FINDING THE 'I' IN TEAM: DEVELOPMENT OF A SNAPSHOT INDIVIDUAL PERFORMANCE METRIC

By

Matthew-Donald Sangster

A Dissertation Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

Major Subject: COGNITIVE SCIENCE

Examining Committee:

____

Dr. Wayne Gray, Dissertation Adviser

____

Dr. David Mendonca, Member

____

Dr. Stanley Dunn, Member

Rensselaer Polytechnic Institute
Troy, New York

January 2017
(For Graduation May 2017)

Contents

List of Tables

List of Figures

ACKNOWLEDGMENT

Thank you to my adviser and readers, respectively Dr. Wayne Gray, Dr. David Mendona, and Dr. Stanley Dunn. Additional thanks to Riot Games and our League of Legends working group. And finally, thank you to my friends, family, and fianc.

ABSTRACT

Though both individual and team performance are widely studied, evaluating individual performance within the context of the team usually relies on relative measures of performance. However, in order to apply much of the findings and theories from individual performance literature it is crucial to develop an absolute measure of individual performance–a “Snapshot” individual performance metric. With such a measure, it becomes possible to determine the contribution of a individual towards the goals of the team. This paper uses Big Data (1.9 million records from 539 thousand matches) from League of Legends, a widely popular competitive team game to look at individual performance based on the overall priorities different members of a team should have. This research applies exploratory factor analysis and logistic regression to determine the latent factor structure of behavioral variables that are predictive of team performance. Through this, we develop the first steps of establishing an individual priority-based measure of performance that can be used to evaluate individual performance for a single observation.

Keywords: Multiplayer Online Battle Arena Game, League of Legends, Individual Performance, Exploratory Factor Analysis, Logistic Regression, Team Performance, Absolute Performance

Chapter 1
INTRODUCTION

Theories of learning and memory often rely on individual measures of performance for tasks in a given domain. However, when expanding these works to individuals in a team, many difficulties arise. The very nature of team tasks assume that individuals collaborate for a common goal. Thereby, it is complicated to try to differentiate between contributions of one team member and another. The National Basketball Association (NBA) has encountered this issue in that it becomes difficult to evaluate the merit of players that do not explicitly contribute to the task outcome (scoring points) (Concepcion, 2015). The NBA addressed this issue by developing a composite "hustle stat" or "hustle rating" statistic that evaluates certain 'intangible' and behavioral elements of gameplay not captured by typical "points scored", "assists", and "rebounds" measures. Thus, one method to bridge the difficulty of differentiating team member contribution is to look at large amounts of behavioral team data for a single domain, and determine commonalities for different 'roles' on the teams.

Multiplayer Online Battle Arena (MOBA) games such as League of Legends (LoL) provide an opportunity to unobtrusively evaluate large amounts of behavioral data from teams as there are millions of games available for use through their publicly accessible application program interface (API) (Sangster, Mendonça, & Gray, 2016). MOBA games generally consist of three or five member teams fighting for control over a map. Each member of the two opposing teams controls one unit that has varying abilities for combat and typically fulfills one of 5 roles for the team. Due to the availability of data, the structure of teams, and the distribution of roles, League of Legends, and MOBAs, in general offer an untapped wealth of opportunity for individual-focused assessments of team performance.

In this paper, we use Big Data from League of Legends matches to develop an 'absolute' behavior-based performance metric that is effective at assessing individual contribution to task outcome in a team environment. The goal of this research is to begin groundwork on applying theories of learning to individuals within a team setting, and eventually applying theories of individual learning to team learning.

Chapter 2
BACKGROUND

Multiplayer Online Battle Arena (MOBA) games are team-based dynamic task environments where the only overall task objective is to destroy the enemy's base. However, since its inception, the genre of MOBA has been difficult to define (Ferrari, 2013). The genre contains elements of role-playing games (RPGs) because the player takes on the identity and abilities of an established character. MOBAs are also strictly-competitive games because in order to win a MOBA your opponent must lose. MOBAs are team games. Though one-on-one scenarios are common in course of a game, a single player's actions are rarely decisive. It often requires a full 'team effort' for a game to end. Lastly, MOBAs are strategic games, they require strategy from team composition to team coordination all the way until through the final moments of the game. This last point is of most importance to this research. MOBAs are a relatively rare synthesis of Team-based Real-Time-Strategy games. Many First Person Shooters (FPS) include both team and strategy elements, but they are not traditional team or strategy games as the player can only ever control a single unit and the units generally have the same or extremely similar capabilities. MOBAs often afford players the option of controlling additional units through spells, items, or abilities. However, one of the primary arguments for MOBAs as strategy games is the idea of minion equilibrium. The minions are the fully autonomous creatures that spawn every 30 seconds in each base, fight in each 'lane', and 'push' towards their respective enemy's base (see Figure 2.1). With no interference, the minions of each side will meet in their respective lanes along the black dotted diagonal on the figure. The minions will then attack each other until they die and eventually be replaced by the next 'wave' (name for the group of minions that spawn every 30 seconds) of minions, thus maintaining the minion equilibrium along the black dotted line. Maintaining and affecting this minion equilibrium is done by killing enemy minions (and in some MOBAs, 'denying' allied minions – e.g. landing the killing blow on an allied minion and 'denying' the enemy the experience and gold for killing the minion). A team that kills an enemy minion will have one more allied minion attacking the enemy's minions and therefore the allied minions will clear the 'wave' of enemy minions before the reinforcements arrive to replace the minion wave (as is what typically occurs). This results in the minion equilibrium 'pushing' towards the enemy's base and is referred to as 'pushing the lane'. So, though MOBA players don't necessarily have control additional units, they can impose on additional units strategically, and thereby cause changes to the game state through means outside of their individual character.

In addition to the particularities of gameplay across teams and players, there are many intricacies as to the individual's contribution to team performance in a game of League of Legends (Donaldson, 2015). Further, there are many ways that team performance can be affected through non-intuitive behavioral decisions both in and out of game. Thus, due to the complex nature of the task, MOBAs are not easily conducive to the development of an 'absolute' measure of performance.

Through much of the individual performance literature, theories are developed based on tasks that have obvious performance metrics (e.g. score). However, tasks that do not lend themselves easily to a quantifiable score must rely on ranking systems such as Elo to evaluate performance over time (Huang, Yan, Cheung, Nagappan, & Zimmermann, 2017; Huang, Zimmermann, Nagappan, Harrison, & Phillips, 2013; Maas & Wagenmakers, 2005; Gong et al., 2015). The Elo rating system was developed by Arpad Elo for chess as means to both predict match winners and losers, but also as a mechanism by which to rank chess players (Elo, 1978). While Elo is a very effective system for evaluating a player's skill relative to other players, it does little to explain how well a player played in a given match-up. Throughout video game literature and the video game industry, Elo is essentially the gold standard (Latham, Patston, & Tippett, 2013; Imbeault, Bouchard, & Bouzouane, 2011; Weber, Mateas, & Jhala, 2011). Often game developers will work to create their own iteration of ELO or extend ELO to team games, but the concept is relatively the same. A player/team wins a game, their ranking improves, and if they lose, their ranking decreases. Very rarely are in-game performance details used in conjunction with the player's game outcome to effect their Elo. Thereby, a player can have their best performance ever and the loss would be treated the same as if they had played horribly throughout the game. While there are many pitfalls to ELO, score too suffers from the lack of fidelity as to the process of completing the task.

For tasks where performance is typically evaluated through task outcome (typically by way of Win/Loss as with Chess) it is obvious that task outcome is a crucial element to individual performance. However, in team based games such as Basketball task outcome is less inherently indicative of individual performance. When the NBA developed their hustle stats they were interested in quantifying the game's intangibles (Concepcion, 2015). According to Kiki Vandeweghe, NBA's vice-president of Basketball operations, in that same interview, "These are the things that actually win games for you". Looking at the game's behavioral variables reveals insights into team and individual performance.

2.1 INDIVIDUAL PERFORMANCE RESEARCH

While controlled experiments reign supreme in Cognitive Science, there is much that can be learned from large naturally occurring data sets (NODS), such as the massive set of publicly accessible data from League of Legends (Goldstone & Lupyan, 2016). Due to their size and domain specificity, NODS are well-suited for validation of existing theory and models. NODS allow for researchers to generalize previously domain specific findings to new domains with minimal resource cost and with the robustness desired for replication studies.

Individual performance literature often focuses on an individual performing a single task in isolation. This is highly reflected in the typical tasks used for individual performance literature (Macnamara, Hambrick, & Oswald, 2014) – Chess (e.g. Gobet & Simon, 2000; Simon & Gilmartin, 1973; Ericsson & Lehmann, 1996; Howard, 2014), Music (e.g. Braasch, 2011; Patston, Hogg, & Tippett, 2007; Brochard, Dufour, & Despres, 2004; Schellenberg, 2001; Sloboda, 1991), and, to some extent, single player video games (e.g. Destefano, Lindstedt, & Gray, 2011; Lindstedt & Gray, 2013; Kirsh & Maglio, 1994; Sims & Mayer, 2002). These tasks all look at individual performance in isolation. However, none of these tasks occur in isolation in the real world; in chess you have opponents, in music you have accompaniments and play in orchestras, and with video games you often have friends that you discuss the game with. When we study individuals, we really are, to an extent, studying teams. Thus, a measure of individual performance will always contain the contribution of the team to that individual's performance. Looking at individual performance *over time* further complicates the issue. Most models that look at performance over time are very simple and often overlook well established principles of cognitive science and psychology. For example, the spacing effect which was established over a century ago (Ebbinghaus, 1913) is absent from all but a few models. There has, however, been some success with a model applying this effect of mass versus distributed practice to Air Force squadrons (Jastrzembski, Gluck, & Rodgers, 2009). Indeed, shows, through their research on the Predictive Performance Optimizer (PPO), how some of the variance in team performance of complex tasks can be brought under control by applying cognitive science theory and modeling techniques. Research on the PPO provides initial support for modeling learning over time through inclusion of a decay parameter that acts to symbolically represent the spacing effect. The complication with extending the research on the PPO is that it requires, as an input, an absolute performance metric.

2.2 BIG DATA AND MOBAS

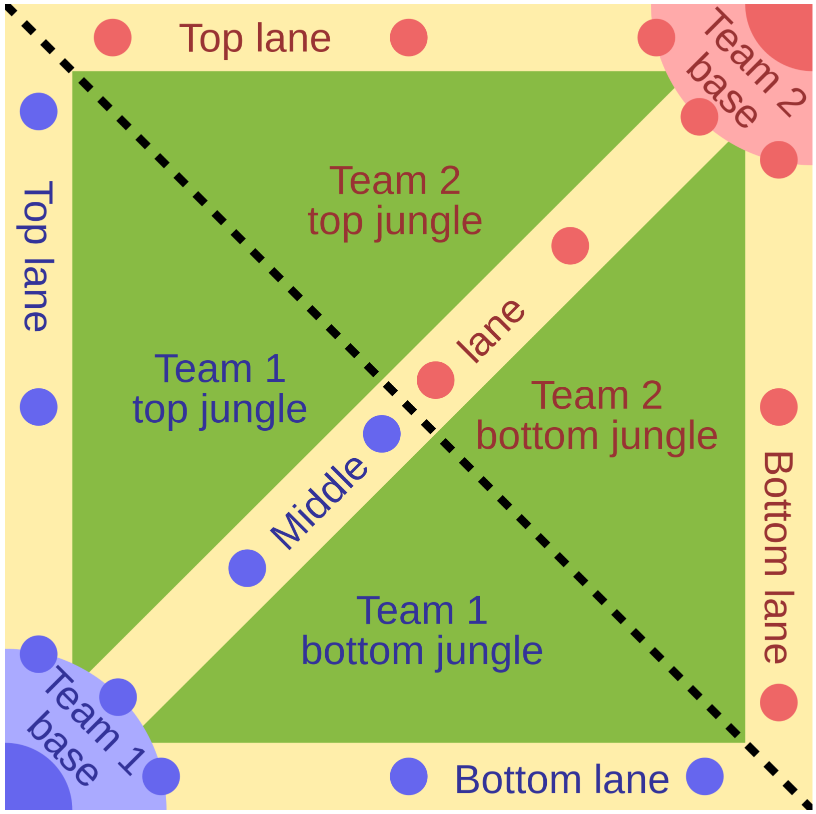
The loss of experimental control is a large fear for big data, but the sacrifice is not without its benefits (Stafford & Dewar, 2014). With large sample sizes, researchers can conduct quasi-experimental research designs that provide much greater statistical power than traditional experiments (Shadish, Cook, & Campbell, 2002). Cohort analyses are a common extension of this quasi-experimental paradigm and have yielded interesting results.

One crucial finding from big data in games was that expert skill can be determined from even the most, seemingly, meaningless actions (Huang et al., 2017). Huang looks at the actions players took in the beginning portion of a video game, when "units are moving automatically to collect resources without any input needed from the player". Through this, he finds that players are engaging in fast-paced repetitive keystrokes during this period of time when there is no reason to do so – the keystrokes do not even affect the units. However, these seemingly meaningless actions were indicative of player expertise as they seemed to have a more intangible purpose that was not readily identifiable. It turns out that expert players were engaging in "warmup" behavior by practicing keystrokes that would become important during the course of the game. Our research intends to take the lessons learned by Huang and use Big Data to conduct an exploratory search of behavioral variables to assess individual performance within a team game.

While our interest is with team games, performance literature that uses games, typically, will use single player games. First Person Shooter (FPS) games are common in performance literature as they have clearly defined surface-level objectives that are relatively easy to evaluate (Huang et al., 2013). However, even these games suffer from the focus on relative ranking information in lieu of absolute performance metrics. Relative ranking is important for determining comparative skill, however it is difficult to use ranking to compare within an individual over time to determine learning. The reason for this is that most video game matchmaking systems are based on the players' rankings and therefore bootstrap with rankings. This makes determining growth difficult because rankings only improve when a player wins, however a player can still improve and learn with a loss.

While MOBA games do not have an easily identifiable measure for individual performance, these games have been studied with increasing frequency. League of Legends, and MOBAs in general, lends itself well to the study of *team* performance: (a) it is a team-based game with high demands for coordinated action across team members; (b) it is highly instrumented, with detailed records kept on many aspects of performance (in fact, detail is sufficient to enable complete post-match playback of the matches); (c) its view of performance is multi-faceted, with many explicit measures both at process and outcome levels; and (d) it enables various measures of team composition to be extracted or derived from match records, such as the working history of team members.

Figure 2.1 – Map of battlefield for Multiplayer Online Battle Arena (MOBA) games (Multiplayer online battle arena, n.d.).



2.3 LEAGUE OF LEGENDS

League of Legends is a Multiplayer Online Battle Arena game and as the genre name implies, the map (i.e. the Battle Arena) the players play on is integral to understanding the game. MOBA games are played on a, roughly, square map with two sides divided diagonally by some change in terrain, for League of Legends that is a river (see Figure 2.1).

The goal of League of Legends is to destroy the central structure in the team base (denoted by the quarter circle at the corners of the map). Each lane (yellow lines on map) spawns one group of lane minions every 30 seconds throughout the game. These minions travel down the lane until they meet a unit viewed as an enemy (team 1 lane minions will attack any unit allied with team 2, and vice versa). Additionally, each lane has three towers (a.k.a. turrets) for each team, denoted by circles (blue for team 1 and red for, their opponent, team 2). In order to destroy their opponent's base, the team must first destroy each successive tower in at least one of the lanes and the towers within the base that guard the central structure (the Nexus). The team that destroys their opponent's nexus first wins. The three towers on the edge of each team's base guard an additional structure called the Inhibitor. If team 1 destroys team 2's middle lane Inhibitor, team 1's lane minions in the middle lane become stronger until the Inhibitor respawns (after a fixed duration). Lastly, within each team's jungle exist 7 camps where neutral (jungle) minions spawn.

2.4 SUMMARY

This study is intended to investigate individual performance in the context of a team environment in order to develop useful measures for studying individuals in the context of a team. Specifically, it is meant to determine whether we can, with the aid of big data from League of Legends, differentiate between the contributions of one team member over another and develop a means by which to assess individual performance relative to the role a person plays on the team. To do this we will search for individual behavior that can serve as a correlate or predictor of individual expertise. Through this effort we will begin the process of making big data and team performance accessible through the lens of individual performance.

Chapter 3
METHODS

In the following section we will discuss the plan for data collection and analysis. This research uses big data from a Multi-player Online Battle Arena game called League of Legends.

3.1 BIG DATA STITCHING

The data we will use comes from the League of Legends public API. We have collected data from over 500,000 matches, and will use a large subset of this for the present research. Each match has a single record for each player in the match, thus some matches have 6 (3 players vs. 3 players) records while some have 10 (5 players vs. 5 players) records. Within each record are summary statistics from the match (e.g. gold earned, experience earned, damage dealt) in addition to some identifying information (e.g. player ID, champion ID, lane, and role played). As each player in a match controls a champion, the full data set can be sorted by the champion played. For the present study, we randomly selected 50 champions (out of the current total of 133 champions) and looked at all the records in our full data set where those 50 champions were played. As each team selects champions from the full pool of 133 champions (without replacement) some matches may have all 10 or 6 players represented in our subset, and some matches may not be represented at all.

3.2 EXPLORATORY FACTOR ANALYSIS

After data collection, this research used Exploratory Factor Analysis (EFA) to determine the latent factors underlying performance. Our analysis relies on the mass collection of a wide array of behavioral variables (see Table A.2) which generally lends itself well to exploratory analysis. Typically, factor analysis is used through Confirmatory Factor Analysis, to confirm a subset of variables to include in sub-scale (Henson & Roberts, 2006). While EFA is by no means uncommon, it is not used as frequently as Principal Component Analysis (PCA). EFA was chosen over PCA because we do not merely want to reduce the dimensionality of our variable set, we believe there may be latent factors causing the observed variable of task performance. Our method of Exploratory Factor Analysis used Ordinary Least Squares (OLS) to find the minimum residual ("minres") solution through the 'fa' function in R (Revelle, 2016). This typically produces similar results to maximum likelihood but has the additional benefit of performing well with bad matricies. Variables were log-scaled as many of the scales for the variables are vastly different. Specifically, we applied a $log(x + 1)$ transformation to the data as many variables include 0 in their range and $log(0)$ is undefined.

3.2.1 SUBSETTING THE DATA

We understand, a priori, that there are differing goals for individuals in a given match of League of Legends. The game is primarily focused on the 'lanes' of the battlefield (see Figure 2.1. Each lane has a series of towers that must be destroyed in succession in order to to destroy the enemy's base (the objective of the game). There are different objectives for the players in each of the three lanes (in addition to the Jungle). Hence, each player's goals are largely determined by the lane s/he chooses to play. In a typical 5 vs. 5 game, one player occupies each of the Top lane, Middle lane, and the team's Jungle while two players occupy the Bottom lane. Though the jungle is not technically a lane, as it contains no towers, it is labeled as one within the dataset. Each lane (or area in the case of the Jungle) has a different set of objective priorities and thus attempts to accomplish different things in support of the main objective of the game (destroy the enemy base). Therefore, we will be separating the dataset by the lane that the player played in for the given match. However, one complication with this is that for the "Bottom" lane there are generally two players that occupy the lane for each team. Within a team, each player is expected to fulfill a different role so it is important to further separate those who played in the "Bottom" lane into groups based on role. Thus, in order to account for this without including bias, we allow for any lane (not just the "Bottom" lane) to be separated by role. To do so, we looked at the role distribution for each lane (Top, Middle, Bottom and Jungle). Any role comprising less than 15% of the lane-subset population will be set aside and grouped as "Other". However, if only one role is played in greater than 15% of the lane-subset's population, the lane-subset will be evaluated as a whole. This resulted in six groupings. Only the bottom lane-subset had two different roles used by more than 15% of its population (Support and Carry roles)– the remaining 5.39% of the bottom lane-subset that did not fall into either Support or Carry roles were lumped into the "Other" role (see Table 4.1 for more details).

3.2.2 FACTOR SCORES

In order to calculate factor scores we will be using one of the simplest methods because of the likelihood of overfitting from such a large data set. Thus, we will use a factor cuttoff level of 0.3 whereby variables with loading < 0.3 or loading > -0.3 on a factor are weighted 0 for the factor. The remaining variables will be considered having 'loaded' on the factor and will receive a standard weight of 1 regardless of loading magnitude. Negative loadings will be given a weight of -1 to reflect the direction of loading.

3.3 LOGISTIC REGRESSION

Logistic regression is used when the response variable for a regression is categorical. In our case, we rely on a binary response variable of task outcome – win/loss. Though individual performance is not expected to perfectly predict task outcome, we do expect that it will be possible to predict task outcome with better accuracy than chance. Rather than rely on c-statistic or other means of validating the logistic regression, we will compare prediction results with actual results for a test set (20% of our data set). The results of the logistic regressions will aide us to know whether the factors determined by the EFA are actual predictors of performance and success. One complication with logistic regression is that there does not exist a perfect parallel to R^2. Thus, most logistic regression models make use of pseudo-R^2 from (1991) or (1989) and many others. Recently, however, steps have been made to re-evaluate pseudo-R^2 and a new candidate has emerged — 's Coefficient of Discrimination (CoD) (2009). Tjur's CoD is a much more intuitive measure that closely relates to R^2 as it is just the difference between the means of the predicted probabilities of each event. This value ranges from 0 to 1 where D = 0 means the model has no discriminatory power (all estimated probabilities are equal) and D = 1 means the model discriminates perfectly (all observed and estimated probabilities are equal for all observations)(Tjur, 2009). Our logistic regressions will make use of Tjur's CoD to enable further assessment of model fit.

3.4 SUMMARY

Our analysis is primarily hinged on the formation of factors to determine the latent structure of variables for individual performance. With the goal of developing a measure of individual performance within the context of a team, it is believed that the latent structure of the behavioral variables for the individual members of the team will shed some light on the differences between roles on the team. Further, this latent structure can be used to predict the performance of the player and the team within the context of a given match. Thus, through use of exploratory factor analysis, a model can be made to predict, with some degree of accuracy, performance of the team from the behavior of an individual.

Chapter 4
RESULTS

Throughout the following results it is important to note that we are less interested in the specifics of how any individual variables are related to the factors resulting from our Exploratory Factor Analysis (EFA) as we are more interested in whether a set of factors *can* be found to be indicative of performance. Thus, in order to assess the efficacy of the factors that result from our EFA, we felt it necessary to use resulting factor scores to attempt to predict game outcome. Though we have mentioned that individual performance should not be expected to be a perfect indicator of 'team-level performance' (assessed through game outcome), it is expected that individual performance will predict team performance with better accuracy than chance (50% in this case).

The analysis follows a particular order as presented in Figure 4.1. We begin by splitting the data into subsets based on the lane that the player was in for the game. After, this, we perform an EFA on a Training subset (80% of the lane subset) and use the resulting loadings to calculate factor scores for the full lane subset. The factor scores for the Training subset were then used to perform a logistic regression to predict game outcome. After, we applied the resulting model coefficients to the Test set (remaining 20% of the lane subset) to test the prediction accuracy.

4.1 CORRELATIONS

Before completing the Exploratory Factor Analysis for each subset, we looked at an overall correlation plot to look for trends and relationships that may exist regardless of role or lane. Figure 4.2 shows the pairwise correlations for each of the variables that are included in the EFAs.

This plot shows some obvious positive correlations between dealing damage, in general, and dealing damage to specific targets. However, it also shows some non-intuitive correlations such as a positive correlation between from deaths to both gold spent and earned.

In addition to the correlation plot for the dataset as a whole, we looked at the

Figure 4.1 – Flowchart of Analysis

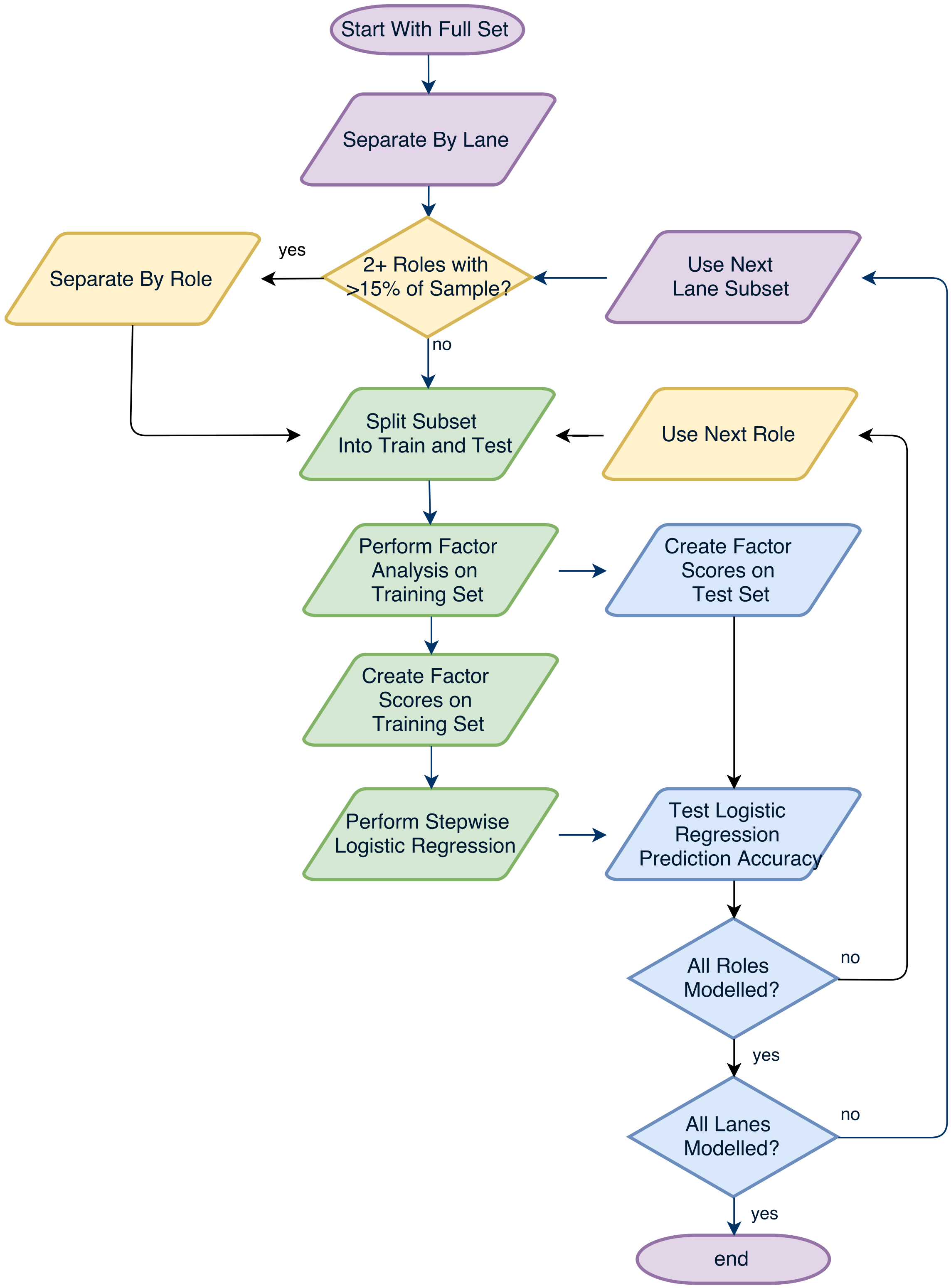
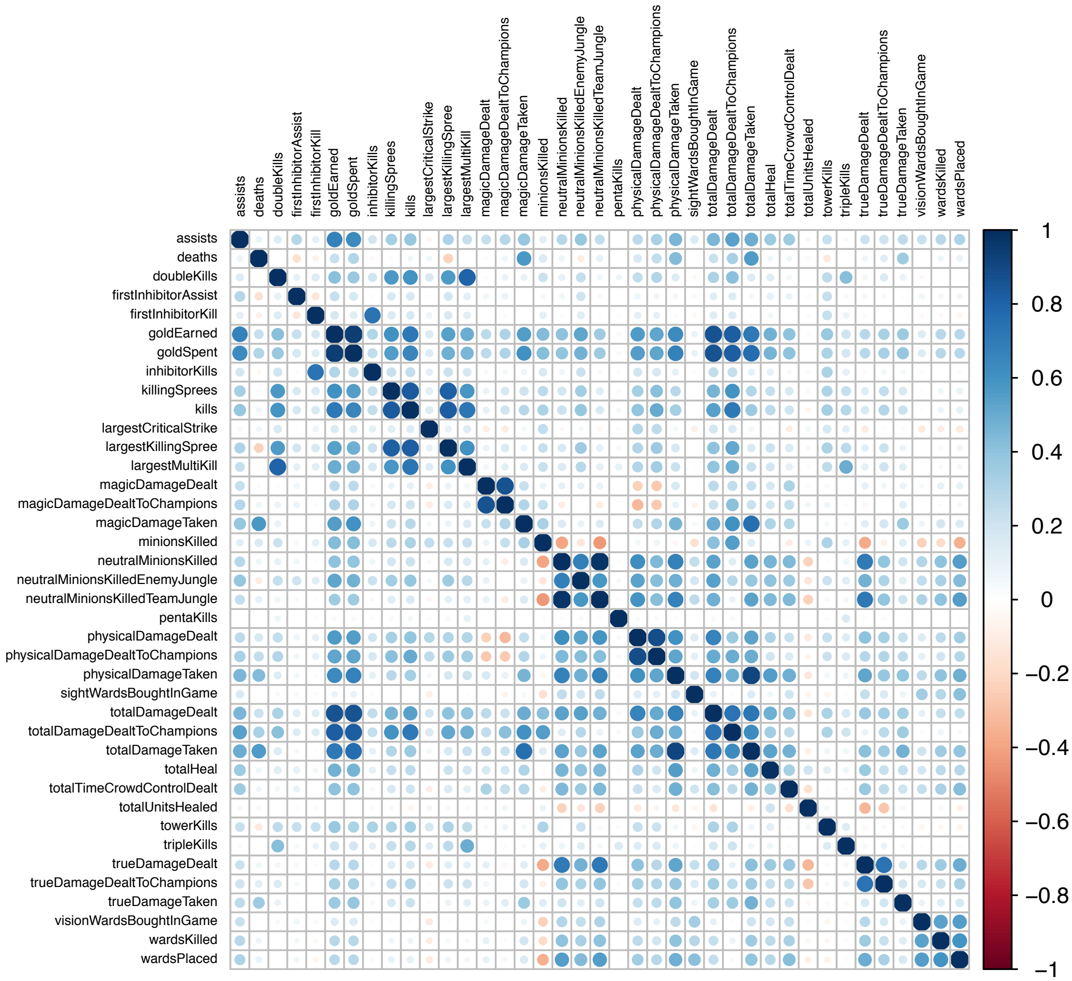


Figure 4.2 – Correlation Plot for Match-level Behavioral Variables for full dataset.



individual correlation plots for each of the Model Subsets (see Appendix, Figures B.1-B.6). These plots show different trends for each of the model subsets and demonstrate the need for subsetting the full data set in this manner. Specifically, there are much more prevalent and stronger correlations between the a wide array of variables for the "Carry" players in the "Bottom" lane (Figure B.3) than for any other player. The pattern of correlations for the "Carry" player are in sharp contrast to that of the other player in the "Bottom" lane – the Support (Figure B.2). While we are not particularly interested in the specifics of which variables are correlated there are some obvious trends throughout the plots; different types of damage (magic, physical, true, and total) correlating with each other as well as similar types of damage correlating with different targets (damage dealt to champions specifically and total damage dealt). However, what is striking about these plots is not what variables are correlated with each other, but which variables are not strongly correlated or are even negatively correlated. Each of the negative correlations thus indicate some element of tradeoff between the variables for the players in the Lane/Role.

Model 1: Middle Lane – Figure B.1. The middle lane player is an interesting case because on first glance, the correlation plot looks nearly identical to the Top Lane player (Figure B.5). However, one of the big differences is the negative correlation between magic damage dealt (both overall and to champions) and 'largest critical strike'. This actually demonstrates a deeper dichotomy in the middle lane that champions who play there are generally weaker (as far as ability to take a lot of damage) but can either deal a high amount of damage quickly (through critical strikes) or provide great assistance and 'battlefield control' (through magical damage).

Model 2: Bottom Lane [Support] – Figure B.2. There is a negative correlation between 'healing' variables and 'magic damage' variables that is not evident in any of the other plots. Many support champions have abilities that have different effects depending on who the target is; typically, when these abilities target an ally they provide healing whereas when they target an enemy they cause magical damage. Thus, it is likely there exists some tradeoff for supports between aiding an ally and assisting in killing an enemy.

Model 3: Bottom Lane [Carry] – Figure B.3. The Carry player in the bottom lane has an interesting role, as they are – in the late game – the primary damage dealer. Thus, our attention is brought to the prevalence of strong correlations between all damage types. This is best viewed in contrast to the Top Lane (Figure B.5) and Middle Lane (Figure B.1). For both of those players, there is a negative correlation between both of the magic damage dealt (total and to champions) variables and both of the physical damage dealt (total and to champions) variables. Comparing the Middle and Top lanes to the Bottom Carry suggests that the Bottom Lane Carry tends to have some manner to deal all three types of damage ("True", "Magic", and "Physical") whereas other lanes may be more specialized or focused on a single damage type.

Model 4: Bottom Lane [Other] – Figure B.4. This subset is particularly interesting as it is more of an anomaly than an established role. This subset consists of players that were neither the support nor the carry for the bottom lane. Thus, this player may have been more of a roaming support or some hybrid role. This is evidenced by the negative correlation between wards placed (a support behavior) and neutral minions killed (a jungle player behavior).

Model 5: Top Lane – Figure B.5. The most prevalent negative correlation for the Top lane players exist between Physical damage and Magical damage which suggests that there is more damage specialization for the Top lane and that the player must focus on one or the other.

Model 6: Jungle – Figure B.6. While the Jungle players also have some degree of the negative correlation that was most evident in the Top Lane, they have an additional negative correlation between Minion Kills and neutral Minion Kills. This represents the tradeoff between time spent in a lane (where you would typically get more total minion kills) and time spent in the jungle (fulfilling the player's role).

Overall, the correlation plots for each of the lane-role subsets indicates different underlying structures based on the player's lane and role. These results were used as a form of subset validation to ensure that the subsets were actually capturing different behavior.

4.2 EXPLORATORY FACTOR ANALYSES

The appropriate number of factors per EFA was determined by taking the average of factor estimates from the eigenvalues (Kaiser, 1960), parallel analysis (Horn, 1965), and optimal coordinates (Raîche, Walls, Magis, Riopel, & Blais, 2013) using the 'nScree' function from the 'nFactors' package in R (Raîche, 2010).

This average was then rounded to the nearest whole number. Table 4.1 shows the final number of factors for each modeled subset.

After determining the appropriate number of factors, we preformed the factor analyses for each subset. The Bartlett's test of sphericity for each model was significant ($p < 0.01$) indicating the suitability of the factor models for the data.

Table 4.1 – Summary Information for the Exploratory Factor Analyses and subset criteria. For each model we provide the number of factors used for the Exploratory Factor Analysis, the number of samples used to construct the model, and the percentage of the data for the lane subset, that the role subset accounts for. The values for models constructed with lanes that did not have more than one role accounting for > 15% of the data (Models 1, 5, and 6) are the proportion of the data that is accounted for by the most popular role.

Model	Lane	Role	% of Subset	Sample Size	# of Factors
Model 1	MIDDLE	—	77.10%	566,755	10
Model 2	BOTTOM	SUPPORT	44.51%	281,486	9
Model 3	BOTTOM	CARRY	50.10%	316,788	10
Model 4	BOTTOM	OTHER	5.39%	34,089	8
Model 5	TOP	—	98.28%	252,567	11
Model 6	JUNGLE	—	100%	467,502	9

Kaiser-Meyer-Olkin (KMO) Sampling Adequacy coefficients, which displays the appropriateness of the data for performing factor analysis, were each greater than 0.70 overall. Each model was also examined for the percentage of non-redundant residuals (no model had KMO $> 6\%$). The factor loadings for each model are presented in Figures C.1 through C.6 in the Appendix. This paper is less interested in the specifics of how individual variables loaded on given factors and is more interested in whether a set of factors *could* be found to be indicative of performance. Thus, for each model we will look at only the non-intercept factor that resulted in the highest Odds Ratios (by extension, the highest Estimates as well) for the resulting logistic regressions (Figure 4.2). It is important to note that while the factors have the same names, they are not the same across models (e.g. 'MR1' from Model 1 is not the same as 'MR1' from Model 2).

Model 1: Middle Lane – MR4. If there were a comprehensive description of the role of the Middle lane player, it would do no better than the factor 'MR4'. Nearly every element that is of paramount importance to a Middle lane player is represented in this factor, from multi-kills and gold earning, to dealing damage to champions and destroying towers. The one variable of note is the neutral (Jungle) minions killed in the enemy jungle. One of the primary tasks for Middle lane players is to roam around the map and provide assistance to the other two lanes when they are attacking towers or are making an attempt to kill enemy champions. Killing neutral minions on the path to these objectives demonstrates a level of efficiency that is beneficial for Middle lane players.

Model 2: Bottom Lane [Support] – MR7. The factor 'MR7' consists of just two variables, the total number of neutral (jungle) minions killed and, interestingly, the number of neutral (jungle) minions killed in the enemy's jungle. Though the relation to team performance may not seem inherently intuitive, this factor actually relates to the idea of positioning. A support player wants to, ideally, stay out of the way of the rest of their team and only be there when necessary. Thus, by positioning themselves in the jungle, and specifically the enemy jungle, they are in an advantageous position for surprise attacks on two lanes (the bottom and middle lanes - refer to Figure 2.1).

Model 3: Bottom Lane [Carry] – MR9. While accomplishing objectives is obviously a team level goal, it is interesting to note that for the Bottom lane Carry player it is more of a personal responsibility. As there is a gold bonus to destroying the first inhibitor, and destroying inhibitors and towers in general, it follows that an experienced Carry player would focus on achieving game objectives as quickly as possible. Further, we can see the connection to neutral (Jungle) minions killed in the enemy jungle because by achieving game objectives, the Bottom Lane Carry player affords him/herself the opportunity to kill jungle minions in the enemy jungle safely. Lastly, by killing enemy jungle minions, the Carry player hinders the ability for the enemy team to stage a comeback as they cannot safely acquire gold and experience.

Model 4: Bottom Lane [Other] – MR5. The 'Other' roles for the Bottom lane, as mentioned before, are likely roaming players or hybrid roles. Thus, as a player who is attempting to have good map coverage and involvement in team skirmishes, this player would ideally have a high number of assists, kills and gold. Additionally, it is worth noting that the 'MR5' factor suggests that this type of player, when successful, is seen as likely using physical damage.

Model 5: Top Lane – MR3. This factor, 'MR3' for the Top Lane, is best viewed by the looking at the presence of 'firstInhibitorAssist' and the absence of 'firstInhibitorKill'. It seems that while assisting a teammate in destroying the first Inhibitor is important, it is not advantageous to land the killing blow on the inhibitor. This leads to the idea that this factor captures the more 'utility' aspects

of the top lane player. They are not meant to get the most kills, or deal the most damage, but are more so expected to contribute to the team though other means. This is further evidenced by looking at 'MR3' in comparison to the Bottom lane 'Other' player's 'MR5' factor previously discussed. These factors are very similar as they both reward assist and team play involvement (kills, killing sprees, and dealing damage). However, what is striking are the variables that are in Bottom [Other] 'MR5' and not in Top lane 'MR3'; namely, the wide array of 'damage dealt' variables. The factor for the roaming player rewards dealing damage of nearly any type while only total damage dealt to champions is rewarded for the Top lane player. Thus, this lane, when played successfully, typically involves the player being involved in team fights and, primarily, attacking enemy champions and objectives.
Model 6: Jungle – MR8. Technically, the factor MR8 has the lowest Odds Ratio, but that is because it's coefficient estimate is negative. This factor is truly the most influential in this model's prediction of game outcome and is, strangely, the simplest feature. This feature contains only one variable 'physicalDamageTaken'. It is worth noting, at this time, that all neutral (jungle) minions deal physical damage. Thus, because this factor is penalized it is truly a measure of the Jungle player's effectiveness. Effective jungle players recognize the need to mitigate the damage of jungle minions as early as possible in order to perform their role efficiently. The earlier a player begins to purchase items and level abilities that reduce physical damage, the longer the player can stay in the jungle without risking death and the less physical damage the player will take over the course of the game (resulting in low values of physical damage taken).

4.3 LOGISTIC REGRESSION

For each subset, a stepwise logistic regression was used to model the influence of the factors (from the EFA) on match outcome. Factors were included in or excluded from the model if the AIC of the resulting model was improved. Through this process we did not look for interaction effects. The resulting factor coefficients from these models, as well as Odds Ratios, are located in Table 4.2. The results of the Logistic Regression for each modeled subset are in Table 4.3. These results show better-than-chance prediction accuracy for the logistic regression models trained on the factors developed through Exploratory Factor Analysis. Each model was developed using step-wise regression including factors when the AIC was improved. Though 'backwards' steps (removing a factor) were allowed, all factors that were possible to be included for a given model, were included in the subset's final model. Included in Table 4.3 are the accuracy of the test-set prediction. The correlation coefficients from the logistic regressions preformed on the training set (80% of the data randomly sampled by subset) were applied to the subset's test set (remaining 20% of the data). These values were adjusted such that a value of .5 or greater is scored as a 1 and a value below .5 is scored 0. The adjusted values are compared to the actual values (Win = 1 and Loss = 0).

Model 1: Middle Lane. This model successfully predicted team performance for 83.23% of our test set (and 83.3 % internal accuracy) with a sensitivity of 83.7% and a specificity of 82.9%. The middle lane model was the second most successful and had the second highest CoD. The factor 'MR2' statistically contributes significantly to the model, however with logistic regression significance is less important than Odds Ratios. The Odds Ratio for this variable is very nearly one meaning that with a 1 unit increase in 'MR2' the odds of winning increases 1.003 times (when controlling for the other factors).

Model 2: Bottom Lane [Support]. This model successfully predicted team performance for 74.89% of our test set (and 75.1% internal accuracy) with a sensitivity of 74.9% and a specificity of 75.3%. The Bottom lane Support player model was the least successful and also contained one factor that did not significantly contribute to prediction accuracy. For factor 'MR6' the Odds Ratio is nearly 1 meaning the factor does not meaningfully contribute to prediction accuracy despite improving the AIC of the model (due to the variable selection method, the factor would not have been included if it did not improve the AIC). 'MR6' rewards the Bottom lane Support player for dealing 'True' damage (see Table C.2).

Model 3: Bottom Lane [Carry]. This model successfully predicted team performance for 87.44% of our test set (and 87.7% internal accuracy) with a sensitivity of 87.5% and a specificity of 87.9%. This bottom lane model was the most successful model,

Table 4.2 – Logistic Regression summary results for all 6 models.

	Model 1 Middle					Model 2 Bottom–Support					Model 3 Bottom–Carry				
	Fact.Id	Estimate		Std. Error	OR	Fact.Id	Estimate		Std. Error	OR	Fact.Id	Estimate		Std. Error	OR
(Intercept)	—	-1.174	***	0.085	0.309	—	-4.851	***	0.108	0.008	—	29.477	***	0.217	6.336×10^{12}
Factor 1:	MR9	0.325	***	0.006	1.384	MR5	0.411	***	0.002	1.509	MR9	1.451	***	0.008	4.266
Factor 2:	MR10	-0.608	***	0.005	0.544	MR1	-0.389	***	0.002	0.678	MR2	-0.347	***	0.003	0.707
Factor 3:	MR1	-0.258	***	0.001	0.773	MR3	0.376	***	0.005	1.457	MR10	1.446	***	0.010	4.246
Factor 4:	MR4	1.084	***	0.005	2.956	MR4	0.165	***	0.004	1.180	MR6	-0.441	***	0.006	0.643
Factor 5:	MR6	-0.208	***	0.002	0.812	MR9	0.257	***	0.006	1.293	MR3	-0.216	***	0.005	0.806
Factor 6:	MR3	-0.190	***	0.002	0.827	MR7	0.957	***	0.014	2.603	MR4	0.337	***	0.017	1.401
Factor 7:	MR7	-0.039	***	0.001	0.962	MR2	-0.755	***	0.012	0.470	MR5	-0.349	***	0.012	0.705
Factor 8:	MR5	-0.009	***	0.001	0.991	MR8	0.026	***	0.005	1.027	MR1	0.052	***	0.002	1.053
Factor 9:	MR8	0.075	***	0.007	1.077	MR6	0.002		0.001	1.002	MR8	-0.046	***	0.004	0.955
Factor 10:	MR2	0.003	*	0.001	1.003						MR7	0.002		0.001	1.002
Factor 11:															

	Model 4 Bottom–Other					Model 5 Top					Model 6 Jungle				
	Fact.Id	Estimate		Std. Error	OR	Fact.Id	Estimate		Std. Error	OR	Fact.Id	Estimate		Std. Error	OR
(Intercept)	—	-1.850	***	0.178	0.157	—	0.261		0.180	1.298	—	-0.287	**	0.092	0.751
Factor 1:	MR5	0.445	***	0.007	1.560	MR10	0.685	***	0.008	1.984	MR5	0.264	***	0.002	1.302
Factor 2:	MR3	-0.143	***	0.005	0.867	MR4	-0.323	***	0.004	0.724	MR6	-0.189	***	0.002	0.828
Factor 3:	MR1	-0.185	***	0.004	0.831	MR3	0.879	***	0.007	2.410	MR2	0.436	***	0.003	1.547
Factor 4:	MR7	0.262	***	0.007	1.299	MR8	-0.870	***	0.013	0.419	MR1	0.110	***	0.002	1.116
Factor 5:	MR4	0.255	***	0.008	1.290	MR11	-0.427	***	0.006	0.652	MR8	-2.119	***	0.016	0.120
Factor 6:	MR6	-0.363	***	0.022	0.696	MR1	-0.304	***	0.003	0.738	MR9	0.274	***	0.002	1.315
Factor 7:	MR2	0.150	***	0.007	1.161	MR9	0.289	***	0.004	1.335	MR7	0.158	***	0.001	1.171
Factor 8:	MR8	-0.101	***	0.005	0.904	MR5	-0.035	***	0.001	0.966	MR3	-0.272	***	0.004	0.762
Factor 9:						MR7	-0.024	***	0.001	0.976	MR4	-0.030	***	0.002	0.971
Factor 10:						MR2	0.020	***	0.001	1.020					
Factor 11:						MR6	-0.068	***	0.007	0.934					

by over 4% accuracy, but had the second lowest CoD. Despite this, the model contains one factor with non-significant contributions, 'MR7'. The Odds Ratio for this factor also suggests this factor does not meaningfully contribute to the model. 'MR7' for the Bottom lane Carry player rewards dealing 'True' damage as well as using 'crowd control' abilities

Model 4: Bottom Lane [Other]. This model successfully predicted team performance for 78.31% of our test set (and 78.5% internal accuracy) with a sensitivity of 67.1% and a specificity of 85.5%. The Bottom lane 'Other' model is the least sensitive in that it has a larger tendency to mis-classify wins as losses than the other models. The lackluster performance of this model may be because the model was defined by the roles that did not fit with the larger groups, compounded with the fact that the sample size for this model was nearly one-tenth the size of the other subsets.

Model 5: Top Lane. This model successfully predicted team performance for 82.19% of our test set (and 81.8% internal accuracy) with a sensitivity of 81.9% and a specificity of 81.6%. The Top lane model involved the largest number of factors (11). However, several factors, though significant, may not be meaningful contributors to the model – 'MR7', 'MR2', and 'MR5'. 'MR7' is identical to 'MR6' for the Bottom lane Support players and rewards dealing true damage. However, for the Top lane this factor negatively contributes to the model. The other two factors 'MR2' and 'MR5' penalize and reward physical damage, respectively. As 'MR5' contributes negatively to the model, a Top lane player dealing physical damage reduces their odds of winning.

Model 6: Jungle. This model successfully predicted team performance for 75.05% of our test set (and 74.9% internal accuracy) with a sensitivity of 74.8% and a specificity of 75.0%. This model is one of the most well-specified as all of the factors seem to contribute meaningfully to the model. However, the Odds Ratio of .971 for 'MR4' can be argued against as that is a relatively small contribution to the model. This factor seems to capture the balance between dealing magical damage and physical damage by rewarding the former and penalizing the later. As this factor contributes negatively to the model the desired values of these variable is reverse in that magical damage reduces the odds of winning and physical damage improves the odds of winning.

Overall each model predicted match outcome with better accuracy than chance. In addition, the models' Coefficients of Discrimination, due to the measure's relation to R^2 (Tjur, 2009), suggest that each model provides a reasonable fit for the data.

Table 4.3 – Logistic Regression Results. This table provides the test-set Prediction Accuracy for each model's regression. Additionally included are the Coefficient of Discrimination (Tjur, 2009) and the Area Under the Receiver Operating Characteristic Curve (AUROC).

Model	Accuracy	CoD	AUROC
Model 1: Middle	83.23%	0.518	0.909
Model 2: Bottom [Support]	74.89%	0.631	0.830
Model 3: Bottom [Carry]	87.44%	0.328	0.944
Model 4: Bottom [Other]	78.31%	0.380	0.860
Model 5: Top	82.19%	0.483	0.901
Model 6: Jungle	75.05%	0.312	0.826

4.4 SUMMARY

We set out to develop a method by which individual performance can be assessed within the context of a team based on behavioral data. To accomplish this goal we applied exploratory factor analysis to unique subsets, based on the lane and role the player fulfilled, from 1,919,187 match records (resulting from 539,832 distinct matches). We assessed the performance of the set of factors created for each subset based on whether these factors could be fit to and predict the game outcome. Each model was able to successfully predict game outcome for a test set (20% sample of each subset aside) with accuracy much greater than chance.

Chapter 5
DISCUSSION

This paper used massive amounts of game data from League of Legends matches to develop an 'absolute' behavior-based performance metric that is effective at assessing individual contribution to task outcome in a team environment. The goal of this research was to begin laying groundwork to apply theories of learning to individuals within a team setting, and eventually applying theories of individual learning to team learning. In order to do so, we investigated individual performance based on the behavioral data resulting from gameplay for 484,926 different players in 539,832 matches.

The exploration of the data set revealed 6 distinct groupings (subsets) based on the part of the map the player chose to play in (the middle, bottom, top, or jungle) that had different correlation structures (see Figures B.1 through B.6). We found evidence for a different latent factor structure of the behavioral variables for each subset that was able to capture interesting trade offs for performance in a given lane. For instance, there is some balance that must be struck between magic damage taken and physical damage taken for Middle lane players (See Table C.1).

We evaluated the efficacy of the factors for predicting game outcome by modeling the results of the EFA on the match outcome. This was based on the assumption that, good behaviors tend to result in winning, but are not perfectly indicative of team performance. Thus, we expected to find reasonably successful prediction accuracies for each model, with some instances being 'misclassified'. Our results suggest that we accomplished this goal as our models were 74-87% accurate at predicting game outcome.

Looking at the model structures in combination with the factor structures we see additional interesting tradeoffs such as the one for Bottom lane Support players (see figure C.2) for factors 'MR7' and 'MR2'. Both factors reward neutral minions killed, however 'MR7' rewards neutral minions killed in the enemy jungle while 'MR2' rewards neutral minions killed in the team's jungle. As mentioned, 'MR7' likely represents a measure of good positioning for the Support player. However, as 'MR2' is penalized in the model structure it seems that it is a measure of poor positioning. Within the context of the game, this seems plausible because a support who spends too much time in the team's jungle may interfere with the Jungle player's role and also is not as able to provide assistance to lanes in the case of surprise attacks.

The differences between model success can easily be attributed to the differences in sample size, or even the differences in proportion of the sample represented by the most popular role. This interpretation, however, may dismiss one very important possibility– certain roles on the team may be more individually important than other roles. For instance, the most successful model, the "Middle Lane" model, makes intuitive sense to be the most predictive of team success. Much of the early game (the first 10 minutes) for a MOBA hinges on the success of the player in the Middle lane. The player's entire job is to level as fast as they can so that they may participate in team fights and help other lanes accomplish objectives. So it would follow that a player 'achieving success' in the middle lane would be most predictive of team-level success.

There are a few limitations to this research that come about as a result of the manner in which the data was collected. The variables used are all collected after the end of the game which means that many of the value reflected may be a result of the fact that the team was already winning and was, thus, able to accomplish different things. For example, the variable 'neutralMinionsKilledEnemyJungle' loaded on the most influential factor for predicting game outcome, for each of the six models. While we were able to give some insight into why this happened for each model, it very well may be a result of the fact that a team that wins the game, by the very nature of winning, has spent more time on the enemy's side of the map. This is to say that, the act of winning the match provided the team the opportunity to kill more neutral (Jungle) minions in the enemy Jungle.

This research continues the work of (Sangster et al., 2016) and offers the next step in individual-focused evaluation of team performance with the aid of Big Data. Through this, we provide initial exploration of a method to determine an absolute metric of individual performance within the context of a team. The development of such a metric offers a necessary tool for future individual-focused assessments of teamwork and team performance and offers a solution to the difficulties raised by relative measures of performance.

REFERENCES

Braasch, J. (2011). A cybernetic model approach for free jazz improvisations. *Kybernetes*, *40*, 984–994. doi:10.1108/03684921111160214

Brochard, R., Dufour, A., & Despres, O. (2004). Effect of musical expertise on visuospatial abilities: evidence from reaction times and mental imagery. *Brain and Cognition*, *54*(2), 103–109.

Concepcion, J. (2015). How do you measure hustle in the NBA? there's a stat for that. Retrieved September 28, 2016, from http://grantland.com/the-triangle/how-do-you-measure-hustle-in-the-nba-theres-a-stat-for-that/

Destefano, M., Lindstedt, J. K., & Gray, W. D. (2011). Use of complementary actions decreases with expertise. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *33rd annual conference of the cognitive science society* (pp. 2014–2709). Austin, TX: Cognitive Science Society.

Donaldson, S. (2015). Mechanics and metagame: exploring binary expertise in League of Legends. *Games and Culture*. doi:10.1177/1555412015590063

Ebbinghaus, H. (1913). *Memory* (H. A. Rueger & C. E. Bussenius, Trans.). New York, NY: Teachers College, Columbia University. (Original work published 1885).

Elo, A. (1978). *The rating of chess players, past and present*. London, England: Batsford.

Ericsson, K. A. & Lehmann, A. C. (1996). Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273–305. doi:10.1146/annurev.psych.47.1.273

Ferrari, S. (2013). From generative to conventional play: MOBA and League of Legends. *Proceedings of DiGRA 2013: defragging game studies.*

Gobet, F. & Simon, H. A. (2000). Five seconds or sixty? presentation time in expert memory. *Cognitive Science*, *24*, 651–682. doi:10.1016/S0364-0213(00)00031-8

Goldstone, R. L. & Lupyan, G. (2016). Discovering psychological principles by mining naturally occurring data sets. *Topics in Cognitive Science*, *8*(3).

Gong, D., He, H., Liu, D., Ma, W., Dong, L., Luo, C., & Yao, D. (2015). Enhanced functional connectivity and increased gray matter volume of insula related to action video game playing. *Scientific Reports*, *5*. doi:10.1038/srep09763

Henson, R. K. & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: common errors and some comment on improved practice. *Educational and Psychological Measurement*, *66*, 393–416. doi:10.1177/0013164405282485

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179–185. doi:10.1007/BF02289447

Howard, R. W. (2014). Learning curves in highly skilled chess players: a test of the generality of the power law of practice. *Acta Psychologica*, *151*, 16–23. doi:10.1016/j.actpsy.2014.05.013

Huang, J., Yan, E., Cheung, G., Nagappan, N., & Zimmermann, T. (2017). Master maker: understanding gaming skill through practice and habit from gameplay behavior. *Topics in Cognitive Science*, 1–30.

Huang, J., Zimmermann, T., Nagappan, N., Harrison, C., & Phillips, B. (2013). Mastering the art of war: how patterns of gameplay influence skill in Halo. In *Proceedings of the international conference on human factors in computing systems (CHI 2013)* (pp. 695–704). New York, NY: ACM Press.

Imbeault, F., Bouchard, B., & Bouzouane, A. (2011). Serious games in cognitive training for alzheimer's patients. *2011 IEEE 1st international conference on serious games and applications for health (SeGAH)*, Braga, 1–8. IEEE.

Jastrzembski, T. S., Gluck, K. A., & Rodgers, S. (2009). Improving military readiness: a state-of-the-art cognitive tool to predict performance and optimize training effectiveness. In *The interservice/industry training, simulation, and education conference (I/ITSEC)* (9024, pp. 1–11). Orlando, FL: NTSA.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151. doi:10.1177/001316446002000116

Kirsh, D. & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513–549. doi:10.1207/s15516709cog1804_1

Latham, A. J., Patston, L. L. M., & Tippett, L. J. (2013). Just how expert are "expert" video-game players? assessing the experience and expertise of video-game players across "action" video-game genres. *Frontiers in Psychology*, *4*, 1–3. doi:10.3389/fpsyg.2013.00941

Lindstedt, J. K. & Gray, W. D. (2013). Extreme expertise: exploring expert behavior in Tetris. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual conference of the cognitive science society* (pp. 912–917). Cognitive Science Society. Austin, TX. Retrieved from http://csjarchive.cogsci.rpi.edu/Proceedings/2013/papers/0183/paper0183.pdf

Maas, H. L. J. V. D. & Wagenmakers, E.-J. (2005). A psychometric analysis of chess expertise. *American Journal of Psychology*, *118*(1), 29–60.

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: a meta-analysis. *Psychological Science*, *25*, 1608–1618. doi:10.1177/0956797614535810

Multiplayer online battle arena. (n.d.). In Wikipedia. Retrieved from https://en.wikipedia.org/w/index.php?title=Multiplayer_online_battle_arena&oldid=739229947

Patston, L. L. M., Hogg, S. L., & Tippett, L. J. (2007). Attention in musicians is more bilateral than in non-musicians. *Laterality*, *12*(3), 262–272.

Raîche, G. (2010). *An r package for parallel analysis and non graphical solutions to the cattell scree test*. Retrieved from http://cran.r-project.org/package=nFactors

Raîche, G., Walls, T. A., Magis, D., Riopel, M., & Blais, J.-G. (2013). Non graphical solutions for the Cattell's scree test. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *9*, 23–29. doi:10.1027/1614-2241/a000051

Revelle, W. (2016). *Psych: procedures for psychological, psychometric, and personality research*. Northwestern University. Evanston, Illinois. Retrieved from https://cran.r-project.org/package=psych

Sangster, M.-D. D., Mendonça, D. J., & Gray, W. D. (2016). Big data meets team expertise in a dynamic task environment. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 60, *1*, pp. 158–162). doi:10.1177/1541931213601036

Schellenberg, E. G. (2001). Music and nonmusical abilities. *Annals of the New York Academy of Sciences*, *930*(1), 355–371.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Simon, H. A. & Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Psychology*, *5*, 29–46. Retrieved from http://www.psyc.nott.ac.uk/research/credit/projects/CHREST/techrep74.pdf

Sims, V. K. & Mayer, R. E. (2002). Domain specificity of spatial expertise: the case of video game players. *Applied Cognitive Psychology*, *16*, 97–115. doi:10.1002/acp.759

Sloboda, J. A. (1991). Musical expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: prospects and limits* (Chap. 6, pp. 153–171). Cambridge, England: Cambridge University Press.

Stafford, T. & Dewar, M. (2014). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychological Science*, *25*(2), 511–518.

Tjur, T. (2009). Coefficients of determination in logistic regression models—a new proposal: the coefficient of discrimination. *The American Statistician*, *63*, 366–372. doi:10.1198/tast.2009.08210

Weber, B. G., Mateas, M., & Jhala, A. (2011). Building human-level AI for real-time strategy games. In *Proceedings of AIIDE fall symposium on advances in cognitive systems* (Vol. 11, pp. 329–336). Stanford, Palo Alto, California: AAAI Press.

Appendix A
VARIABLE DEFINITIONS

The following table presents a set of useful terms and their definitions. The second table, Table A.2, provides the variables of interest for our Exploratory Factor Analysis and the explanation of their relevance in the context of a match. Each variable is provided for each player, individually, for each match.

Table A.1 – Useful Terms. This table provides explanations for common terms that are important for understanding Table A.2.

Variable	Explanation
assist	Awarded to any allied player who assisted (dealt damage, used an ability, etc.) with the kill in the preceding ten seconds.
crowd control	An ability that hinders an enemy's control over their own unit.
death	Occurs when a player is killed. The dead player respawns in his/her team's base after a short duration. This duration increases over the course of a match.
kill	Occurs when a player on one team deals fatal damage to an opponent.
ward	An object that removes the Fog of War for a certain duration to the surrounding radius when placed.

Table A.2 – Variable Definitions. Variables denoted with an '*' did not load on any factors for any of the models.

Variable	Explanation
assists	Total number of assists by player for match.
deaths	Total number of kills by player for match.
doubleKills	Number of times player killed two enemy players within 10 seconds.
firstBloodAssist*	Boolean of whether player received an assist for the first kill of the game.
firstBloodKill	Boolean of whether player received the first kill of the game.
firstInhibitorAssist	Boolean of whether player received an assist for destroying the first enemy inhibitor (building).
firstInhibitorKill	Boolean of whether player received a kill for destroying the first enemy inhibitor (building).
firstTowerAssist*	Boolean of whether player received an assist for destroying the first tower of the game.
firstTowerKill*	Boolean of whether player received a kill for destroying the first tower of the game.
goldEarned	Total gold earned for the match.
goldSpent	Total gold spent for the match.
inhibitorKills	Number of times player destroyed an enemy inhibitor.
killingSprees	Number of times player killed at least three enemy players without dying.
kills	Total number of kills by player for the match.
largestCriticalStrike	Amount of damage dealt with largest critical strike.
largestKillingSpree	Largest number of kills (¿2) for player in a single life.
largestMultiKill	Largest number of kills (¿1) for player within 10 second window.
magicDamageDealt	Total amount of Magic-type damage dealt by the player in the match.
magicDamageDealtToChampions	Total amount of Magic-type damage dealt by the player to other Champions in the match.
magicDamageTaken	Total amount of Magic-type damage received in the match.
minionsKilled	Total number of minions killed by the player in the match.
neutralMinionsKilled	Total number of neutral (jungle) minions killed by the player in the match.
neutralMinionsKilledEnemyJungle	Number of neutral (jungle) minions killed by the player in the enemy jungle.
neutralMinionsKilledTeamJungle	Number of neutral (jungle) minions killed by the player in the team's jungle.
pentaKills	Number of times player killed 5 enemy players within 30 seconds.
physicalDamageDealt	Total amount of Physical-type damage dealt by the player in the match.
physicalDamageDealtToChampions	Total amount of Physical-type damage dealt by the player to other Champions in the match.
physicalDamageTaken	Total amount of Physical-type damage received in the match.
quadraKills*	Number of times player killed 4 enemy players within 30 seconds.
sightWardsBoughtInGame	Number of Sight wards purchased throughout the game.
totalDamageDealt	Total amount of damage dealt by the player in the match.
totalDamageDealtToChampions	Total amount of damage dealt by the player to other Champions in the match.
totalDamageTaken	Total amount of damage received in the match.
totalHeal	Total amount of healing by the player.
totalTimeCrowdControlDealt	Total time player had enemy units under crowd control effects.
totalUnitsHealed	Number of units healed throughout the match.
towerKills	Number of times player landed killing blow on a tower.
tripleKills	Number of times player killed 3 enemy players within 10 seconds.
trueDamageDealt	Total amount of True-type damage dealt by the player in the match.
trueDamageDealtToChampions	Total amount of True-type damage dealt by the player to other Champions in the match.
trueDamageTaken	Total amount of True-type damage received in the match.
visionWardsBoughtInGame	Number of Vision wards purchased throughout the game.
wardsKilled	Number of wards destroyed by the player.
wardsPlaced	Number of wards placed by the player.

Appendix B
CORRELATION PLOTS

The following plots are the individual correlation plots for each modeled subset. The upper and lower diagonals are mirror images. Larger and darker circles represent stronger correlations.

Figure B.1 – Correlation Plot for Match-level Behavioral Variables for the "Middle Lane" Model (Model 1). For variable explanations see Table A.2.

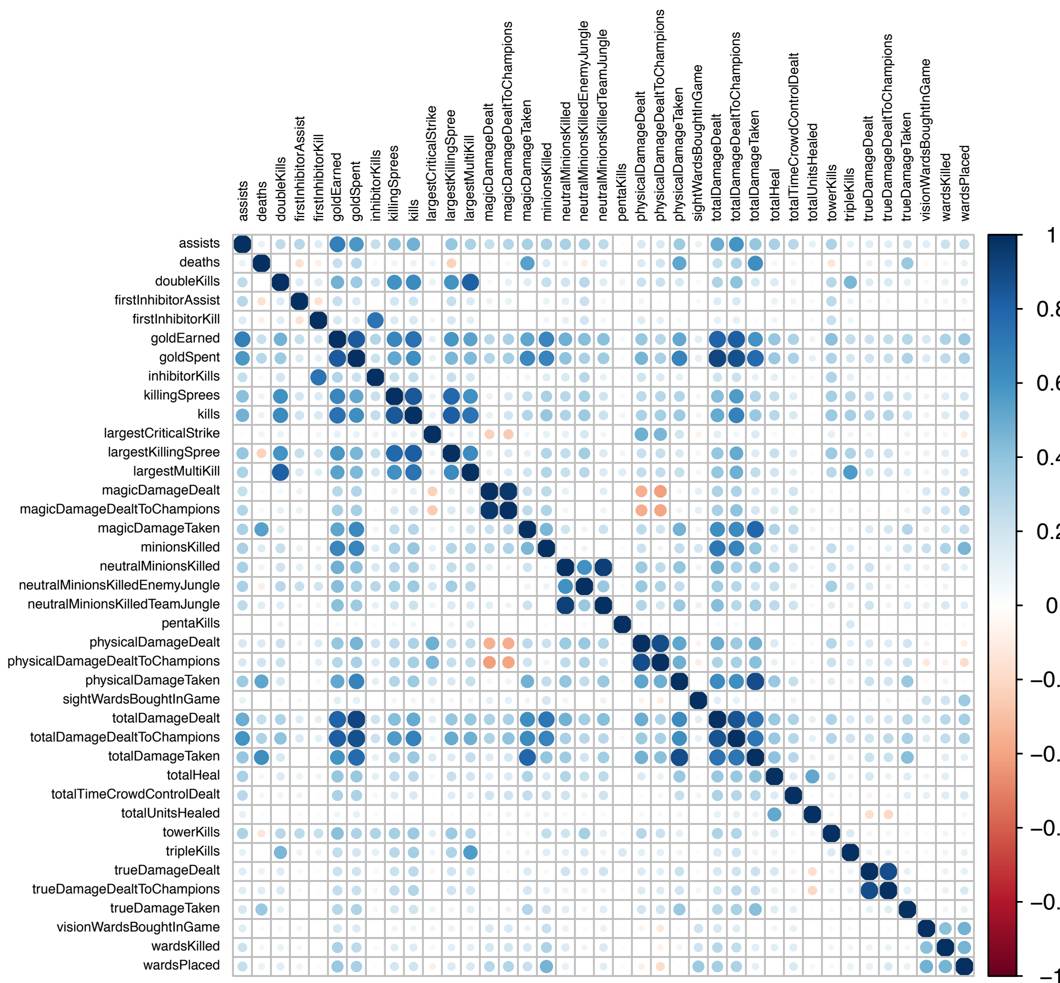


Figure B.2 – Correlation Plot for Match-level Behavioral Variables for the "Bottom [Support] Lane" Model (Model 2). For variable explanations see Table A.2.

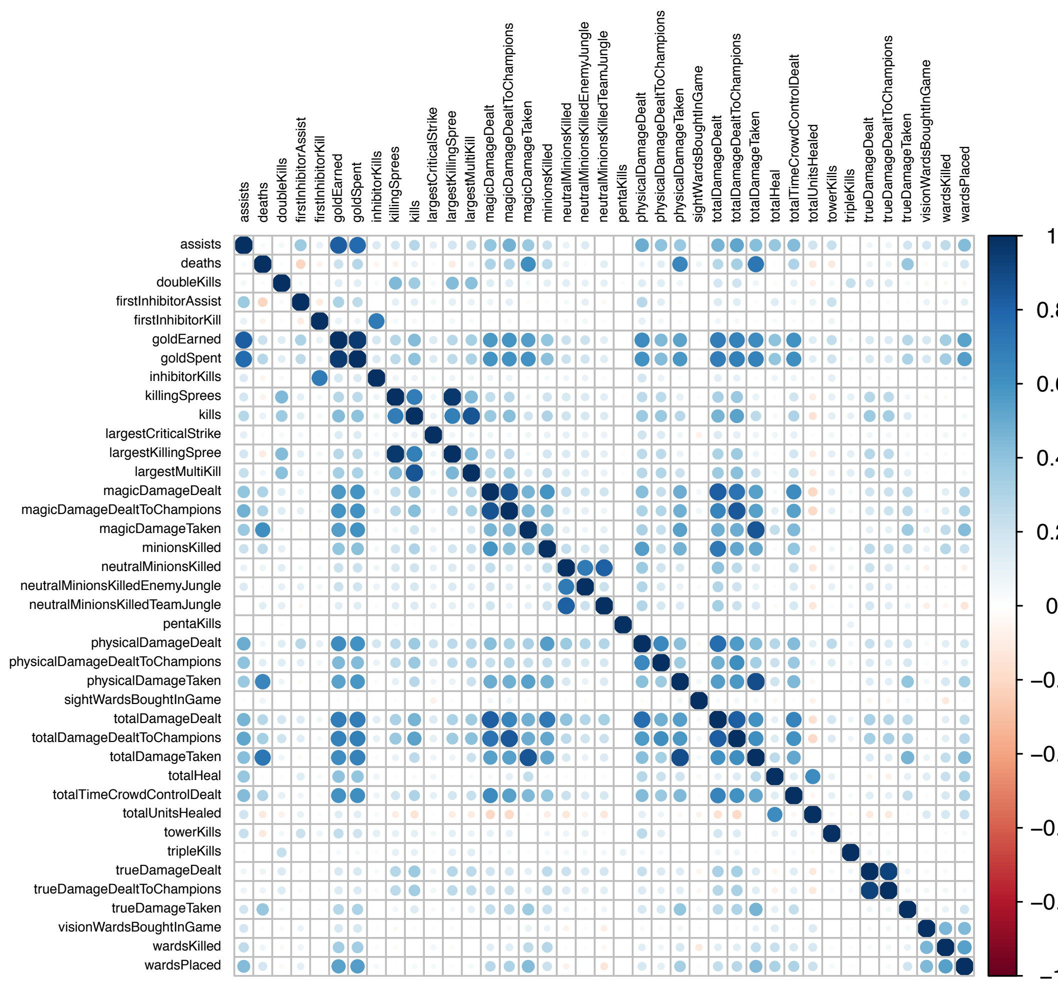


Figure B.3 – Correlation Plot for Match-level Behavioral Variables for the "Bottom [Carry] Lane" Model (Model 3). For variable explanations see Table A.2.

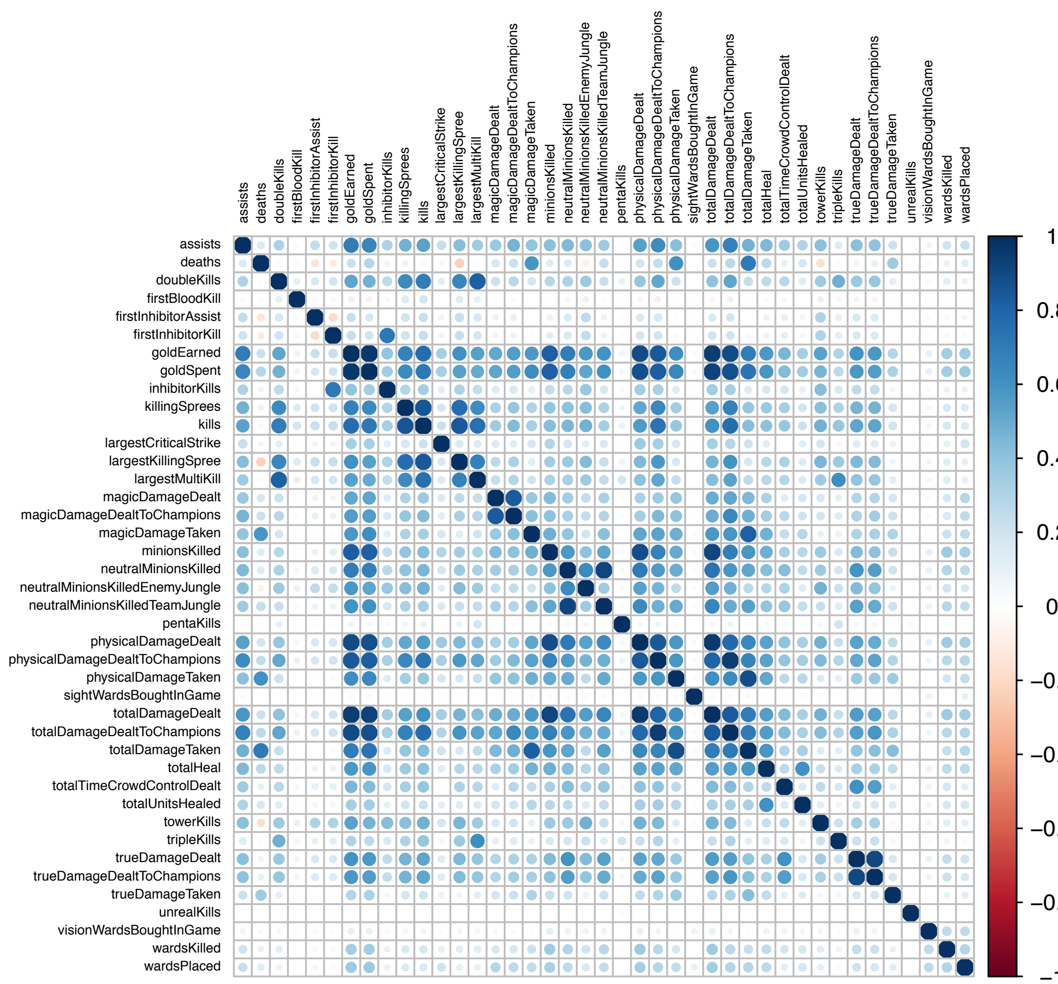


Figure B.4 – Correlation Plot for Match-level Behavioral Variables for the "Bottom [Other] Lane" Model (Model 4). For variable explanations see Table A.2.

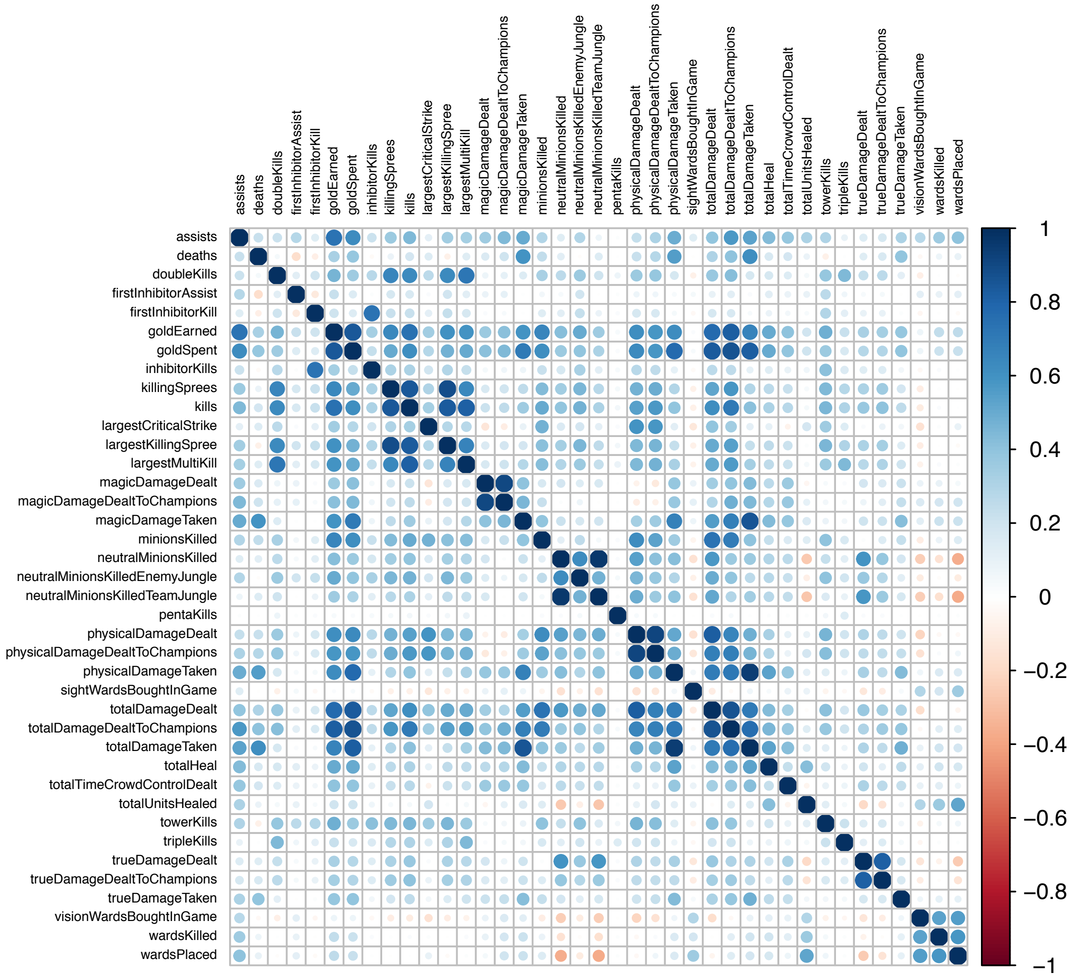


Figure B.5 – Correlation Plot for Match-level Behavioral Variables for the “Top Lane” Model (Model 5).For variable explanations see Table A.2.

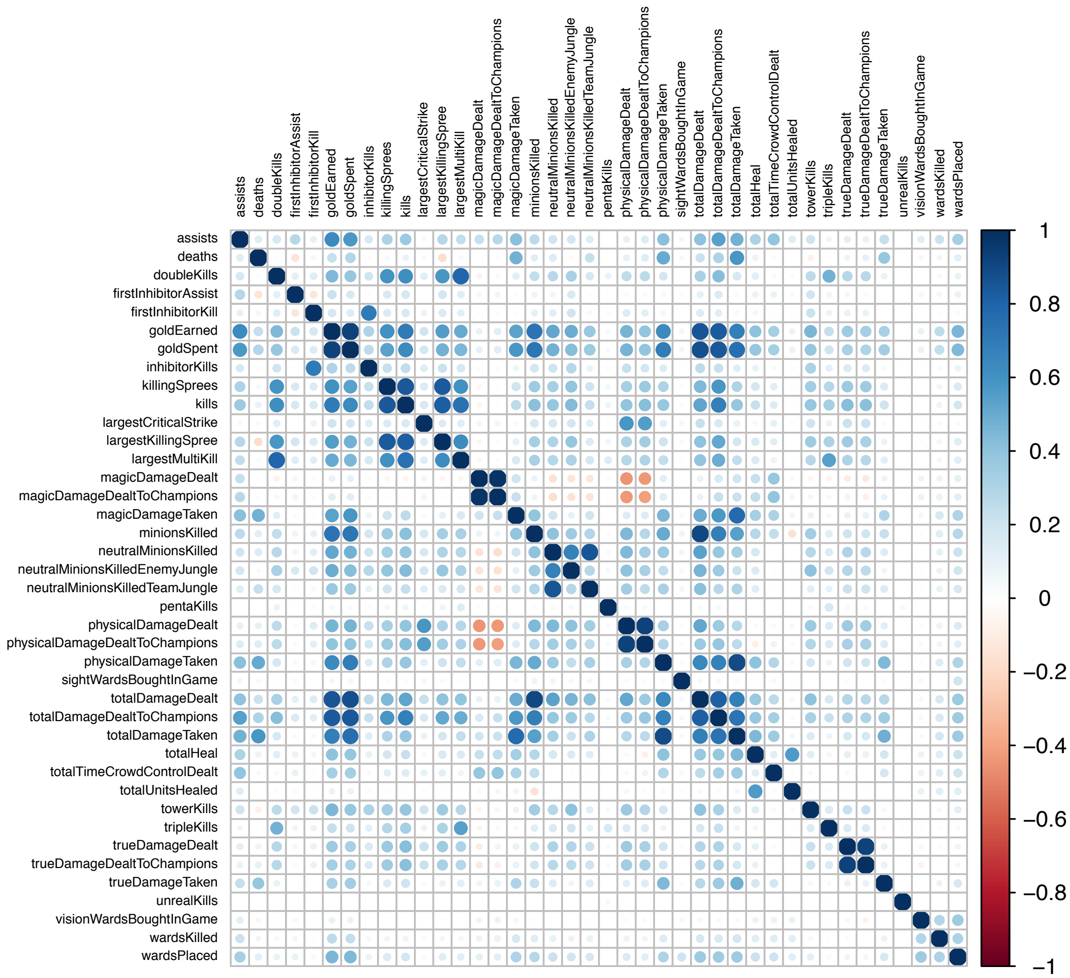
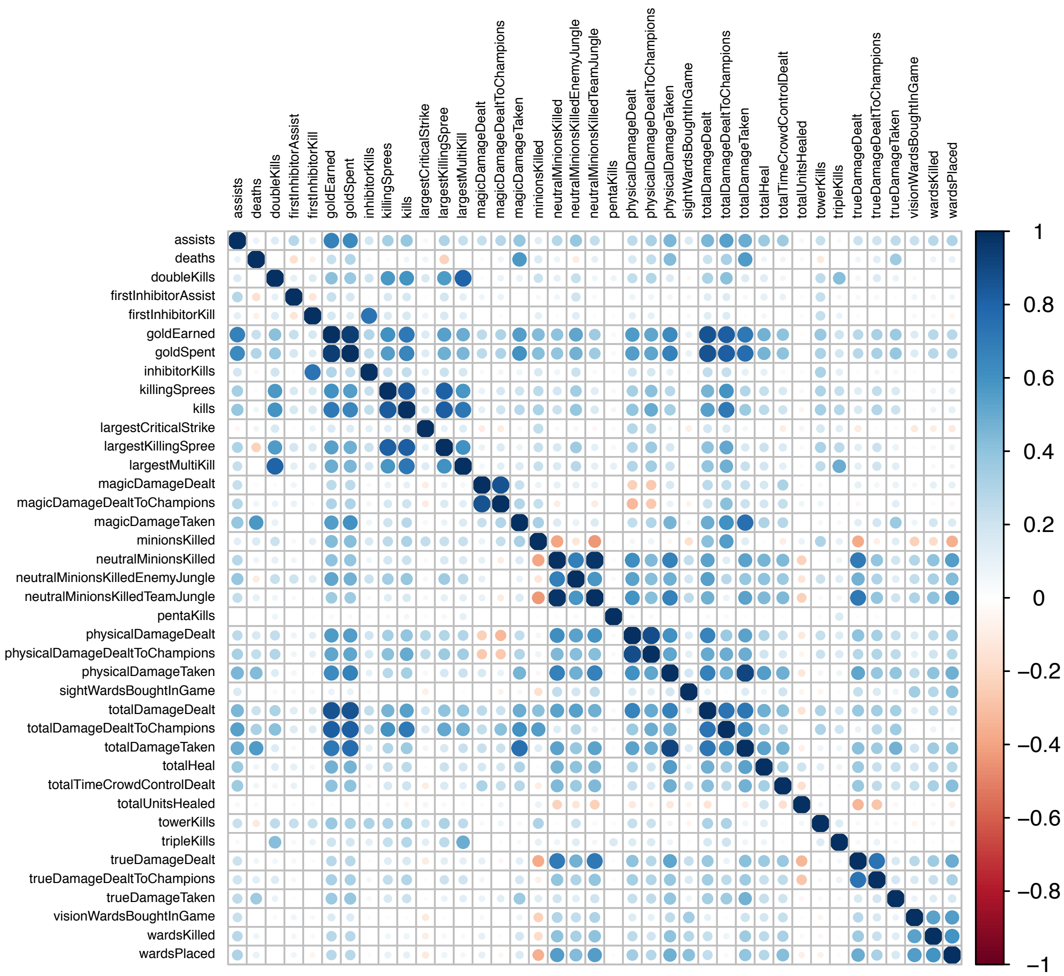


Figure B.6 – Correlation Plot for Match-level Behavioral Variables for the "Jungle Lane" Model (Model 6).For variable explanations see Table A.2.



Appendix C
FACTOR LOADING

The following tables show the factor loading scores for each of the six models. Variables that had no variance throughout the subset were excluded in the Exploratory Factor Analysis. Loading values are colored green if the variable loading was $> .3$, red if the variable loading was $< -.3$, and omitted if the variable did not load on the factor. Factors are listed in order of eigenvalues (i.e.the variance in all variables accounted for by the factor) from left to right. SS Loadings Adjusted are the adjusted values accounting for variable loading scores that were omitted.

Table C.1 – Factor Loadings for 'Middle Lane' Model (Model 1). For variable explanations see Table A.2.

	MR1	MR4	MR6	MR10	MR7	MR3	MR2	MR5	MR9	MR8
assists	0.33	0.53								
deaths	0.69									
doubleKills				0.77						
firstInhibitorAssist		0.33								
firstInhibitorKill									0.72	
goldEarned	0.46	0.54	0.45							
goldSpent	0.64	0.33	0.47							
inhibitorKills									0.73	
killingSprees		0.67		0.46						
kills		0.70		0.52						
largestCriticalStrike					0.49					
largestKillingSpree		0.70		0.48						
largestMultiKill		0.33		0.80						
magicDamageDealt							0.90			
magicDamageDealtToChampions							0.88			
magicDamageTaken	0.77		0.30							-0.45
minionsKilled			0.72							
neutralMinionsKilled						0.92				
neutralMinionsKilledEnemyJungle		0.33				0.45			0.39	
neutralMinionsKilledTeamJungle						0.92				
pentaKills										
physicalDamageDealt					0.85					
physicalDamageDealtToChampions					0.80					
physicalDamageTaken	0.84									0.35
sightWardsBoughtInGame										
totalDamageDealt	0.57		0.58		0.32					
totalDamageDealtToChampions	0.63	0.45	0.44							
totalDamageTaken	0.95									
totalHeal	0.38									
totalTimeCrowdControlDealt										
totalUnitsHealed										
towerKills		0.35							0.33	
tripleKills				0.56						
trueDamageDealt								0.90		
trueDamageDealtToChampions								0.90		
trueDamageTaken	0.42									
visionWardsBoughtInGame			0.35							
wardsKilled			0.39							
wardsPlaced			0.57							
SS Loadings	5.104	3.246	2.750	2.590	2.313	2.255	1.811	1.793	1.722	0.406
SS Loadings Adjusted	4.431	2.762	2.169	2.277	1.700	1.891	1.580	1.618	1.309	0.321

Table C.2 – Factor Loadings for 'Bottom [Support] Lane' Model (Model 2). For variable explanations see Table A.2.

	MR1	MR4	MR5	MR3	MR8	MR6	MR2	MR7	MR9
assists			0.77						
deaths	0.79								
doubleKills		0.47							
firstInhibitorAssist			0.39						
firstInhibitorKill									
goldEarned	0.37		0.78						
goldSpent	0.45		0.74						
inhibitorKills									
killingSprees		0.95							
kills		0.75							
largestCriticalStrike									
largestKillingSpree		0.96							
largestMultiKill		0.55							
magicDamageDealt	0.33			0.74	0.44				
magicDamageDealtToChampions	0.34			0.86					
magicDamageTaken	0.81								
minionsKilled	0.37				0.62				
neutralMinionsKilled							0.78	0.56	
neutralMinionsKilledEnemyJungle								0.92	
neutralMinionsKilledTeamJungle							0.95		
pentaKills									
physicalDamageDealt			0.36		0.64				0.43
physicalDamageDealtToChampions									0.72
physicalDamageTaken	0.83								
sightWardsBoughtInGame									
totalDamageDealt	0.34			0.44	0.68				
totalDamageDealtToChampions	0.39	0.32		0.62					0.38
totalDamageTaken	0.94								
totalHeal			0.52						
totalTimeCrowdControlDealt	0.33		0.30	0.37	0.36				
totalUnitsHealed			0.40	-0.31					
towerKills									
tripleKills									
trueDamageDealt						0.90			
trueDamageDealtToChampions						0.91			
trueDamageTaken	0.45								
visionWardsBoughtInGame									
wardsKilled			0.41						
wardsPlaced	0.34		0.56						
SS Loadings	4.545	3.658	3.653	2.535	2.087	1.828	1.697	1.228	1.185
SS Loadings Adjusted	4.253	3.015	3.017	2.094	1.597	1.647	1.521	1.165	0.849

Table C.3 – Factor Loadings for 'Bottom [Carry] Lane' Model (Model 3). For variable explanations see Table A.2.

	MR1	MR6	MR2	MR7	MR8	MR3	MR9	MR10	MR5	MR4
assists	0.38							0.44		
deaths			0.79							
doubleKills		0.80								
firstBloodKill										
firstInhibitorAssist								0.33		
firstInhibitorKill							0.77			
goldEarned	0.70	0.37	0.36							
goldSpent	0.68	0.32	0.42							
inhibitorKills							0.77			
killingSprees		0.66						0.33		
kills		0.74						0.36		
largestCriticalStrike	0.35									
largestKillingSpree		0.72						0.34		
largestMultiKill		0.85								
magicDamageDealt					0.78					
magicDamageDealtToChampions					0.81					
magicDamageTaken	0.32		0.71						0.54	
minionsKilled	0.86									
neutralMinionsKilled	0.41					0.80				
neutralMinionsKilledEnemyJungle	0.35					0.41	0.30			
neutralMinionsKilledTeamJungle	0.36					0.75				
pentaKills										
physicalDamageDealt	0.86									
physicalDamageDealtToChampions	0.60	0.41	0.41							0.45
physicalDamageTaken			0.80						-0.44	
sightWardsBoughtInGame										
totalDamageDealt	0.81		0.32							
totalDamageDealtToChampions	0.54	0.41	0.43		0.36					0.33
totalDamageTaken	0.35		0.89							
totalHeal	0.36		0.44							
totalTimeCrowdControlDealt				0.51						
totalUnitsHealed										
towerKills	0.38						0.35			
tripleKills		0.57								
trueDamageDealt				0.86						
trueDamageDealtToChampions				0.84						
trueDamageTaken			0.38							
unrealKills										
visionWardsBoughtInGame										
wardsKilled	0.32									
wardsPlaced	0.31									
SS Loadings	5.892	4.503	4.394	2.482	1.946	1.764	1.717	1.073	0.497	0.375
SS Loadings Adjusted	5.134	3.768	3.657	1.700	1.391	1.371	1.408	0.656	0.486	0.309

Table C.4 – Factor Loadings for 'Bottom [Other] Lane' Model (Model 4). For variable explanations see Table A.2.

	MR1	MR5	MR3	MR7	MR2	MR4	MR8	MR6
assists	0.39	0.33		0.60				
deaths	0.66							
doubleKills		0.72						
firstInhibitorAssist								
firstInhibitorKill								
goldEarned	0.49	0.55	0.30	0.31				
goldSpent	0.69	0.35	0.33					
inhibitorKills		0.32						
killingSprees		0.86						
kills		0.87						
largestCriticalStrike			0.60					
largestKillingSpree		0.87						
largestMultiKill		0.77						
magicDamageDealt	0.31					0.88		
magicDamageDealtToChampions	0.32					0.89		
magicDamageTaken	0.81							
minionsKilled		0.36	0.54					0.45
neutralMinionsKilled					0.86			
neutralMinionsKilledEnemyJungle		0.41			0.55			
neutralMinionsKilledTeamJungle				-0.33	0.84			
pentaKills								
physicalDamageDealt	0.38	0.32	0.76					
physicalDamageDealtToChampions	0.38	0.39	0.77					
physicalDamageTaken	0.90							
sightWardsBoughtInGame								
totalDamageDealt	0.57	0.34	0.51					0.39
totalDamageDealtToChampions	0.64	0.45	0.42					
totalDamageTaken	0.96							
totalHeal	0.44							
totalTimeCrowdControlDealt	0.35							
totalUnitsHealed				0.57				
towerKills		0.45	0.35					
tripleKills		0.38						
trueDamageDealt					0.34		0.76	
trueDamageDealtToChampions							0.81	
trueDamageTaken	0.46							
visionWardsBoughtInGame				0.59				
wardsKilled				0.62				
wardsPlaced				0.73				
SS Loadings	6.017	5.662	2.945	2.782	2.496	2.061	1.535	0.700
SS Loadings Adjusted	5.442	5.206	2.585	2.151	1.865	1.557	1.234	0.361

Table C.5 – Factor Loadings for 'Top Lane' Model (Model 5). For variable explanations see Table A.2.

	MR1	MR11	MR6	MR2	MR5	MR3	MR4	MR7	MR9	MR10	MR8
assists	0.33					0.40			0.50		
deaths	0.66										
doubleKills		0.77									
firstInhibitorAssist						0.31					
firstInhibitorKill										0.47	
goldEarned	0.43		0.51			0.40			0.35		
goldSpent	0.52		0.53			0.32			0.32		
inhibitorKills										0.52	
killingSprees		0.57				0.58					
kills		0.63				0.57					
largestCriticalStrike					0.65						
largestKillingSpree		0.59				0.60					
largestMultiKill		0.82									
magicDamageDealt				0.96							
magicDamageDealtToChampions				0.95							
magicDamageTaken	0.63										0.63
minionsKilled			0.85								
neutralMinionsKilled							0.85				
neutralMinionsKilledEnemyJungle						0.38	0.37			0.66	
neutralMinionsKilledTeamJungle							0.89				
pentaKills											
physicalDamageDealt				-0.31	0.82						
physicalDamageDealtToChampions					0.86						
physicalDamageTaken	0.87										
sightWardsBoughtInGame											
totalDamageDealt	0.41		0.76								
totalDamageDealtToChampions	0.54		0.46			0.39					
totalDamageTaken	0.91										
totalHeal	0.33								0.47		
totalTimeCrowdControlDealt				0.37							
totalUnitsHealed									0.59		
towerKills										0.33	
tripleKills		0.54									
trueDamageDealt								0.89			
trueDamageDealtToChampions								0.90			
trueDamageTaken	0.48										
unrealKills											
visionWardsBoughtInGame											
wardsKilled									0.32		
wardsPlaced									0.39		
SS Loadings	4.265	3.251	2.708	2.335	2.293	2.225	1.937	1.926	1.659	1.420	0.576
SS Loadings Adjusted	3.781	2.638	2.059	2.060	1.815	1.839	1.657	1.597	1.303	1.027	0.403

Table C.6 – Factor Loadings for 'Jungle Lane' Model (Model 6). For variable explanations see Table A.2.

	MR5	MR6	MR7	MR1	MR2	MR4	MR3	MR9	MR8
assists		0.33	0.57						
deaths		0.73							
doubleKills	0.76								
firstInhibitorAssist									
firstInhibitorKill								0.36	
goldEarned	0.48	0.49	0.33		0.53				
goldSpent	0.41	0.57			0.51				
inhibitorKills								0.37	
killingSprees	0.80								
kills	0.86								
largestCriticalStrike									
largestKillingSpree	0.83								
largestMultiKill	0.82								
magicDamageDealt						0.89			
magicDamageDealtToChampions						0.90			
magicDamageTaken		0.87							
minionsKilled			-0.34	-0.33	0.64				
neutralMinionsKilled			0.41	0.86					
neutralMinionsKilledEnemyJungle			0.47	0.50					
neutralMinionsKilledTeamJungle			0.39	0.86					
pentaKills									
physicalDamageDealt				0.45		-0.32	0.64		
physicalDamageDealtToChampions	0.34	0.33					0.75		
physicalDamageTaken		0.64	0.38	0.45					0.39
sightWardsBoughtInGame			0.37						
totalDamageDealt	0.32	0.46		0.43	0.61				
totalDamageDealtToChampions	0.52	0.58			0.40				
totalDamageTaken		0.85		0.31					
totalHeal		0.32		0.36					
totalTimeCrowdControlDealt			0.38						
totalUnitsHealed								0.41	
towerKills					0.38				
tripleKills	0.40								
trueDamageDealt			0.47	0.48				-0.55	
trueDamageDealtToChampions			0.36					-0.62	
trueDamageTaken		0.44							
visionWardsBoughtInGame			0.54						
wardsKilled			0.56						
wardsPlaced			0.66						
SS Loadings	4.951	4.533	3.535	3.309	2.158	2.127	1.408	1.364	0.333
SS Loadings Adjusted	4.357	4.055	2.909	2.866	1.614	1.707	0.975	1.122	0.153