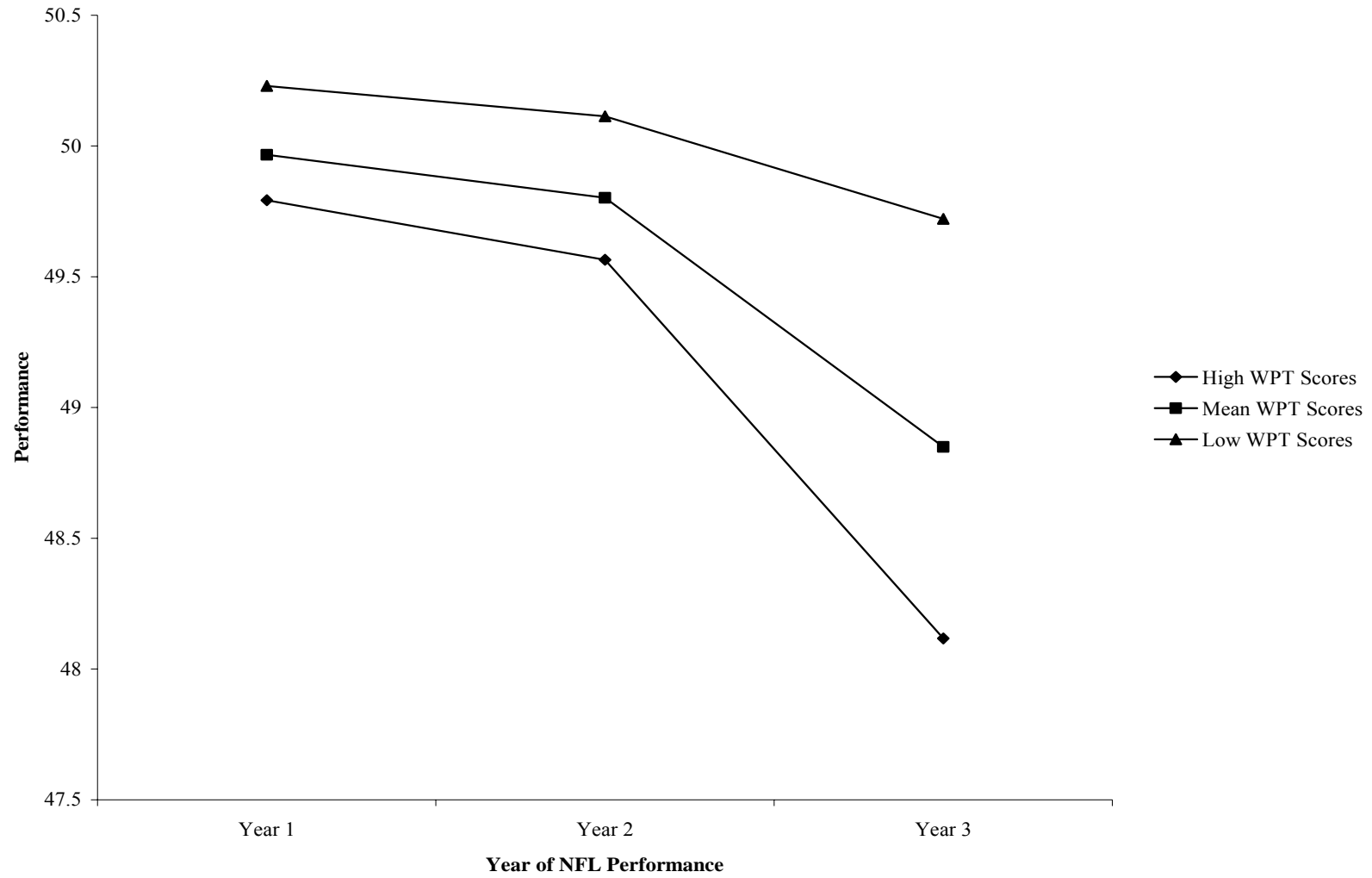
Logistic Regression Analyses Predicting NFL Retention Status

	<i>N</i>	Step	Variables	Year 3 NFL Status				Year 4 NFL Status			
				<i>B</i>	<i>SE(B)</i>	Exp(<i>B</i>)	Fit Statistic	<i>B</i>	<i>SE(B)</i>	Exp(<i>B</i>)	Fit Statistic
1	Y3: 495 Y4: 495	1	Draft Selection Nagelkerke R^2	-.01	.00	.99*		-.004	.00	1.00*	
		2	Year 1 NFL Performance Model $\Delta\chi^2$ Model χ^2 Nagelkerke R^2	.25	.05	1.29**	.11 40.15** 72.14** .24	.21	.04	1.23**	.11 50.92** 87.33** .25
2	Y3: 488 Y4: 488	1	Draft Selection Nagelkerke R^2	-.01	.00	1.00*		-.004	.00	1.00*	
		2	Year 2 NFL Performance Model $\Delta\chi^2$ Model χ^2 Nagelkerke R^2	.16	.05	1.17**	.08 14.94** 33.57** .15	.21	.04	1.24**	.12 48.40** 86.11** .26
3	Y4: 433	1	Draft Selection Nagelkerke R^2					-.004	.00	1.00	.08
		2	Year 3 NFL Performance Model $\Delta\chi^2$ Model χ^2 Nagelkerke R^2			Not applicable		.26	.06	1.30**	33.36** 50.02** .24
4	Y3: 559 Y4: 559	1	Draft Selection Nagelkerke R^2	-.003	.00	1.00		-.002	.00	1.00	
		2	Overall NFL Performance Model $\Delta\chi^2$ Model χ^2 Nagelkerke R^2	.27	.05	1.32**	.10 45.18** 77.44** .23	.34	.04	1.41**	.13 96.49** 146.82** .34

Note. Y3 = Year 3 NFL status sample size. Y4 = Year 4 NFL status sample size. * $p < .05$. ** $p < .01$.

Figure 1

Examining the Relationship between Different Levels of Wonderlic Personnel Test (WPT) Scores and NFL Performance Over Time



Appendix

Performance Criteria by Position

Position	Rating ^a	Yards Passing	Yards Per Attempt	Completion Percentage	Total Completions	Fumbles ^b	Rush Yards	Carries	Receptions	Reception Yards
QB	X	X	X	X	X	X				
RB						X	X	X	X	X
WR						X			X	X
TE						X			X	X
DL										
LB										
DB										

Note. Cells with an “X” represent measured criteria. ^aRating produces a value that estimates a quarterback’s efficiency. ^bFumbles were negatively coded for all positions. ^cInterceptions were negatively coded for quarterbacks.

Appendix (cont.)

Performance Criteria by Position

Position	Touch-downs	Total Tackles	Forced Fumbles	Sacks	Inter-ceptions ^c	Passes Defended	Games Started
QB	X				X		X
RB	X						X
WR	X						X
TE	X						X
DL		X	X	X	X	X	X
LB		X	X	X	X	X	X
DB		X	X	X	X	X	X

Note. Cells with an “X” represent measured criteria. ^aRating produces a value that estimates a quarterback’s efficiency. ^bFumbles were negatively coded for all positions. ^cInterceptions were negatively coded for quarterbacks.