Social Network Analysis of College and Professional Basketball

Carol Xu

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Renee Ha, Chair

James Ha

Massimo Maoret

Program Authorized to Offer Degree:

Psychology

University of Washington

Abstract

Social Network Analysis of College and Professional Basketball

Carol Xu

Chair of the Supervisory Committee:

Renee Ha

Department of Psychology

Social network analysis (SNA) allows us to quantify cooperative team sports such as basketball, and compare different passing patterns' effects on a successful game. We analyzed games at both the college and professional level to identify common network metrics that predict a win. For both levels of competition, individual players who scored the most points per game were unlikely to also occupy the central position in their team's network; however, we found opposite effects of out-degree centralization on college vs. professional basketball. Other factors that significantly increase chances of success at the professional level include a team's total number of passes and level of reciprocity. These results may be used as a tool for coaching staff to improve their teams' strategies in concrete, measurable ways, and help track the level of cooperation among their players more easily.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1. BACKGROUND AND SIGNIFICANCE

1.1 HISTORY AND DEVELOPMENT

Despite its recent resurgence in popularity, social network analysis (SNA) has a rich history as a metric for measuring social structures and interactions spanning nearly a century. Even as early as the 19th century, sociologists have been interested in studying the relationships between different social actors (see Freeman, 2004, for review); Auguste Comte first began writing about different levels of "social systems" and the interconnections between actors at each level in 1835 (Martineau, 1835). Throughout the 19th and early 20th centuries, various sociologists agreed on the main idea of two different types of relations among groups of individuals: personal ties between people who know each other on an intimate level, and impersonal ones linking people to those they have more formal relationships with, such as colleagues or members of a large club (Maine, 1861; Tönnies, 1855; Durkheim, 1893; Spencer, 1897; Cooley, 1909/1962). Other early researchers of network structure include Gustave LeBon (1897), who studied information flow through crowds, and Georg Simmel (1908), who believed sociology should study patterns of interactions. The earliest empirical data on human networks was also being collected by the 1800's (Morgan, 1871; Hobson, 1919), which included graphical representations of those networks. In their studies of kinship and inheritance, Macfarlane (1883), Galton and Watson (1875), and Bienaymé (1845) were among the pioneers of network modeling and mathematical analysis.

The 20th century saw the development of "sociometry" by Jacob Levy Moreno and Helen Hall Jennings (1934; 1938), which included the first modern use of the term "network", and refocused the field of sociology on structural analysis of networks. Their later work further included all four defining features of SNA (Freeman, 2004): structural analysis, systematic empirical data, graphical imagery, and explicit mathematical modeling. By the middle of the 20th century, however, progress on SNA seemed to have stagnated somewhat. There were fewer majorbreakthroughs in the 1940's, 1950;s, and 1960's, although it never completely faded into the background. Kurt Lewin formed the Research Center for Group Dynamics at MIT in 1945, which in the late 1940's and early 1950's conducted several studies on human behavior in groups (Patnoe, 1988). Alex Bavelas, one of Lewin's graduate students, published on the effects of node centrality on individual and network performance (1948, 1950). This era also saw the landmark paper which introduced the notion of a "clique", a subset of a larger network in which actors share stronger connections than the rest of the population (Luce & Perry, 1949). Cartwright and Harary (1956) applied network analysis to Fritz Heider's cognitive balance theory (Heider, 1946; Heider 1958), which focused on how individuals tend to seek a balance between their perceptions of other individuals, objects, or concepts, and their experiences with them. James Davis (1967) later adapted this model to reflect social structure. Charles Loomis, Leo Katz, and Claude Lévi-Strauss were also key contributors in keeping the sociometric/social network tradition alive (Loomis, 1946; Katz, 1947; Lévi-Strauss, 1949, as examples).

These periodic innovations continued into the next few decades. Other notable breakthroughs include Coleman, Katz, and Menzel's (1957) study of interpersonal factors in the spread of information between physicians, Peter Blau's book on homophily (1977), and Charles Kadushin's (1966) work on social circles, continuing the structural tradition of analysis. Alfred Radcliffe-Brown also strongly advocated the structural perspective of social systems (1957), and inspired numerous other social scientists to embrace this view; Freeman (1989) published a paper identifying the common mathematical methods in many of these structural studies. At MIT, Karl Deutsch and Ithiel de Sola Pool were among the first to apply networks to political science (Deutsch, 1953; Pool & Kochen, 1978); in fact, Pool and Kochen's paper influenced Stanley Milgram's development of his "small world" theory (Milgram, 1967). Other key contributors during the 1960's include Claude Flament, whose book *Applications of Graph Theory to Group Structure* (1963) is considered a foundation of modern SNA; Edward Laumann, who has studied topics ranging from politics (Laumann & Knoke, 1987) to sexual behavior (Laumann, Gagnon, Michael, & Michaels, 1994), and many of whose students are now also respected SNA researchers; James Davis's study on transitivity in relationships (Davis, 1970; Davis & Leinhardt, 1971); andJac Anthonisse and Frans Stokman, who developed computer programs using graph theory to analyze structural data; and Harrison White, who not only wrote seminal papers on subjects such as structural equivalence (Lorrain & Whtie, 1971), but also excelled as a teacher of structural analysis at Harvard University (Freeman, 2004). At last, social network analysis was exiting its "dark age" and emerging as the field we know in present day.

1.2 CURRENT APPLICATIONS

In its current form, SNA has cast a wide net in terms of providing a way to quantify social interactions and other relationships between players in a given group or population, ranging from business corporations to theoretical and methodological studies in sociology, psychology, physics, epidemiology, economics, and more. It has been called on to explain everything from the motivations, evolution, and favorable conditions of cooperation among humans and other species (Ohtsuki, Hauert, Lieberman, & Nowak, 2006; Apicella, Marlowe, Fowler, & Christaksi, 2012; Rand, Arbesman, & Christaksi, 2011), to how often specific chemicals interact with each other in biological processes (Guimera & Amaral, 2005; Vendruscolo, Dokholyan, Paci, & Karplus, 2002), to trade agreements between whole nations (Serano & Boguná, 2003; Fagiolo, Reyes, & Schiavo, 2008). We will now review a few of the largest fields using this analysis and how their various theories and methods compare to each other.

1.2.1 *In sports*

Social sports are an excellent candidate for SNA because each game or match can be categorized as a closed population made up of players as nodes, and the passing of a ball between those players as edges. It is usually easy to track every pass during a game, as well as its direction. Various studies have already been conducted in the field of sports psychology using social networks, including soccer (Clemente, Couceiro, Martins, & Mendes, 2015; Kooij et al., 2009) and basketball (Bourbousson, Poizat, Saury, & Seve, 2010; Fewell et al. 2012; Vaz de Melo, Almeida, & Loureiro, 2008). Clemente, Couceiro, Martins, & Mendes (2015) found that among the First Portuguese Soccer League, there was no single “star” player who monopolized the most central position of the team’s network, and that all players decreased their number of connections in the second half, suggesting a lower level of participation for the entire team. Similarly, Fewell et al. (2012) measured the degree centrality, clustering (whether there were any “subgroups” in a team’s overall network), entropy (a measure of the unpredictability in a team’s passes), flow centrality (a measure of a player’s importance within the network), and uphill/downhill flux (whether teams consistently passed to their best shooters) among the 2010 NBA play-offs’ first round as proxies of offensive strategies, and their effects on team success. They too discovered there was often no star player among these teams, and that those teams with higher entropy (aka more unpredictable passing patterns) performed better. Grund (2012) found similar results with men’s soccer teams; lower degree centralization was associated with better performance.

Taken together, these studies would seem to imply that having a less centralized network, one in which important decisions and plays were not left up to just a few select individuals, is more successful. This theory persists outside the realm of sports psychology; Sparrowe, Liden, and Kraimer (2001) tested similar hypotheses in the workplace, on both individual and group levels. Interestingly, they only found weak support for the hypothesis that centralization in a work group's advice network negatively influences its performance. However, they showed that more central individuals within those groups do better than their peers, creating an interesting potential paradox wherein better-connected individuals may benefit even if their own group does not. This is similar to the theory of multilevel selection in evolutionary psychology: at the individual level, more selfish individuals are selected for over their less selfish counterparts, while at the group level it is better to be less selfish than other groups (Wilson & Sober, 1994).

Another factor that would obviously affect a team's passing behavior is the identity of the five specific players that are on the court at a given time. Both McPherson, Smith-Lovin, and Cook (2016) and Sparrowe, Liden, and Kraimer (2001) note the tendency of nodes to form subgroups within their larger networks based on personal preferences such as homophily among members, and the ability of these subgroups or "cliques" to affect performance of the overall network. Core-periphery analysis, which is often used in business settings, has also been applied to men's basketball (Fonti & Maoret, 2015), finding that stable relationships between dyads where both players are in the core of their networks positively affect team success significantly more than those between dyads involving at least one peripheral member. However, this effect was found only to a certain extent, after which increasing peripheral stability instead had more of a positive consequence on team performance. It could be that relying too heavily on a team's core five players makes it more predictable and thus vulnerable to opposing teams developing effective strategies to block said players, again emphasizing the importance of variation in a team's network structure.

Though still a relatively new field, the use of SNA in sports has great potential in helping teams and coaches come up with new strategies that involve the entire team, not just individual players. Tracking the path of the ball down the court can provide far more information than just simple player statistics and team scoring records. Knowledge of how and why some networks of passes lead to more success than others can be applied to devising new plays to maximize a team's effectiveness. In fact, an exciting future direction for this research to pursue would be factoring in opposing teams into analyses and exploring, for example, if certain types of networks are more effective against various defensive plays. If so, coaches may alter their team's passing behavior based on what type of defensive strategy the other team is likely to use. Another interesting question is whether teams tend to change their network structures between halves or quarters, particularly in the last part of a game when the pressure to score increases. Some networks may work better under low-pressure situations when the goal is more to hold on to the lead and prevent the other team from catching up, while others lead to more opportunities to score quickly and often, as is often necessary in the last half or quarter, particularly when a team is behind.

SNA can also help with modeling games, both for those involved in the sport itself and professionals such as commentators or owners who work at forecasting the outcomes of a game or season. Social networks offers another source of statistics and analysis that could prove handy when predicting results.One can easily track metrics across seasons and even years, to see if they remain relatively consistent within teams. If so, it may be easier to predict more vs. less successful teams based on their metrics. More should also be done on the success of individuals within their networks, as most SNA research so far has looked at the team's wins as a whole. For instance, do players who occupy a more central position in their team's network score more points, or earn more money or awards than their less-central teammates? Are there players who would be more successful if given an opportunity to increase their centrality, maybe by receiving more passes or simply being on the court more? Having this knowledge could inform coaches, owners, and other professionals whether each player is being utilized to their full potential, which could be key in predicting success as well.

Finally, we can ask how these results apply to other small networks, such as in business, the military, and even other species. Do we see the same patterns when the network is cooperating in reaching a less physical and more abstract goal, such as coming up with effective advertising strategies, or plans of attack? Unlike sports, groups of individuals in many other areas such as companies, politics, or social settings are not bound by a specific set of rules. Does this lack of concrete proscriptions on nodes' behaviors affect network stability or predictability? And how do sports networks compare to those in which there is no particular opposing team to win against in single matches, such as when companies compete against each other for customers every day?

Another characteristic mostly unique to sports networks is the rigid assignment of roles to each node. Players very rarely cover more than one or two positions their entire careers. But this strictness is not often seen in other settings. People get promoted or demoted frequently in companies and the military, and many social or community-based networks tend to have fluid or rapidly-changing roles for their members. The greatest contrast may be between sports and animal networks, where no roles are formally assigned. Although also an emergent field, much promising research into network structures in other species is currently being conducted (see

Wey, Blumstein, Shen, & Jordan, 2008; Krause, James, Franks, & Croft, 2014, for overviews), focusing both on sociality between network members, and the role of the individual within its network. Many if not most animal networks are fluid with no specific roles explicitly assigned to each member and no set opponent to compete against, almost the opposite of organized team sports. Comparing the advantages of network structures in these two very different settings would no doubt reveal some illuminating insights.

1.2.2 *In business/economics*

Plenty of studies have been conducted in the corporate world on how interactions between coworkers, or how different leadership styles, can create different network structures within a company and affect its success on many levels. Some focus on the individual scale, where individual employees are the unit of interest. As mentioned in the previous section, according to Sparrowe, Liden, Wayne, and Kraimer (2001), decentralized business networks consisting of colleagues do at least marginally better than centralized ones (because everyone has access to and can help each other equally), similar to several findings in sports (Clemente, Couceiro, Martins, & Mendes,2015; Fewell et al. 2012). Having many competent players who are equally likely to receive and pass on the ball can make it harder for opposing teams to identify and “double-team” a few star players. They also found that those employees occupying central positions in advice networks receive better performance reviews; basketball players in more centralized positions on their teams may score more points or win more awards, perhaps simply because they have increased chances to shoot the ball.

SNA can also be used to determine the nature of the relationships between employees at a given company. Social exchange theory (Emerson, 1972b) bases its predictions on the characteristics of edges between nodes, not the nodes themselves. According to this theory, the nature of our interactions with other people (for instance, whether reciprocity has occurred or not) direct the outcomes of those relationships, such as how likely we are to cooperate again in the future; hence we engage in a constant cost-benefit analysis when dealing with others. Molm (1994) further classifies these exchanges between nodes as *dependence*, in which each actor's outcome depends solely on the actions of its partner(s), or *interdependence*, where all actors involved have an effect their collective outcome. Because in interdependent networks no one actor has direct control over another's fate and one does exercise some control over one's own outcome, these structures are seen as more stable. We can analyze sports networks in this way by strictly focusing on the outcomes of passes and passing patterns rather than the identity of the players themselves, assuming that their common goal is to win their match. Their interactions would fall under *interdependence* because the outcome depends both on their own actions (who to pass to) and others' actions (who is open, who has the ball), and because players are not exchanging goods amongst themselves but rather all working together toward a common end. Interdependent networks foster more cooperation than dependent ones do, which is important for a team setting. However, in larger groups it is more difficult to maintain complete interdependence between all members, especially of only the contributions of a few members are able to achieve the overall goal, thus enabling a "free ride" for some nodes on the periphery. So again, high centralization in a network may in fact be costly. Nonetheless, this could be avoided in sports, individual stats are tracked and thus those players contributing the most to their team's success may still be recognized.

Another property of business networks is that they may be organized or affected by job hierarchies (Cross, Borgatti, & Parker, 2002), whereas in sports all positions treated more or less equally; players are ranked more by ability than position. Some positions might receive the ball more often than others but this is usually due to the traditional way the teams and game rules are set up; otherwise the amount of access each player has to the ball is more likely due to their own skill level rather than what position they play. Also according to Cross, Borgatti, and Parker (2002), business networks often display homophily, wherein employees are more likely to associate with coworkers within the same department or position, but in sports different positions regularly pass to each other depending on where in the play the ball is, and who is in the best position to receive it. This is not to say that homophily cannot occur on the court or field, perhaps in the form of preferential passing to a teammate who has similar seniority, playing style, interests off-field, or even the same race or ethnicity.

Due to their structure, most companies' networks allow nodes to share information by passing on data to several different others at once, even regardless of the formalized divisions and hierarchies within the company (Sparrowe, Liden, Wayne, & Kraimer, 2001; Cross, Borgatti, & Parker, 2002). Sport teams "share" the ball amongst themselves, but as there is only ever one ball in play, each possession describes a strictly linear path between single dyads at a time.

In terms of collection methods, passing data in sports would be akin to the decisional method (Tichy, Tushman, & Fombrun, 1979) because it concretely looks at the behavior of main players in making important decisions, and the outcome of those decisions, similar to identifying "star" players who control the passes and recording whether each possession/game is successful or not. However, it could be an improvement on classical decisional studies because of the clarity of key players and issues (those that get passed to or shoot the most, and the goal of winning the match). This data need not be based on reputation or self-report either, which can make it easier to measure/quantify objectively, and eliminate self-report bias.

Another way sports networks differ from business networks is in how business networks are often constrained by geographic location whereas all members of a sport team are necessarily in the same setting during a game. However, both frequently undergo structural changes: businesses from mergers/acquisitions or restructuring, sports from graduation/trading.

1.2.3 *In social sciences*

As noted above, much of SNA has its roots in the social sciences, including sociology and social psychology, and continues to inspire many studies from these fields. Many social psychology networks track concepts such as communication or support, rather than concrete, quantifiable interaction such as passes between teammates or collaborations between coworkers. They tend to rely on self-report as well. However, these conceptual networks can still tell us much about the behavior of humans in social situations.

Sometimes social psychologists construct agonistic networks involving disagreement or aggressive interactions (Guidetti, Cavazza, & Graziani, 2016), similar to how some business networks chart “hindrance” among colleagues rather than more positive interactions such as advice-giving (Sparrowe, Liden, Wayne, and Kraimer, 2001). Guidetti, Cavazza, & Graziani (2016) in fact looked at two types of networks featuring differences of political opinion between Italian citizens: heterogeneity of opinions in general and actual disagreement between discussion partners. They discovered that network heterogeneity encouraged political participation by increasing citizens’ political interest and knowledge, while disagreement had the opposite effect. This indicates that not all networks need be homogenous in order to increase cooperation, as long as they are not actively agonistic. In fact, heterogeneity between nodes and a more open network structure can increase knowledge flow between members and from outside nodes (Uzzi & Spiro, 2005); among Broadway artists, too little connection between “small world” groups prevented exchange of ideas between said groups, but too much connectivity tended to homogenize the entire network, thus stagnating the creation of novel ideas. This is similar to how Fonti & Maoret (2015) found that networks featuring more interaction between core and peripheral members can increase NBA teams' success through unpredictability to opponents. These agonistic networks can be just as important as networks of support in identifying problematic relationships and how to avoid or improve them. We see this kind of network less often in sports because teams, by their nature, should be cooperative among players (and coaches).

Hidalgo (2016) brings to light an interesting distinction between SNA as used by social scientists versus by natural scientists, such as physicists and mathematicians. Overall, social scientists are more context-specific in terms of constructing their networks, focusing and specific populations or groups of individuals, while natural scientists are more universal, concerning themselves with networks that can be generalized beyond their immediate sample. Furthermore, social scientists tend to make edges based on interpersonal relationships (trust, friendship, etc.), and within a network, the personal identity of the nodes involved matter in terms of determining the nature of ties to other nodes; natural scientists use more quantifiable links based on actual recorded interactions, and these links are usually predicted stochastically through the network structure itself, without taking the characteristics or identities of nodes into account. Sports networks straddle the line between these two types of analysis: they can fall into the natural science camp because passes are impersonal and concrete, with no special definitions for the relationship between teammates (AKA all players pass the ball in the same way); these networks are also not based on social contexts and should be comparable across teams and different sports, or even with other types of networks, like business models. At the same time, they can also be considered under social science since they are strategic games in which the identities and characteristics of individual players do matter, so nodes have just as much influence on the forming of edges as network structure does. Nonetheless, social and natural scientists often overlap in their use of SNA, such as in tracking the spread of phenomena such as diseases (Eubank et al., 2004) and language (Scott, 2012) within a population.

1.3 CONCLUSIONS

Studying the behavior and relationships that occur in groups (human and otherwise) has long interested researchers from fields as diverse as psychology, sociology, biology, chemistry, physics, political science, economics, sports, and more. Social network theory has thus been applied to many different areas and yielded answers to questions from how best to organize a company to which passing strategies are most successful in team sports. Throughout its lengthy history, SNA has undergone several major changes in statistical modeling, structural focuses, and more, but it has remained widely applicable across various sciences. In fact, it most likely owes its long-term relevance to its adaptability and usefulness to researchers interested in all levels of group interaction; and because of these strengths, the theory will likely persist long into the future as a reliable way to quantify and measure even the most complex interactions between different actors.

Chapter 2. EFFECTS OF DEGREE CENTRALIZATION AND TOTAL DEGREE ON COLLEGE WOMEN’S BASKETBALL TEAMS’ SUCCESS

2.1 INTRODUCTION

The growing field of social network analysis (SNA) focuses on how social interactions between individuals or small groups affects their larger populations, or vice versa (Jamali & Abolhassani, 2006; Edwards, 2010). It allows for social behaviors between pairwise and more complex structures to be quantified, and is applicable to a diverse set of disciplines. SNA is useful for both analyzing the overall ties between individuals or groups in a given population, as with simpler undirected networks, or to track specific behaviors targeted toward others in a directional network. In the latter case it is necessary identify and distinguish between the initiator and the recipient of the behavior. Directed networks are also helpful in tracking the flow of information or disease within a population.

Regardless of the exact type of network, the analysis works best in closed systems in which all individuals (or nodes) and their interactions (or edges) of interest are known to us. This eliminates the possibility of observer bias – that we are only reporting the interactions we see, while there could be much more occurring out of sight. Social sports are an excellent candidate for SNA because we may categorize each game or match as a closed population made up of known players, and if we are recording the passing of a ball between those players as edges in our network, it is usually fairly easy to track every pass during a game, as well as its direction. Some work has already been done in the field of sports psychology using social networks, including soccer (Clemente, Couceiro, Martins, & Mendes, 2015; Kooij et al., 2009) and basketball (Bourbousson, Poizat, Saury, & Seve, 2010; Fewell et al. 2012; Vaz de Melo, Almeida, & Loureiro, 2008). Clemente, Couceiro, Martins, & Mendes (2015) found that among the First Portuguese Soccer League, there was no single “star” player who monopolized the most central position of the team’s network, and that all players decreased their number of connections in the second half, suggesting a lower level of participation for the entire team. Similarly, Fewell et al. (2012) measured the degree centrality, clustering (whether there were any “subgroups” in a team’s overall network), entropy (a measure of the unpredictability in a team’s passes), flow centrality (a measure of a player’s importance within the network), and uphill/downhill flux (whether teams consistently passed to their best shooters) among the 2010 NBA play-offs’ first round as proxies of offensive strategies, and their effects on team success. They also discovered there was often no star player among these teams, and that those teams with higher entropy (aka more unpredictable passing patterns) performed better. Grund (2012) found similar results with men’s soccer teams; lower degree centralization was associated with better performance.

Taken together, these studies imply that having a less centralized network, one in which important decisions and plays were not left up to just a few select individuals, is more successful. This theory persists outside the realm of sports psychology; Sparrowe, Liden, Wayne, and Kraimer (2001) tested similar hypotheses in the workplace, on both individual and group levels. Interestingly, they only found weak support for the hypothesis that centralization in a work group’s advice network negatively influences its performance. However, they showed that more central individuals within those groups do better than their peers, creating an interesting potential paradox wherein better-connected individuals may benefit even if their own group does not. Similarly, Brass (1981) discovered that employees’ centrality in a company’s work flow network had an indirect effect on their performance, while Baldwin, Bedell, and Johnson (1997) reported that M.B.A. students with higher centrality within their class groups earned higher grades. While Fewell et al. (2002) looked at the centralization of basketball teams as a whole in relation to their overall success, she did not measure whether individuals occupying more central positions in the network (thus being more "selfish" with their passing behavior) were more successful in scoring points than their teammates.

Sparrowe et al. (2001) also examined how the density of a network affected its success. Though they predicted that a denser advice (affiliative) network would increase group success, this was not supported by their results. However, they found that density in hindrance (agonistic) networks did negatively impact success, showing that network density can still have an effect on performance. Team sports are necessarily cooperative networks which are more similar to Sparrowe et al.'s advice rather than hindrance networks, but perhaps the different atmosphere of an athletic competition compared to the workplace is more conducive to significant positive results in affiliative network density in relation to success. According to Molm (1994), mutual interdependence between group members encourages more cooperation and stability, which are key to succeed at team sports. Collectivism, small group size, high identifiability, and low shared vs. personal responsibility also foster cooperation (Wagner, 1995). All these factors are characteristic of basketball teams: players are expected to put the team's needs and goals above their own, there are only five active players on the court at any time, each athlete can be individually identified easily, and everyone's personal contribution is framed as important to achieve success. Increasing the density of a network can make these conditions even more salient. Additionally, teams that engage in more passes and ball-sharing between teammates may score more because doing so makes it harder for the defense to follow the ball's path to the basket, and prevents them from teaming up against any single player. Therefore we expect that teams with higher network density, as measured by the sum of all players' degrees per game, perform better than those with lower density.

The first half of our study focuses on comparing metrics between pairs of opposing teams per game, specifically the National Collegiate Athletic Association (NCAA) women's basketball teams during the 2015 championships. We have combined the findings above into a new hypothesis: that less centralized teams do indeed score more points than more centralized ones, but players who receive the most passes on those teams score higher or make more assists (both reflections of personal success) than those receiving less, possibly because they simply allow their teammates less opportunities to attempt shots in the first place. They may also receive more passes because as the best shooters on the team, other players are more likely to pass to them. In either case, we would expect those players with higher centrality in their networks to score more points or assists than their teammates.

In regards to team centrality, because previous studies indicate that networks with low overall degree centrality perform better, we predict that, when split into in- and out-degree measures, teams with lower centrality in both will also win more games. Further, if successful teams really are sharing the ball evenly in terms of both passing and receiving, we should see in- and out-degree centralizations that are not only low, but fairly similar to each other. Thus, when we compare the rate of a team's in-degree centralization to out-degree centralization, this ratio should not be significantly different from 1. That is, not only do we expect that more successful teams will distribute passes more evenly as a whole, they will not differ significantly in the amount of passes each player receives compared to the number of teammates a player passes to. Again, both these strategies should lead to less predictable ball movement and prevent opponents from easily identifying star passers or receivers to "double-team".

Another factor that could affect a team's passing behavior is the identity of the five specific players that are on the court at a given time. Both McPherson, Smith-Lovin, and Cook (2016) and Sparrowe, Liden, Wayne, and Kraimer (2001) note the tendency of nodes to form subgroups within their larger networks based on personal preferences such as homophily among members, and the ability of these subgroups or "cliques" to affect performance of the overall network. Not only may individual preferences for basketball teammates vary as stated above, these different "squads" contain players with varying levels of skill in different offensive strategies and thus may be used at differing rates – more skillful combinations of five players (the so-called "A team") most likely see more play time than other combinations, thus creating an opportunity for these lesser-used squads to skew our data. Therefore, we've broken our analyses up to both reflect the network metrics of the entire team as well as its "A squad" as a check against this bias.

Coaches and teams that utilize high numbers of substitutions between these different squads also run the risk of introducing unwanted instability into their game strategy. If a particular squad of 5 teammates have not practiced or played together very often, they may be less familiar with each other's skills and tactics, which could impede ball movement and allow the rival team more opportunities to intercept the ball. Social network researchers have studied the evolving dynamics of members entering and leaving a network (McPherson, Popielarz, & Drobnic, 1992) and how the strength of both internal and external ties influence network turnover rates (Wellman, Wong, Tindall, & Nazer, 1997; Moynihan & Pandey, 2008), but virtually no studies have yet looked into how turnover rates and different compositions of nonvoluntary subnetworks (i.e. players exert less choice over play time and squad combination than coaches do) affect overall network success. While unpredictability generated from decentralized passing patterns between teammates may help keep the opponent guessing, the unpredictability from large numbers of substitutions may serve to confuse players, as they are constantly forced to readjust to a changing roster of teammates. Further, this could decrease the playing time of a team's best players while increasing the likelihood that a "rookie" will make a mistake on the court. If this is the case, teams with lower amounts of substitutions and unique squad makeups should perform better than those with higher numbers.

The second half of our study focuses on how network metrics influence success within a single team. This helps eliminate individual differences when comparing across different teams. To this end, we also analyzed the 2006-2007 season of the University of Washington (UW) Huskies women's team. Doing so allowed us to hold such factors such as coaching staff and team roster constant, thus reducing variability between games that could otherwise confound our results. This allows clearer analyses on which metrics lead to success vs. failure.

In summary, we predicted that teams with lower centralization scores are more successful while individual players with higher centrality within their teams (those receiving or making the most passes) will score more points per game. We also predicted that, if successful teams do exhibit lower centralization, the ratio of their in-degree to their out-degree centralization scores will not vary significantly from 1. Finally, we hypothesized that winning teams will have higher total degrees and less unique squads than losing teams. We expected to see these same results from the A Squads, controlled for individuals with very little play time, and in our within-team analysis of the UW team.

2.1 METHODS

For the between-teams half of our study, we coded the 2015 NCAA women's basketball championships tournament (March Madness), from the Sweet 16 rounds to the final (15 games total). Each game was coded twice, for both teams playing, for a total of 30 networks. For the within-team portion, we included the 12 games of the UW women’s 2006-2007 season, from 12/21/2006 to 2/10/2007 (Table 2.1). Players were coded as nodes and passes between them as weighted, directional edges. We also recorded the start of each possession (Inbound, Rebound, or Turnover) and its outcome (Good 2, Good 3, Miss 2, Miss 3, Offensive Foul, Defensive Foul, Turnover Out of Bounds, Turnover Steal, or Turnover Sloppy); however, these were excluded from our analyses because we wanted to confine our metrics to players and did not wish for these “extra” nodes to skew our results.

Each player was coded as an individual node; for any five groups of teammates that appear on the court together during the March Madness games, we coded that unique combination as a “squad”, and recorded the total number of squads in each game per team. In order for our centralization measures to not be skewed by players or combinations of players that are rarely on the court, we ran separate analyses on the most commonly used squad combination (“A squads”) for our focal team to ensure we were confining our analyses to players who have substantial amounts of play time. Our measure of network centralization was based on Borgatti, Everett, & Freeman ‘s formula for UCINET (1992). Weighted and directed degree centralizations for each team were calculated in R by finding the in-, out-, and total degrees for each node, then dividing it by the value of the node with the maximum score to standardize it against bias towards teams that may have a higher number of passes simply because they go through more possessions than other teams. From there we subtracted each individual degree from the max degree value, as a measure of how much they vary from the potential "star player" on the team with the most passes, and then that value is divided by the total number of nodes minus 1, to control for networks of different sizes; finally all weighted degrees for each team are summed together. The result is a number between 0 and 1, with 0 indicating that everyone on the team is receiving the exact same number of passes, and 1 indicating that there is one player receiving all the edges from everyone else and no one else is connected, i.e. a wheel network. It should be noted that total centralization was not simply a sum of in-degree and out-degree centralizations: each type of centralization was calculated separately as comparisons of variance among players for that particular degree type. For example, a team in which four players all only passed to the fifth but never to each other would receive an in-degree centralization of 1 since only Player 5 has any in-degree, an out-degree centralization of a little over 0 because all players other than Player 5 share the same out-degree, and a total centralization of 1 because Player 5 is the only player who is connected in some way to their teammates.

We also calculated the ratio of each team's in-/out-degree centralizations as a measure of variation in receiving passes compared to passing among players, and conducted single-sample *t*-tests against a null ratio of 1 (equal proportions of players receiving and passing on the ball) for both the winning and losing teams per round.

Total degrees for each game was calculated by summing up each player on the team's overall degree. To control for the possibility that higher totals may be reflecting greater numbers of possessions rather than amount of passing per possession, we also recorded the total number of possessions for each team in each game and compared those, as well as running a correlation between total degree and number of possessions. Correlations were also performed between the winning and losing team of each game in terms of their degree centralizations and total degrees, to test for match effects. Team ID was also run through the model to control for some teams being represented more than others, through having made it further through the tournament.

For each game, we obtained the identity of the players who scored the most points or made the most assists (passed to a teammate who then made a successful shot) per team from www.NCAA.com, and recorded whether they also possessed the highest degree for that game. Rating percentage index (RPI) scores for the Huskies' opponents were also recorded from www.NCAA.com. This index is ranking of teams every season based on their win percentage as well as that of their opponents and opponents' opponents. The RPI of each opposing team was included as a well-known predictor of success that we can use against our own metrics. We ran all relevant factors through the GLM model-fitting function in R, as well as separate comparisons between winning and losing teams in R and Microsoft Excel. Effect sizes were calculated from odds ratios according to Chinn's (2000) methodology.

2.2 RESULTS

2.2.1 *Between-team analysis*

Figure 2.1 displays an example of the team network graphs, from the NCAA championship match between University of Connecticut and Notre Dame. When taking the whole team into account, individual centrality was the most significant predictor of success (p = 0.01), Table 2.2. Its odds ratio of 0.09 signifies that for every unit increase in individual centrality, the odds of winning decrease by 0.91 or 91%. In-/out-degree ratio (p = 0.14), total degree centralization (p = 0.79) and total degree (p = 0.95) did not significantly predict success between winning and losing teams; these results are echoed in our stepwise regression of the all variables (Table 2.9). However, in-degree centralization's and ratio's log odds (2.40 and 1.70, respectively) indicate that they are strong positive success predictors for all teams. This is supported by their relatively large effect sizes (0.48 and 0.29, respectively), indicating in-degree centralization's and ratio's strong effects on success. In terms of model-fitting, individual centrality (p = 0.009) and ratio (p = 0.09) significantly accounted for the variability of our model, Table 2.3. There was no significant effect of team on success, Table 2.4.

Winning teams were not more likely than losing teams to possess higher in- ($t(14)$ = 2.14, p = 0.47), out- ($t(14)$ = 2.14, p = 0.67), and total ($t(14)$ = 3.18, p = 0.78) degree centrality overall. However, in general the various teams possessed higher in-degree centralization (M = 0.75, SD = 0.12) than out-degree centralization (M = 0.69, SD = 0.10, $t(14)$ = 2.05, p = 0.04). It should be noted that our power was low for all three analyses (0.11, 0.07, and 0.05, respectively). There was also no significant correlation between opposing teams' total centralization scores (r = 0.05, p = 0.85).

There was no significant correlation between the in-degree (r = 0.03, p = 0.92), out-degree (r = 0.11, p = 0.70), or total degree (r = 0.03, p = 0.92) centralizations of the two teams in each game. We did find a significant negative correlation between the total degrees of teams per round (r = -0.56, p = 0.03). When winning teams had higher degree totals, the respective losing teams possessed lower totals.

For each team's top squad, none of the variables predicted success significantly (Table 2.5), nor did they significantly account for any of the variability in our model (Table 2.6). Further analysis revealed was no difference in in- ($t(14)$ = 2.14, p = 0.60), out- ($t(14)$ = 2.14, p = 0.69), and total ($t(14)$ = 3.18, p = 0.52) degree centralizations for winning vs. losing teams over the entire last half of the tournament. Again, power was low (0.08, 0.07, and 0.09, respectively). However, when compared on the level of individual rounds, there is a trend towards significance for winning teams (M = 0.54, SD = 0.06) to show higher degree centralization than the teams that lost to them (M = 0.47, SD = 0.1) in the Sweet 16 round ($t(7)$ = 2.36, p = .075).

We also found a strong correlation (r = 0.97, p = 0.03), in our A Squads, for in-degree centralization to increase as the tournament progressed, with teams in the later rounds steadily displaying higher average centralizations than in the earlier ones (Figure 2.2). Individual analyses revealed that, compared to the championship round, both Sweet 16 ($t(7)$ = -4.46, p = 0.01) and Elite 8 ($t(3)$ = -2.36, p = 0.046) had significantly lower in-degree centralizations (Figure 2.3).

The number of unique squads per game on each team did not significantly predict success in our model (p = 0.18), but it did trend toward being lower on winning (M = 10.33, SD = 3.83) vs. losing teams (M = 12.07, SD = 3.47), $t(14)$ = -2.02, p = 0.06. There was no significant correlation between the number of squads and a team's final score, r = -0.23, p = 0.23, power = 0.82.

Winning teams were significantly more likely to have an in-/out-degree ratio greater than 1 (M = 1.17, SD = 0.26, $t(14)$ = 2.57, p = 0.02, Figure 2.4); however, no significant difference was found for losing teams (M = 1.05, SD = 0.22, $t(14)$ = .85, p = 0.41).

Finally, winning teams show fewer instances of the player with highest in-, out-, or total degree centralization also having highest assist (N = 6) than losing teams (N = 12), though this result was not significant ($\chi^2(1)$ = 2.00, p = 0.16). Winning teams did not have higher total degrees (M = 489.87, SD = 93.38) than losing teams (M = 480, SD = 131.68), $t(15)$ = 0.19, p = 0.85, nor did they go through more possessions (M = 85.73, SD = 4.83) than losing teams (M = 90.2, SD = 17.59), $t(14)$ = -1.13, p = 0.28.

2.2.2 *Within-team analysis*

In analyzing the UW women's 2006-2007 season, our model showed that out-degree centralization most strongly predicted success negatively (p = 0.02), with an odds ratio of 0.001,

Table 7. It also explained most of the variance within our model ($p < .001$), Table 2.8. Total degree centralization ($p = 0.05$) and the opposing team's RPI ($p = 0.07$) both trended toward being positively associated with success; however, degree centralization explained more variability ($p = 0.01$) than RPI did ($p = 0.03$). Individual centrality ($p = 0.90$), home vs. away ($p = 0.46$) and total degree ($p = 0.95$) did not significantly predict success. Though none of our variables returned significant *p*-values in a stepwise regression (Table 2.11), their odds ratios nevertheless support our GLM results, with out-degree centralization (0.003) and total centralization (14.30) most strongly positively and negatively predicting success, respectively. They also possess the largest effect sizes of all variables (-3.06 and 1.47, respectively). Notably, both metrics were better predictors of success than RPI in all models.

When analyzed separately, we found that there was no difference in total degree centralizations between games won ($M = 0.65$, $SD = 0.07$) and games lost ($M = 0.68$, $SD = 0.04$), $t(10) = -0.83$, $p = 0.43$, or in in-degree centralizations between games won ($M = 0.63$, $SD = 0.13$) and games lost ($M = 0.62$, $SD = 0.10$), $t(10) = 0.14$, $p = 0.90$. Power was 0.10 and 0.05, respectively. However, the Huskies did possess a lower out-degree centralization on average when they won ($M = 0.73$, $SD = 0.05$) compared to when they lost ($M = 0.82$, $SD = 0.06$), $t(10) = 2.02$, $p = 0.04$, Figure 2.5. There was no significant correlation between total degree centralization and score ($r = 0.23$, $p = 0.47$, power = 0.77).

For ratios, in winning games the Huskies trended toward a lower ratio ($M = 0.86$, $SD = 0.18$) than 1, $t(7) = -2.22$, $p = 0.06$. However, when they lost the average ratio per game was significantly lower than 1 ($M = 0.74$, $SD = 0.06$), $t(3) = -8.61$, $p = 0.003$, Figure 2.6.

In 4/8 winning games and 1/4 losing games, the player with the highest in-degree also scored the most points. In contrast, none of the players in either successful or unsuccessful games had both the highest out-degree and points scored. For total degree, in only 1/8 won matches and no lost matches did the same player possess both the highest degree and most points.

For total degree per game, UW had significantly higher totals when successful (M = 443.75, SD = 98.30) than unsuccessful (M = 340, SD = 53.14), $t(10) = 2.37$, $p = 0.04$, Figure 2.7. There was no significant difference in total number of possessions between matches won (M = 97.75, SD = 3.01) and lost (M = 98, SD = 0.82), $t(10)$, -0.22, $p = 0.83$, power = 0.05, nor was there a significant correlation between degree and number of possessions, $r(11) = 0.02$, $p = 0.94$, power = 0.08. A post-hoc correlation between total team degree and final score per game also returned nonsignificant results, $r(11) = 0.208$, $p = 0.52$, power = 0.77, further weakening the possibility that simply going through a higher number of possessions in a game leads to success.

The total number of unique squads failed to significantly predict success ($p = 0.44$), but did trend slightly towards being higher in losing teams (M = 24.25, SD = 4.5) than in winning teams (M = 19, SD = 3.55), $t = 2.04$, $p = 0.097$. There was no correlation between number of squads and total score per game ($r = -0.1$, $p = 0.77$, power = 0.18). As the season progressed though, the Huskies did significantly increase the amount of squads substituted in each game (r = 0.79, $p = 0.002$), Figure 2.8.

Correlation matrices are included for between-teams (Table 2.10) and Husky women's (Table 2.12) analyses.

2.3 DISCUSSION

Although explaining a significant amount of variability in our model, we found no evidence that centralization predicted success; there was no difference between winning and losing teams in their in-, out-, or total degree centralizations. This is contrary to findings which indicate sport teams with lower centralizations, aka more evenly distributed passes, score higher (Grund, 2012; Pena & Touchette, 2012) but do support Fewell et al.'s (2012) conclusion that there was no significant difference between winning and losing teams' degree centralizations in the 2010 NBA play-offs; lower centralization therefore leads to success in soccer but not basketball, probably due to inherent differences between the two types of sports and their strategies. For example, soccer involves 11 players on the field at a time and is reflected by a larger network in which even distribution of passes among more nodes is easier to achieve than in a smaller network of 5 players. Being less dependent on physical height can make it less challenging for players to cover a variety of positions in soccer compared to basketball. Our results may be also due to the fact that, at the college level, there is such a difference in skill level among teammates that it is actually a better strategy to concentrate passes and plays on the strongest players. The increasing level of centralization among all teams as the tournament progresses might reflect teams relying more and more on their strongest players as the competition increases as well. Several studies have found that decentralized networks only have an advantage over centralized ones when performing complex tasks; for more simple tasks centralized networks actually perform better (Bavelas, 1950; Bavelas & Barrett, 1951; Leavitt, 1951). If college basketball is indeed less difficult to play than professional basketball, a centralized network would be a better strategy in this case. Or perhaps it is simply more difficult for teammates and coaches to prevent star players from monopolizing the ball in college, due to this discrepancy in skill levels.

However, as the positive significant result for total centralization in our UW model reflects, a higher level of centralization within a single team can contribute to success when applied across multiple games. In fact, total degree centralization was a stronger predictor of the

Huskies' success than their opponents' RPI ranking; thus network metrics can better forecast a team's performance than methods that are currently in popular use. Interestingly, we also see a strong negative relationship between out-degree centralization and success, meaning that when the Huskies distributed their outgoing passes more evenly among all players, they performed better. Taken together, we can conclude that the most effective strategy is to ensure all players are passing the ball at the same rate, followed by endeavoring that those passes are received predominantly by the most skilled players, thus reflecting the lower out-degree and higher total degree centralizations of our results. Though these variables are moderately correlated with each other (r = 0.67, p = 0.02, Table 2.12), they still exert opposite effects on success despite their positive correlation. Furthermore, total centralization still contributes significantly to our model even when accounting for out-degree centralization (p = 0.02, Table 2.8).

Generally, in-/out-degree ratio was not enough to predict success either between or within teams. However, when comparing the ratios separately in our between-team analysis, both winners and losers had ratios higher than one. Within the UW team, the Huskies were much more likely to have a lower ratio when they lost. As a whole, our results show that there is more variation among teammates in who receives the ball compared to who passes, particularly in successful teams. Again, one explanation is that this reflects the practice of everyone passing to their team's best player(s), who then makes the decision on who best to direct their own passes to. Because winning teams were found to possess a higher ratio, the more successful strategy involves greater variation in who receives passes among the team than whom each player passes to. Winning teams concentrate their passes toward one or two "stars" who then presumably make the call on who else to pass to, given changing circumstances as the game progresses. The outcomes of each possession (fouls, successful and unsuccessful shots, and turnovers) were also excluded from this analysis, so it is possible that the star players receiving the most passes on each team were also the ones attempting the most shots, something that would not be reflected in our current results.

Interestingly, we found that on winning teams, it is more likely that the players making the most assists are not at the center of the network. Though this result was not significant between teams when compared separately, it does significantly predict success when included as part of our overall model. Within teams, we see much stronger evidence on an individual level of comparison, in the same direction, particularly with out-degree – in none of the Huskies' 2006-2007 games did the highest scorer also possess the highest out-degree. These results indicate that successful teams do not rely on the same couple players performing all roles – that is, controlling passing behavior, shooting, and assisting. This would prevent the opposing team from easily double-teaming or otherwise incapacitating the most competent athlete, and creates more opportunities for all teammates to contribute to the play in some way. Even if a team's best shooter draws extra defense, the other players can still at least help the ball reach the shooter rather than losing it to a turnover.

Though we did not find a significant increase in total team degree for successful teams among the NCAA championships, we did see this during the UW women's 2006-2007 season. Simply increasing the number of passes per possession is not on its own enough to guarantee a win, but it is one factor that contributes noticeably. At least among the Huskies, other factors such as coaching and team composition held constant, the team performed better when number of passes, but not number of possessions, increased. Having more pathways for the ball to travel to the basket involving more players can force the defending team to spread its resources more thinly as they must be cognizant of more than just one or two players, and having alternate passing routes again helps decrease the chances of a turnover if one route is blocked. The negative correlation between winning and losing teams' total degrees in the NCAA data could also reflect this pattern of higher numbers of passes leading to success, or it could reflect that as one team monopolizes the play time to perform more passes, it effectively limits its opponent's access to the ball and decreases the other team's chances of scoring. This of course is a tried-and-true tactic, and it is encouraging that network metrics are able to correctly reflect it.

In both our within and between team analyses, we saw a tendency for success to be associated with fewer squad substitutions. Despite nonsignificant results, the similar trend in our two datasets points to the possibility that increased unpredictability among teammates brought about by frequent changes in subgroup composition can be detrimental to gameplay. This is not surprising, given previous findings that stronger ties between nodes lead to more stable networks (Wellman, Wong, Tindall, & Nazer, 1997; Moynihan & Pandey, 2008); as a cooperative network, thicker edges between teammates used to playing with each other often would help them gauge each other's skill levels and passing habits, thereby making it easier to form effective and lasting strategies. Too much substitution can break up these stable squads and force players to constantly re-evaluate and anticipate novel playing conditions and partners. On the other hand, UW's increasing substitutions as their season progressed may reflect the need to cycle through more peripheral players as the core squad became more fatigued or even injured after playing through multiple games. If so, substitution should not be relied upon to increase success but as a last resort to replace struggling teammates.

Understanding how social metrics and passing behavior affect a team's success on the court can help coaches and players improve and/or adjust their strategies based on the strategies of opposing team, as well as any vulnerabilities within their own teams. Although previous studies have suggested that less centralization is associated with higher success, this may only be true on a professional level where each player on a team possess similar amounts of skill. In college games, where disproportionate skill levels may occur within a team, the better strategy may just be to allow highly skilled players more access to the ball, perhaps so they can make more efficient decisions on who to pass to and how to direct the ball's path to the basket. Since we found that players with higher degrees were not the ones assisting the most shots, it's possible that the most effective teams have a player who specializes in directing passes so they can most easily reach the best assister(s), who can then set up a successful shot. This would serve to take the pressure off any one player to perform more than one role during the game, and prevent the opposing team from double-teaming star players. This is also supported by our analysis of in-/out-degree ratios. For coaches and athletes, such information provided by an SNA approach to the sport could be invaluable in forming strong, effective strategies that provide them with an edge over their opponents. In fact, as shown by our between-team results, these metrics can be even better predictors of success than traditional statistics currently in use, further reason for their adoption by the industry. Comparing sports networks' metrics to other types of networks such as those in businesses may also prove useful in discovering whether differences in structure and end goals have an effect on the success of certain strategies. For example, do the strict rules of organized sports allow networks to benefit more from unpredictability in edge formation than less formally-organized networks of coworkers in a company? This type of analysis could lead us to a greater understanding of how networks function under a wide range of circumstances and settings, and be broadly applicable across many areas of study.

2.4 TABLES AND FIGURES

Game Date	Team	Location	Score	Outcome	Opponent's RPI
12/21/2006	U of O	Away	60-56	WIN	99
12/23/2006	OSU	Away	71-67	WIN	218
12/29/2006	USC	Home	76-60	WIN	67
12/31/2006	UCLA	Home	72-67	WIN	103
1/4/2007	AZ	Away	83-66	WIN	124
1/6/2007	ASU	Away	83-88	LOSE	15
1/20/2007	WSU	Home	86-64	WIN	278
1/28/2007	USC	Away	70-50	LOSE	67
2/1/2007	ASU	Home	64-75	LOSE	15
2/3/2007	AZ	Away	83-66	WIN	124
2/8/2007	CAL	Away	79-76	WIN	28
2/10/2007	Stanford	Away	54-80	LOSE	10

Table 2.1: 2006-2007 UW women's basketball games

	Estimate	*p*-value	Odds ratio	Effect Size
Intercept	-1.6387616	0.7265	0.1942204	-0.9035
Individual Centrality	-2.3692248 *	0.0136*	0.09355322	-1.30622
Total Centralization	-1.2512900	0.7923	0.2861354	-0.68987
Ratio	3.4322717	0.1418	30.94686	1.892311
Total Degree	-0.0002637	0.9465	0.9997363	-1.5E-04

Table 2.2: Logistic regression of 2015 NCAA games

	Df	Deviance Resid.	Df	Resid. Deviance	*p*-value
Null			29	41.589	
Intercept	1	6.7939	28	34.795	0.009147**
Individual Centrality	1	0.1303	27	34.665	0.718100
Total Centralization	1	2.8444	26	31.820	0.091694.
Ratio	1	0.0045	25	31.816	0.946406

Table 2.3: Model-fitting of 2015 NCAA games

	Estimate	*p*-value
Team ID Baylor	1.787548	0.119
Team ID Dayton	.637630	0.248
Team ID Duke	.439001	0.285
Team ID FSU	.028999	0.168
Team ID Gonzaga	.625905	0.596
Team ID Iowa	.778320	0.215
Team ID Louisville	.592373	0.604
Team ID Maryland	.660591	0.169
Team ID Notre Dame	.161280	0.100
Team ID South Carolina	.445936	0.217
Team ID Stanford	.828543	0.218
Team ID Tennessee	.376054	0.218
Team ID Texas	.627616	0.620
Team ID UConn	.508896	0.142
Team ID UNC	.487398	0.240

Table 2.4: Team effects for NCAA data

	Estimate	*p*-value	Odds ratio	Effect Size
Intercept	1.708400	0.559	5.520122	0.94189
Individual Centrality	0.454584	0.560	1.575518	0.250625
Total Centralization	-3.522426	0.392	0.02952771	-1.94202
Ratio	0.448855	0.682	1.566517	0.247467
Total Degree	-0.003662	0.499	0.9963447	-0.00202

Table 2.5: Logistic regression of 2015 NCAA games A squads

	Df	Deviance Resid.	Df	Resid. Deviance	*p*-value
Null			29	41.589	
Intercept	1	0.55758	28	41.031	0.4552
Individual Centrality	1	0.41899	27	40.612	0.5174
Total Centralization	1	0.15534	26	40.457	0.6935
Ratio	1	0.47737	25	39.980	0.4896

Table 2.6: Model-fitting of 2015 NCAA games A squads

	Estimate	*p*-value	Odds ratio	Effect size
Intercept	2.317	0.1884	10.14519	1.28
Out-Degree Centralization	-6.684	0.0227 *	0.001250765	-3.69
RPI	2.942e-03	0.0517	1.002946	1.62E-03
Total Centralization	4.726	0.0663	112.8433	2.61
Individual Centrality	4.057e-02	0.9038	1.041404	2.24E-02
Total Degree	6.714e-05	0.9542	1.000067	3.70E-05

Table 2.7: Logistic regression of 2006-2007 University of Washington games

	Df	Deviance Resid.	Df	Resid. Deviance	*p*-value
Null			11	2.66667	
In-Degree Centralization	1	1.27048	10	1.39618	0.0001244***
Out-Degree Centralization	1	0.40082	9	0.99536	0.0311351 *
Total Centralization	1	0.47608	8	0.51928	0.0188251 *
Individual Centrality	1	0.00128	7	0.51800	0.9029624
Total degree	1	0.00031	6	0.51769	0.9523057

Table 2.8: Model-fitting of 2006-2007 University of Washington games

	Estimate	*p*-value	Odds ratio	Effect Size
Intercept	-0.0958171	0.9331	0.9086302	-0.05283
In-Degree Centralization	0.8734629	0.2125	2.3951909	0.481565
Out-Degree Centralization	-0.1048644	0.8974	0.9004466	-0.05781
Total Centralization	-0.4604928	0.6297	0.6309726	-0.25388
Ratio	0.5301865	0.1831	1.6992492	0.292307
Total Degree	0.0001188	0.9021	1.0001188	6.55E-05
Centrality	-0.5077969	0.0134 *	0.6018200	-0.27996

Table 2.9: Stepwise regression of between-teams analysis

Row	Column	Correlation
Ratio	Total Degree	0.11000
Ratio	Centrality	0.00250
Total Degree	Centrality	-0.01400
Ratio	Out-Degree Centralization	0.29000
Total Degree	Out-Degree Centralization	-0.43000*
Centrality	Out-Degree Centralization	0.36000*
Ratio	In-Degree Centralization	0.00083
Total Degree	In-Degree Centralization	-0.23000
Centrality	In-Degree Centralization	0.18000
Out-Degree Centralization	In-Degree Centralization	0.25000
Ratio	Total Centralization	-0.04700
Total Degree	Total Centralization	-0.24000
Centrality	Total Centralization	-0.09900
Out-Degree Centralization	Total Centralization	-0.06800
In-Degree Centralization	Total Centralization	0.19000

Table 2.10: Between-teams correlation matrix

	Estimate	*p*-value	Odds ratio	Effect size
Intercept	1.6769782	0.3150	5.349366654	0.924567
In-Degree Centralization	1.5672024	0.3150	4.793219895	0.864044
Out-Degree Centralization	-5.5564774	0.0501 .	0.003862358	-3.06345
Total Centralization	2.6603682	0.3194	14.301554112	1.466738
RPI	0.0037124	0.0301 *	1.003719319	0.002047
Centrality	-0.0546961	0.8639	0.946772787	-0.03016
Total Degree	0.0003201	0.7721	1.000320148	0.000176

Table 2.11: Stepwise regression of Husky women's analysis

Row	Column	Correlation
Out-Degree Centralization	Total Centralization	0.670*
Out-Degree Centralization	Total Degree	-0.510
Total Centralization	Total Degree	-0.240
Out-Degree Centralization	In-Degree Centralization	0.210
Total Centralization	In-Degree Centralization	0.570
Total Degree	In-Degree Centralization	-0.220
Out-Degree Centralization	Ratio	-0.290
Total Centralization	Ratio	0.230
Total Degree	Ratio	0.026
In-Degree Centralization	Ratio	0.870***

Table 2.12: Husky women's correlation matrix

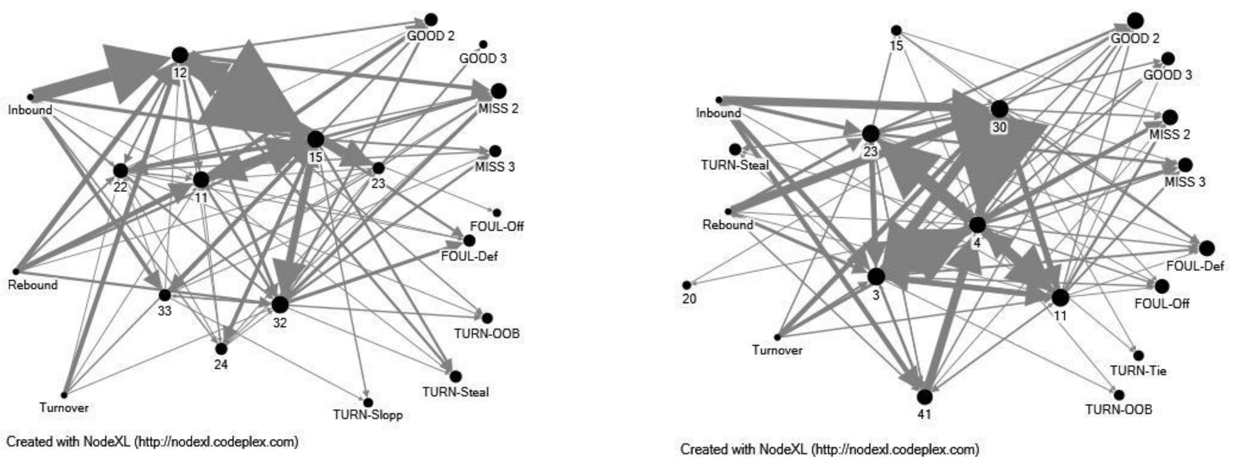

Figure 2.1: Graphs for Notre Dame (left) and UConn (right) in the 2015-2016 NCAA championship finals. Node size is based on in-degree, and edge width on edge weight.

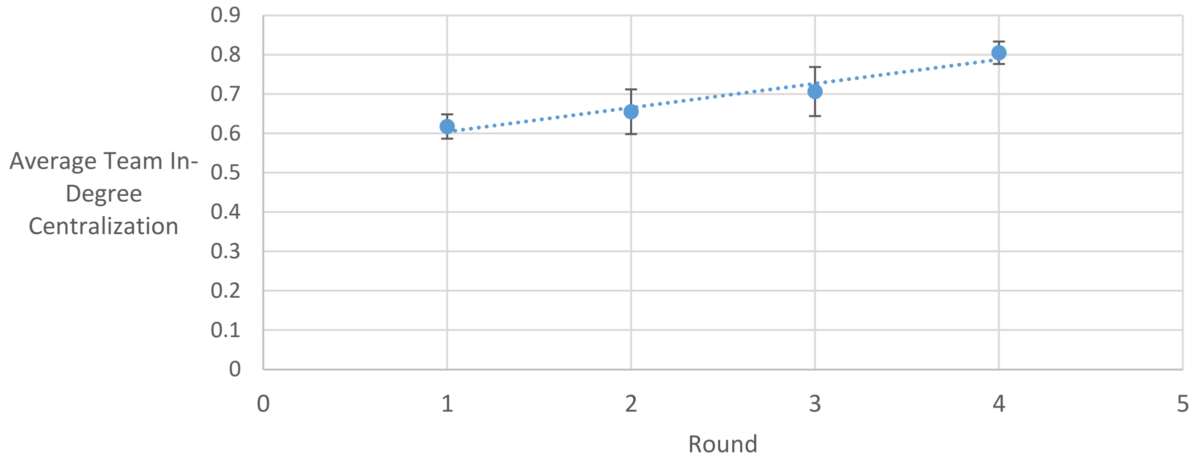

Figure 2.2: Average team in-degree centralization throughout progression of tournament. Error bars represent +/-1 SE

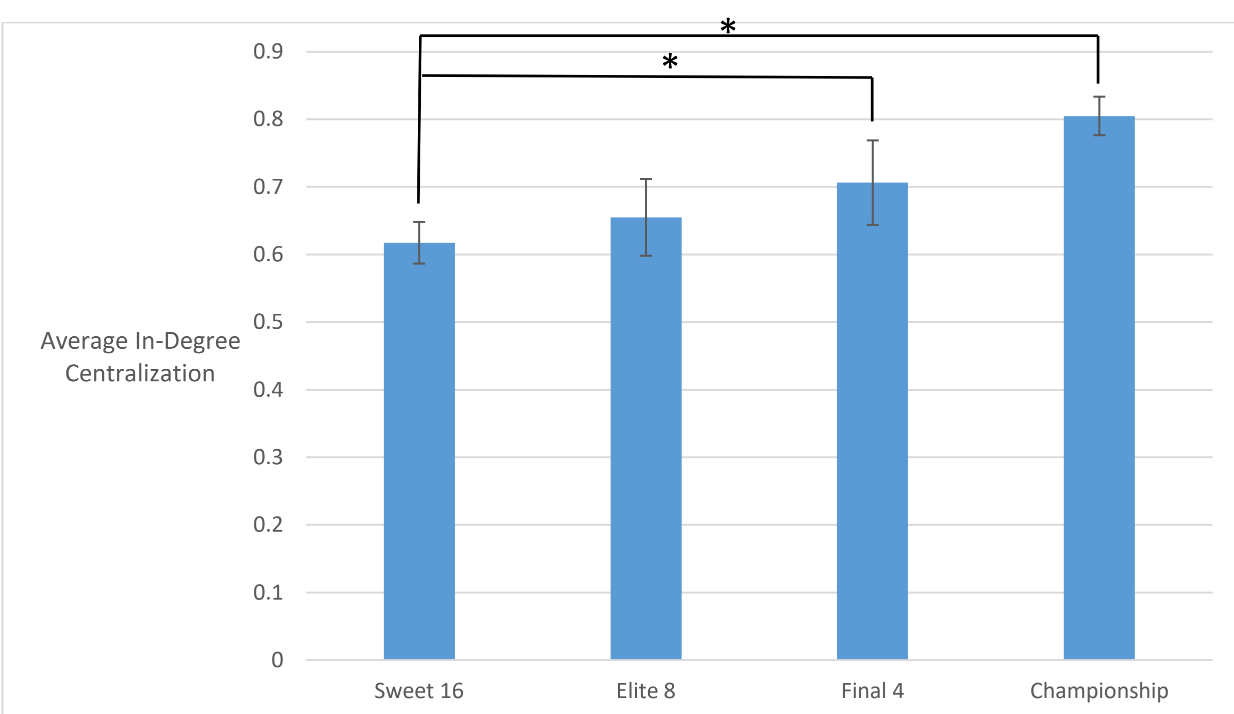

Figure 2.3: Average in-degree centralizations per round. Error bars represent +/-1 SE

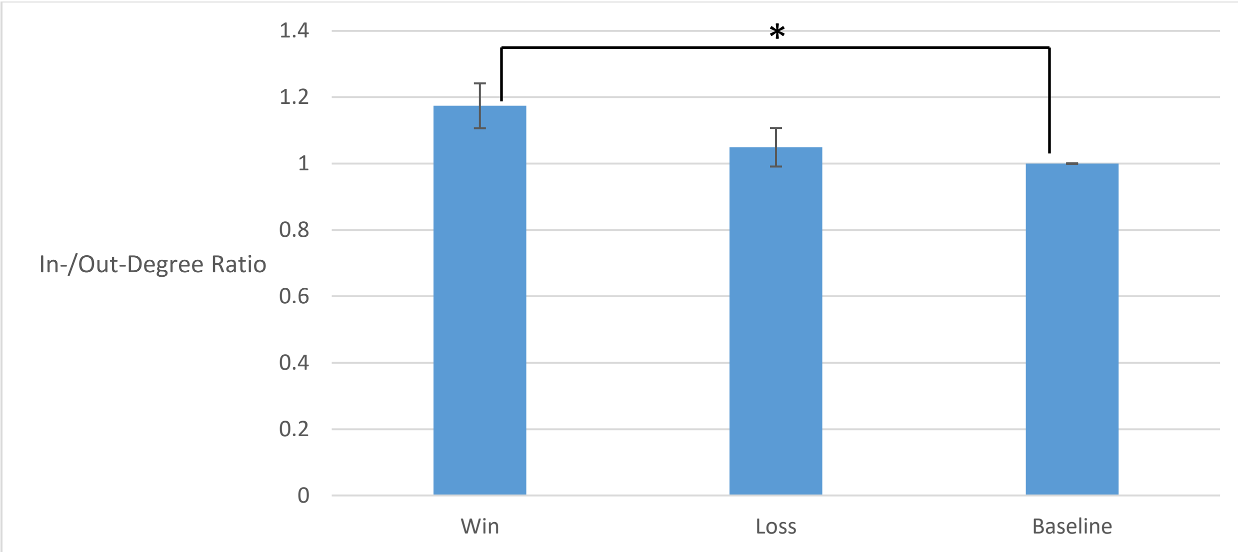

Figure 2.4: Average in-degree centralization/out-degree centralization ratios for winning and losing teams, compared to baseline (1:1)

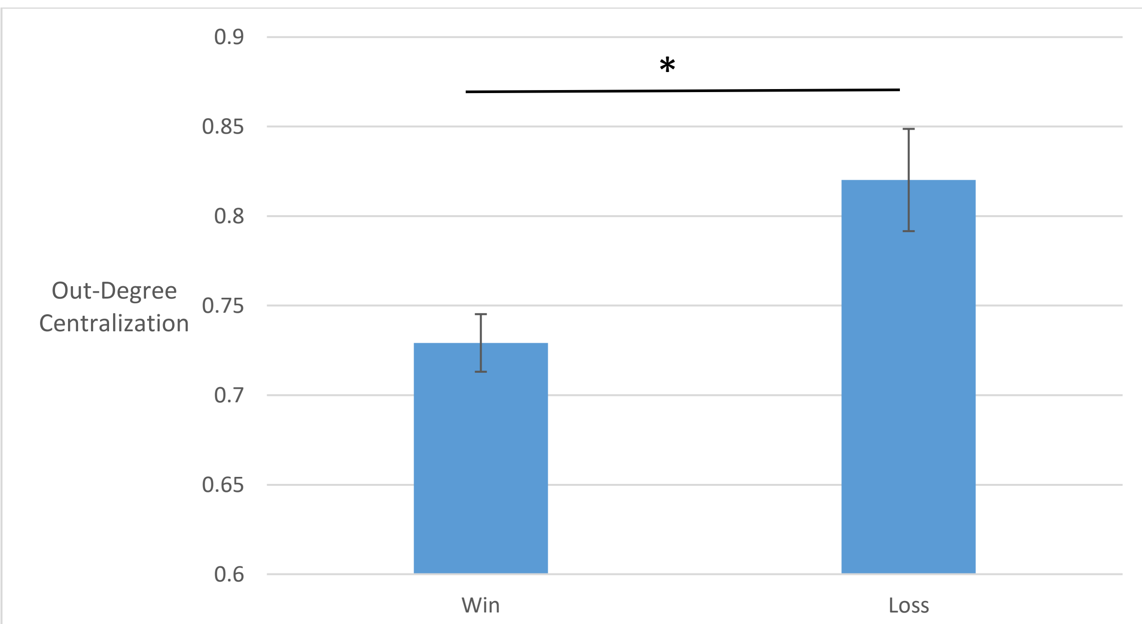

Figure 2.5: Out-degree centralizations for UW wins vs. losses

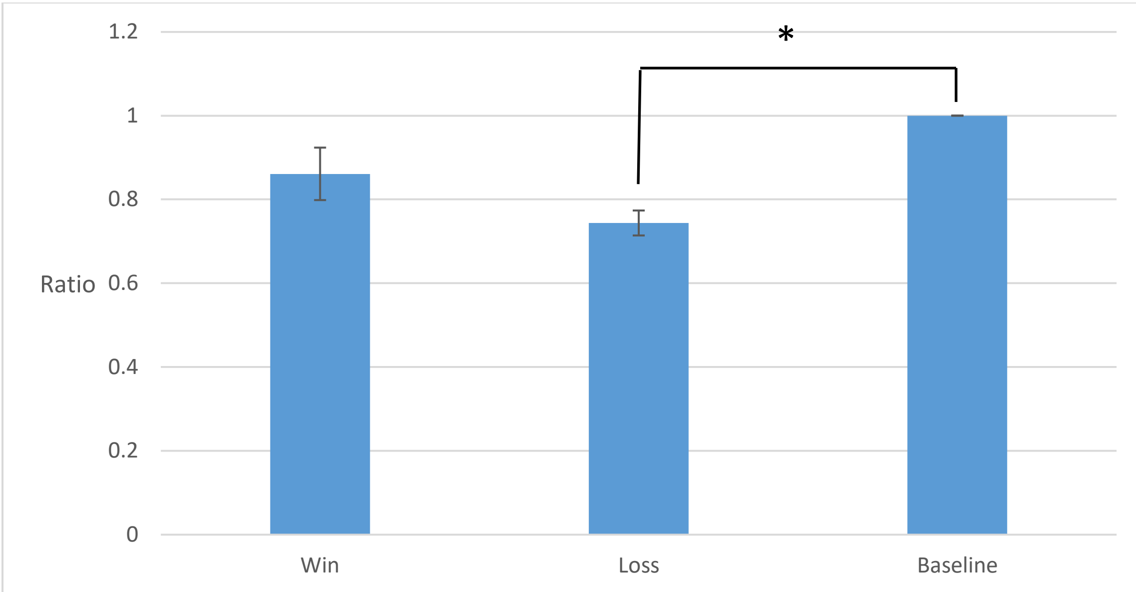

Figure 2.6: In-/out-degree ratios for UW wins and losses compared to baseline of 1:1

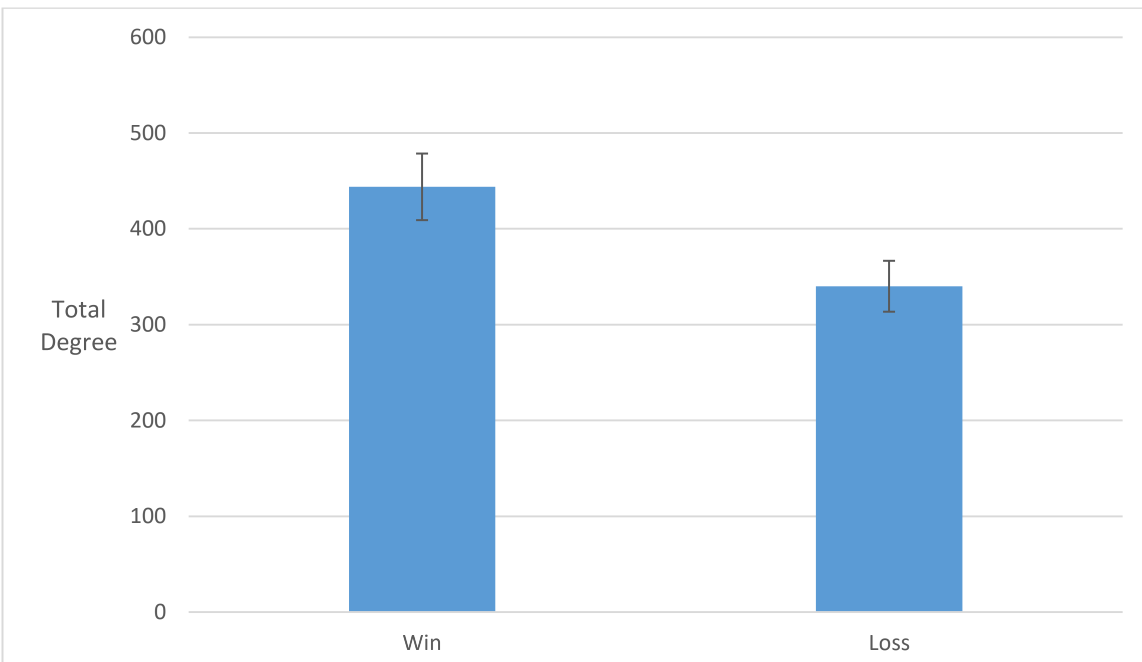

Figure 2.7: Total degrees for UW wins and losses

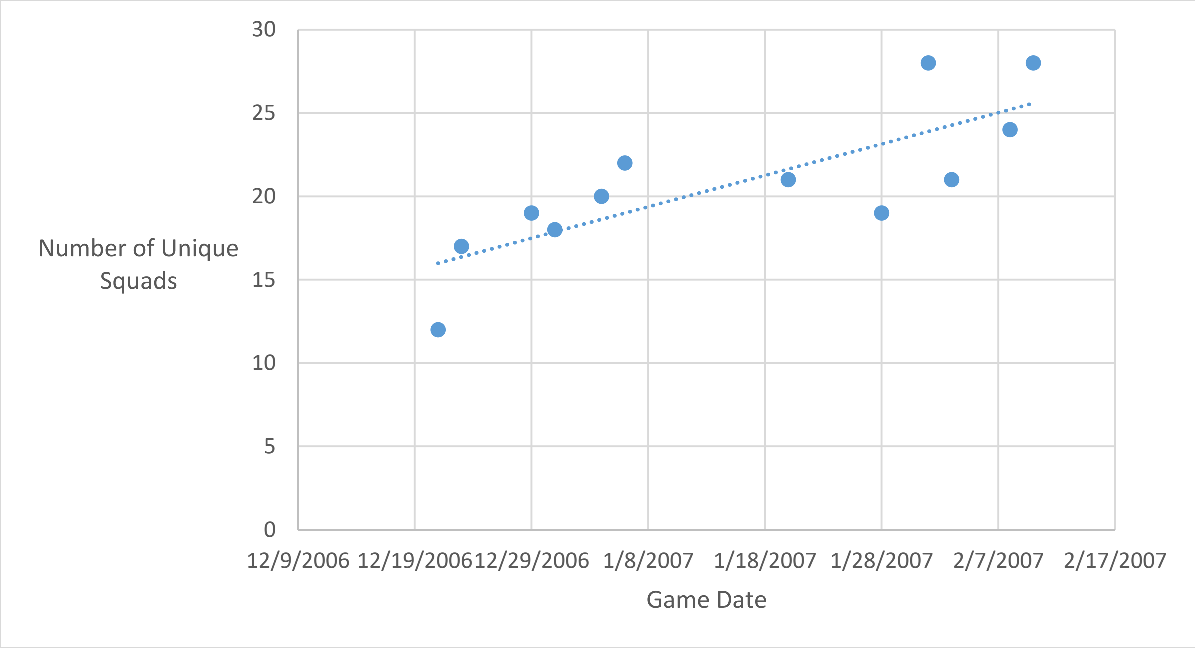

Figure 2.8: UW unique squad combinations over 2006-2007 season

Chapter 3. DEGREE CENTRALIZATIONS, RECIPROCITY, AND TOTAL DEGREE AS PREDICTORS OF SUCCESS AMONG NBA TEAMS

Social network analysis (SNA) is a useful tool in mapping and analyzing interactions between individuals at many levels, from single dyads to entire populations, and can be applied to essentially any situation which involves group relations. SNA has been used to examine interactions between animal populations (Wey, Blumstein, Shen, & Jordán, 2008), coworkers in corporate work groups (Cross, Borgatti, & Parker, 2002), students in educational settings (Carolan, 2014), in the social sciences (Borgatti, Mehra, Brass, & Labianca, 2009), and many more areas. Its versatility makes it valuable to a large number of researchers studying diverse questions.

One area that can be translated into social networks very easily is that of team sports. Each game or match can be coded essentially as a weighted, directed network, with players forming the nodes and passes between said players, the edges. All the participants and their interactions are fairly easy to track, thus avoiding the observer bias that more open networks tend to fall prey to, and for more popular sports, such statistics can be readily obtained online for free. In fact, several studies have already used network metrics to quantify passing patterns in team sports, including men's soccer (Clemente, Couceiro, Martins, & Mendes, 2015; Kooij et al., 2009) and basketball (Bourbousson, Poizat, Saury, & Seve, 2010; Fewell et al. 2012; Vaz de Melo, Almeida, & Loureiro, 2008). Clemente, Couceiro, Martins, & Mendes (2015) graphed matches among teams in the First Portuguese Soccer League and found there was no single "star" player who monopolized a disproportionate amount of each team's passes. Grund (2012) found similar results within the English Premier League; lower degree centralization was associated with better team performance. Fewell et al. discovered too that there was often no star player among these teams, and that those teams with higher entropy (aka more unpredictable passing patterns) performed better. Decentralized teams may perform better because they provide opposing teams less opportunities to “double-team” a single star player, if all players are equally likely to receive passes and/or attempt shots. A lower degree centralization could also reflect that players on these teams have easier access to several of their teammates, and thus more chances to complete a successful pass. Sparrowe, Liden, Wayne, and Kraimer (2001) asked the same question of work groups at business companies, with higher group ratings from managers as their measure of success. Contrary to expectations, they found only weak support that centralization in a work group’s advice network negatively influences its performance. This may point to inherent differences between sports teams and workplace groups. For instance, although both cooperative within their own groups, sports teams engage in more formalized competition between groups, with clearly delineated rules, goals, and wins. Companies on the other hand, while still competitive with each other, have very few codified “rules” on how to succeed, and do not necessarily engage in zero-sum games. There is not always a clear “winner” and “loser” in business campaigns, and competition is not limited to discrete games at a time. These differences can definitely affect the structure of their respective networks, and roles of nodes within that network.

Degree centralization can be divided in a directed network into in- and out-degree centralizations. This could be useful in sports networks because it provides us with more detailed information on whether all players are receiving passes at the same rate (in-degree centralization), or passing on the ball at the same rate (out-degree centralization). Previous data on NCAA women's basketball suggests that successful teams possess a higher in-degree centralization/out-degree centralization ratio than 1:1, indicating that there is more variability among teammates in who receives passes than who makes them (Chapter 2). However, this analysis was conducted on college women's teams, which of course differ substantially from professional men's basketball. We repeat the same ratio comparison on our NBA dataset to test whether the same strategy is equally effective when employed in different professional and gender settings.

In our previous dataset, we also examined the effects of individual centrality within a team network on each player's success per game. Do players who are involved in the greatest number of passes in a given game (AKA sport the highest degrees) also score the most points for that game, compared to their less well-connected teammates? Having more access to the ball could increase a player's chances of being able to shoot that ball if opportunity presents itself. In the workplace, Sparrowe, Liden, Wayne, and Kraimer (2001) measured individual workers and their positions in advice networks relative to each other as well, and found that more central individuals within those groups receive higher ratings than their peers. Thus, better-connected nodes may succeed on the individual level while uniformly well-connected networks succeed at the population level (in other words, it is best to be the most-connected node in the least centralized network). Similarly, Brass (1981) discovered that employees' centrality in a company's work flow network had an indirect effect on their performance, while Baldwin, Bedell, and Johnson (1997) reported that M.B.A. students with higher centrality within their class groups earned higher grades.

On the other hand, these findings may be limited to non-team sports networks, as our NCAA data showed the opposite effect – high scorers do not usually also possess the highest degree. Basketball teams consist of very discrete positions that are defined by specific skillsets that vary little across teams, unlike work groups in which positions could be more flexible and depend upon the company. Coworkers may have significant overlaps in their skillsets and thus be able to swap positions more easily, but among basketball teammates, their various positions could vary much more greatly (especially if dependent on physical characteristics such as height and speed) and be difficult to switch. Someone who is trained as a center may do poorly as a forward or guard. Among the NBA, it is very rare to find “five-tool” players: players that excel in points, rebounds, assists, steals, and blocks (Rose, 2016). In fact, when Isaiah Thomas, a retired NBA player, was hired as president of operations by the New York Knicks in 2003, he selected only from the highest-scoring players in the league, reasoning that such an “all-star team” would outscore everyone else. Contrary to expectations, the Knicks did very poorly under his management, with four straight losing seasons and an overall 66% loss. As it turns out, only focusing on shooting left the team vulnerable to turnovers and unable to effectively defend against their opponents. Even their offense suffered, as players were more focused on scoring themselves than helping their teammates score. Thus, in basketball, it is important for players to specialize and cover all positions, and not rely on one or a few players for everything. In this case, we should expect players directing the most passes between teammates and players scoring the most points to not be the same individuals.

Besides using entire games as units of analysis, we can also apply SNA within games by splitting each up into its first and second half, then comparing their metrics to chart passing behavior as game time progresses. Clemente, Couceiro, Martins, & Mendes (2015) found that in the First Portuguese Soccer League, all players decreased their number of connections in the second half, suggesting the entire team participated less after halftime. This could be a sign of fatigue, change in strategy, or "eating up the clock" as the game winds down and the winning team attempts to hold on to its lead. Given the similarities between team sports such as soccer and basketball, we would expect to see the same pattern in NBA games when comparing first and second halves: teams will have lower in-, out-, and total degree centralizations as well as a lower total degree in the second half.

Because basketball, like all team sports, is inherently a cooperative effort, reciprocity is another useful metric to study due to its close ties with cooperation. Reciprocity refers to the existence of a mutual link between two nodes in a directed network. Significant patterns of reciprocity can reveal important information on network and node characteristics that drive edge formation, which can be lost if the network is treated as undirected (Garlaschelli & Loffredo, 2004). For example, Leider, Möbius, Rosenblat, & Do (2009) discovered that in their real-world social networks, participants formed more reciprocal altruistic ties with friends, and when future interactions were more likely to occur. This hints at an "enforced reciprocity" wherein repeated interactions within the same network drives reciprocity between its members. However, to our knowledge no work has been done on whether a higher level of reciprocity increases a network's effectiveness, particularly compared to other networks with less reciprocity. Given that reciprocity is tied to higher levels of cooperation within groups (Bowles & Gintis, 2004), and team sports depend on cooperation among teammates, we hypothesize that teams with higher reciprocity will outscore those with lower reciprocity.

However, it is also important to take into account inter-team variability that could arise from coaching or player effects. Coaches of course employ different strategies from each other, and a team's play style could also be influenced by the players on that team. Simply comparing the two opposing teams per match may allow these variables to confound our results. With that in mind, we also repeated our analyses within teams by comparing the networks of the same teams across seasons when they won and when they lost. Doing so should reduce individual differences between teams, as we can be reasonably certain each team did not change its coach or roster significantly during the season.

Our current study thus makes the following predictions, both between and within teams: Winning teams will have: 1. lower degree centralizations; 2. in-/out-degree ratios higher than 1; 3. higher network reciprocity; and 4. separate players scoring the most points and possessing the highest degrees, compared to losing teams. We also expect to see our degree measures decrease for each team between the first and second halves of a game.

3.1 METHODS

Our data consisted of 8114 games from the 2001-2010 NBA seasons, downloaded from www.basketball-reference.com. Players were coded as nodes and passes as weighted, directional edges. We confined our analyses only to successful passes between teammates, excluding outcomes such as interception by the other team or turnovers due to fouls; however, we included both successful and unsuccessful shots to capture all teammate passes within a game. The data only included the final dyad of each possession between teammates, but validity testing done on a previously-used dataset indicate that this did not change the results significantly compared to using the entire possession's passes. We used our NCAA college women's basketball data (Chapter 2) to test the validity and found that only including the final dyads per possession did not significantly alter our results. For instance, our *p*-value for comparing degree centralization in winning and losing teams was 0.78 when the entire data set was used, and 0.85 with just the last dyads. When we compared the results of our in-/out-degree ratios for winning teams, we found p = 0.02 and p = 0.24 for all dyads vs. last dyads, respectively. If anything, only using the final dyads of each possession caused our results to be even more conservative.

Weighted and directed degree centralizations for each team were calculated in R, based on Borgatti, Everett, & Freeman's formula for UCINET (1992). The in-, out-, and total degrees for each node was divided by the value of the node with the maximum score to standardize it against bias towards teams that may have a higher number of passes simply because they go through more possessions than other teams. Then we subtracted each individual degree from the max degree value, as a measure of how much they vary from the potential "star player" on the team with the most passes, and divided that value by the total number of nodes minus 1, to control for networks of different sizes; finally all weighted degrees for each team were summed together. The result is a number between 0 and 1, with 0 indicating that everyone on the team is receiving the exact same number of passes, and 1 indicating that there is one player receiving all the edges from everyone else and no one else is connected, i.e. a wheel network (Borgatti, Everett, & Freeman, 1992). . It should be noted that total centralization was not simply a sum of in-degree and out-degree centralizations: each type of centralization was calculated separately as comparisons of variance among players for that particular degree type. For example, a team in which four players all only passed to the fifth but never to each other would receive an in-degree centralization of 1 since only Player 5 has any in-degree, an out-degree centralization of a little over 0 because all players other than Player 5 share the same out-degree, and a total centralization of 1 because Player 5 is the only player who is connected in some way to their teammates.

Each team's ratio of in-/out-degree centralizations was calculated as a measure of variation in receiving passes versus passing among players, and compared to single-sample *t*-

tests with a null ratio of 1 (equal proportions of players receiving and passing on the ball) for both the winning and losing teams per round.

For each game, we recorded the top scorer for both teams (www.stats.nba.com). Network reciprocity was calculated using the R packages igraph and NetData, by dividing the number of mutual dyads over the total number of dyads. We split each game into first (quarters 1 and 2) and second (quarters 3 and up) halves, and compared their metrics against each other. All variables were calculated and run through GLMs in R. Team ID was coded in as a covariate to control for inter-team variability. Effect sizes were calculated from odds ratios according to Chinn's (2000) methodology.

For our within-teams analyses, we averaged all the above metrics across each NBA team separately under two conditions, win and loss, and paired the results by team for comparison.

3.2 RESULTS

3.2.1 *Between teams*

Out-degree centralization positively predicted wins (p = 0.004), with an odds ratio of 0.1.66, Table 3.1. In other words, teams with higher out-degree centralizations were 66% more likely to win. Both total degree and reciprocity significantly predicted success positively (p < 0.001) with odds ratios of 1.06 and 1.60, respectively. Individual centrality significantly negatively predicted success (p = 0.004) with an odds ratio of 0.89; winning teams were 11% less likely to have their high scorer also be the player with the highest degree. Total centralization marginally negatively predicted success (p = 0.05). In-degree centralization did not significantly predict success (p = 0.17). These results were also borne out by a stepwise regression including all variables (Table 3.3), with out-degree centralization and reciprocity sporting the highest odds ratios once again (1.11 for both). Additionally, out-degree centralization and reciprocity returned the largest effect sizes (0.06 for both) as well, reflecting their strong positive effect on success.

Winning teams had lower in-degree centralization ($t(8113) = -6.16$, $p < 0.001$, Figure 3.1), higher out-degree centralization ($t = 5.54$, $p < 0.001$, Figure 1), and higher total degree centralization ($t(8113) = 4.96$, $p < 0.001$, Figure 3.1) than losing teams. Both winning ($t(8113) = -21.58$, $p < 0.001$) and losing teams ($t(8113) = -5.55$, $p < 0.001$) had in-/out-degree ratios significantly lower than 1:1, Figure 2. Winning teams also showed higher total degrees ($t(8113) = 46.12$, $p < 0.001$, Figure 3.3) and reciprocity ($t(8113) = 13.68$, $p < 0.001$, Figure 3.4) than losing teams.

Overall, the player with the highest degree on each team for each game was significantly unlikely to also be the top scorer ($N = 7926$), $\chi^2(1) = 8.71$, $p = 0.003$. The same pattern is reflected among winning teams, players with the highest degree are significantly less likely to score the most points ($N = 3811$), $\chi^2(1) = 29.83$, $p = 0.001$. However, among losing teams players who scored the most points ($N = 4115$) are not significantly more or less likely than their teammates to also boast the highest degree for that game, $\chi^2(1) = 1.66$, $p = 0.20$.

During the second half of the games, teams displayed significantly lower out-degree centralization ($t(16226) = 4.27$, $p < .001$, Figure 3.5), total centralization ($t(16226) = 11.06$, $p < 0.001$, Figure 3.6), and total degree ($t(16226) = 25.71$, $p < 0.001$, Figure 3.7). There was no significant change in in-degree centralization between first ($M = 0.65$, $SD = 0.18$) and second halves ($M = 0.65$, $SD = 0.20$), $t(16226) = 0.10$, $p = 0.92$.

3.2.2 *Within teams*

When comparing within teams, only total degree returned a significant *p*-value ($p = 0.005$), Table 3.2. However, out-degree centralization is actually the strongest positive predictor of success according to its odds ratio (6.94). After running all predictors through a stepwise regression, we still see the same effects of out-degree centralization and total degree (Table 3.4), with out-degree centralization once again most strongly predicting success with an odds ratio of 13.08 and effect size of 1.42.

Teams showed lower in-degree centralization ($t(32) = -6.38$, $p < 0.001$, Figure 3.8), higher out-degree centralization ($t(32) = 5.24$, $p < 0.001$, Figure 3.8), higher total centralization ($t(32) = 3.62$, $p = 0.001$, Figure 3.8), higher total degree ($t(32) = 27.58$, $p < 0.001$, Figure 3.9), and higher reciprocity ($t(32) = 11.21$, $p < 0.001$), Figure 3.10) when they won compared to when they lost. Both wins ($t(32) = -7.24$, $p < 0.001$) and losses ($t(32) = -3.14$, $p = 0.004$) had centralization ratios significantly lower than 1, Figure 3.11.

Correlation matrices for between-teams (Table 3.5) and within-teams (Table 3.6) analyses are included as well.

3.3 DISCUSSION

Of our centralization measures, only out-degree centralization significantly predicted success, in the positive direction. Teams that won matches were more likely to concentrate their passes to one or a few “star” players. Previous findings in men’s soccer (Clemente, Couceiro, Martins, & Mendes, 2015; Grund, 2012) found the opposite effect that less centralized passing patterns were more common among men’s soccer, but they focused on total degree centralization only. We also found total centralization to marginally predict success negatively, suggesting it is not as important as out-degree centralization but can still decrease a team’s chances of winning slightly. This is consistent with both the men’s soccer results and Fewell et al.’s (2012) findings that entropy but not overall centralization significantly impact success. Basketball could differ from soccer, if it is less important in the former sport that all teammates touch the ball equally and more important that the team is less predictable in its passing patterns. This would suggest more of a path-based than player-based strategy, at least for men's basketball. Soccer also involves more players on the field (11) than basketball (5), so it's possible that a lower centralization is easier to achieve or more useful with more players. On the other hand, if soccer analysis is split into separately analyzing in-, out-, and total degree centralization, we may see similar results regarding out-degree but not in-degree or overall centralization affecting success. Our results may also differ from previous studies conducted on task-oriented groups such as Sparrowe, Liden, Wayne, & Kraimer (2001) and Leavitt (1951) due to the physical and strategic differences between performing group tasks in the laboratory and playing sports on a court. Decentralization in a network might be more useful when members are merely seeking advice from their "teammates" to solve puzzles rather than physically engaging in coordinated activity against a rival team. Furthermore, Sparrowe et al. (2001) found only marginal support for less centralized advice networks doing better in workplace settings. Taken together, this suggests that lower centralization in a network does not automatically guarantee success; other factors such as the setting/environment and type of task must be taken into consideration as well. Since we did not find any significant effects of centralization at all within teams, this is further reinforcement that there is no "minimum" or "maximum" value of out-degree centralization that guarantees success; the trick is to adjust the team's passing patterns according to each opponent. The high positive correlation between out-degree centralization and total centralization in both our analyses (r = 0.43 for between teams, r = 0.58 for within teams, Table 3.5) suggests that the two are not completely independent, and that much of total centralization's effect on success can be explained by out-degree centralization.

Further, teams with higher reciprocity performed better than their opponents. Higher reciprocity indicates that a team is less likely to rely on a just a couple players to shoot or assist, reflecting that more of their players are skilled in both passing and scoring. This also shows flexibility in roles, which is a sign of cooperation as teammates are willing to switch off in the shooting and assisting roles with each other. Moreover, increased reciprocity among teammates can prevent the opposing team from simply blocking a team's best shooter each time, or it could signify better team cohesion, with players more likely to share the ball with their teammates. Interestingly, not only was this effect not found within teams at significant levels, we actually see a negative correlation between reciprocity and success. Dropping total degree from the model does create a significant positive effect of reciprocity, suggesting that much of the positive relationship between reciprocity and success even in our between-teams analysis could be moderated by the effect of total degree. This is confirmed by the strong positive correlation between reciprocity and total degree (r = 0.73, Table 3.6). In fact, total degree boasted one of our highest significant values for between teams and was the only significant predictor within teams, strongly supporting that teams with higher number of passes do much better, compared with both their opponents and themselves. Obviously, more passes per possession indicates a that a team is able to control and retain the ball for a longer period of time while also decreasing their opponents' access to the ball. It also could create more chances and pathways for the ball to reach the hoop, and reflect greater stamina and use of strategy by the players.

The ratio of in-degree to out-degree centralization in our NBA games between teams is directly opposite compared to what we saw in our NCAA women's basketball data. Having a ratio less than 1:1 reflects a pattern in which teammates receive passes at relatively equal rates but only one or a couple players are monopolizing out-going passes. NBA teams therefore may be more reliant on their best players making decisions on who to pass to, unlike college teams in which the best players receive a disproportionate amount of passes but then pass the ball equally to their teammates. Since we only used the last dyads of each possession, the players with the highest out-degrees may be the ones with the most assists while lower out-degrees indicate shooters. However, we did not factor in the outcomes of each attempted shot, whether it was successful or not. So in winning teams especially, perhaps once a player receives the ball, it is less important who he passes to than whether he chooses to pass at all or to attempt a shot. Or it is possible that these teams only have a small number of “star” assisters with high out-degrees who then pass to a greater number of shooters, all with the same relative in-degrees. If this is true, then successful teams should rely on their assisters to decide who is in a good position to shoot, rather than on their scorers to set up a good shot. This could be important for coaches in deciding how to structure their plays and where to place their best players to increase the chances of a successful outcome. However, our non-significant results for our within-team analysis suggests that a lower ratio does not always guarantee success. It is probably more important for a team to have a lower ratio than its opponent each game than simply compared to baseline.

As stated above, both winning and losing teams were less likely to rely on the same player for scoring and passing, and having the same player possess the highest centrality and points per game was a significant negative predictor of success. Therefore, the most successful player is not usually the most central one, although only measuring success in terms of points scored is inadequate and one-dimensional, as Rose (2016) points out. Central players are just as important for their roles in facilitating passes and ensuring their teammates have opportunities to score. In fact, not having to depend on just a couple players to cover all roles makes a team more versatile, and prevents their opponents from double-teaming their star players and rendering the rest of the team helpless. It also helps players hone specific skillsets best suited to their abilities rather than having to spread themselves thin on trying to master many disparate skills.

Finally, our results show that all teams decrease their out-degree and total centralization, and total degree, in the second halves of each game on average. Clemente, Couceiro, Martins, & Mendes (2015) found the same in the First Portuguese Soccer League and posited it could be due to decreasing participation to save players from fatigue, or a switch to focusing on defense after halftime. Basketball teams could certainly employ these strategies as well in the second halves of their games, especially for teams that already have a commanding lead. Coaches may also choose to allow playtime to lesser-used players in the second half if their team is far enough ahead, and swapping them in for more regular players can help even out a team's centralization. Winning teams might try to "eat up the clock" by holding the ball for as long as possible to deny their opponents possessions and chances to score; this would drive down the total number of passes (total degree) if players stall for time by dribbling instead of passing or shooting. Future analysis focusing on specific strategies and plays in the first vs. second half of a game, such as the amount of time spent on offense vs. defense, could further help to illuminate such strategies.

Some limitations of our study include only using the final dyads of each possession, though again validity testing did not turn up any significant differences compared to using the entire possession. We were able to control somewhat for inter-team variation by comparing the same team's metrics when they won to when they lost, but this analysis would be greatly enhanced by following the same team through the entire season, or even multiple seasons. Doing so would not only give us more data to compare their passing behavior in successful vs. unsuccessful games, but allow us to track whether their networks change over time. While we did analyze across a decade of NBA games, we did not break the dataset into separate seasons or years. Doing so in future studies and comparing between and within teams across time would reveal even more information on how a team's network metrics change and affect its success dynamically. Additionally, as stated above, we did not include outcomes such as successful/unsuccessful shots, fouls, or turnovers; which could provide even more material to analyze. Finally, our NBA data is not directly comparable to our previous NCAA women's data due to varying in both gender and professional level.

Subsequent studies should run separate analyses on women's and men's games while holding the level of professionalism constant, or vice versa, to isolate gender or play level differences. Doing so may reveal whether the success of various types of networks is affected by these factors, which in turn could prove useful to coaches and players in choosing which network patterns to apply to their own teams. Studying the success rate of different types of networks against varying defensive strategies employed by opposing teams would also provide us with a more holistic view of the game, and further help coaches decide which plays to apply against specific opponents. It would be useful in addition to divide games into individual possessions and analyze each separately for centrality, then study whether different passing patterns per possession within a single game correlates with a successful vs. unsuccessful outcome.

Another area of interest may be how well players' off-court social networks with their teammates map onto their in-game passing patterns. For instance, if a team's friendship network correlates significantly with its passing network, it may indicate preferential passing between close social associates, which of course would have an effect on the efficiency of a team's on-court behavior. Finally, our study is, to our knowledge, the first to apply a network's reciprocity metric to its effectiveness, and this type of analysis can be more widely applied to areas outside of sports, such as businesses and other task-oriented groups, community networks, and more. If increasing reciprocity levels in these other types of networks also leads to improved efficacy, this would be valuable information for managers or community leaders, etc., who wish to increase their networks' success.

3.4 TABLES AND FIGURES

	Estimate	*p*-value	Odds ratio	Effect Size
In-degree Centralization	-0.211142	0.17171	0.80965940	-0.11641
Out-degree Centralization	0.503957	0.00425 **	1.65525840	0.277846
Total degree Centralization	-0.357718	0.05482	0.69926993	-0.19722
Total Degree	0.062946	< 2e-16 ***	1.06496970	0.034704
Reciprocity	0.471769	7.88e-06 ***	1.60282724	0.2601
Centrality	-0.112272	0.00350 **	0.89380092	-0.0619

Table 3.2: Between-teams GLM results

	Estimate	*p*-value	Odds ratio	Effect size
In-degree Centralization	-7.3195	0.86320	6.625064e-04	-4.03545
Out-degree Centralization	50.2921	0.21503	6.943829e+21	27.72749
Total degree Centralization	74.3621	0.24087	5.069368e-33	40.99797
Total Degree	1.4368	0.00486 **	4.207356e+00	0.792149
Reciprocity	-5.0949	0.90954	6.127856e-03	-2.80897

Table 3.2: Within-teams GLM results

	Estimate	*p*-value	Odds ratio	Effect size
In-degree Centralization	-0.0408344	0.221064	0.9599881	-0.02251
Out-degree Centralization	0.1057906	0.004758 **	1.1115891	0.058325
Total degree Centralization	-0.0780107	0.050376 .	0.9249545	-0.04301
Total Degree	0.0137071	< 2e-16 ***	1.0138015	0.007557
Reciprocity	0.1020640	8.12e-06 ***	1.1074544	0.056271

Centrality	-0.0256432	0.002085 **	0.9746828	-0.01414

Table 3.3: Between-teams stepwise regression

	Estimate	*p*-value	Odds ratio	Effect size
In-degree Centralization	-2.24255	0.402	0.10618730	-1.23638
Out-degree Centralization	2.57114	0.143	13.08069228	1.417544
Total degree Centralization	-2.52657	0.383	0.07993275	-1.39297
Total Degree	0.09691	3.2e-10 ***	1.10175591	0.053429
Reciprocity	-0.81743	0.718	0.44156483	-0.45067

Table 3.4: Within-teams stepwise regression

Row	Column	Correlation
In-Degree Centralization	Ratio	0.550*
In-Degree Centralization	Out-Degree Centralization	0.075
Ratio	Out-Degree Centralization	-0.660*
In-Degree Centralization	Total Centralization	0.280
Ratio	Total Centralization	-0.071
Out-Degree Centralization	Total Centralization	0.430*
In-Degree Centralization	Total Degree	-0.071
Ratio	Total Degree	-0.120
Out-Degree Centralization	Total Degree	0.075
Total Centralization	Total Degree	0.130
In-Degree Centralization	Reciprocity	-0.210
Ratio	Reciprocity	-0.024
Out-Degree Centralization	Reciprocity	-0.170
Total Centralization	Reciprocity	0.071
Total Degree	Reciprocity	0.240

Table 3.5: Correlation matrix for between-teams analysis

Row	Column	Correlation
In-Degree Centralization	Ratio	0.500*
In-Degree Centralization	Total Degree	-0.380
Ratio	Total Degree	-0.260
In-Degree Centralization	Reciprocity	-0.560*
Ratio	Reciprocity	-0.050
Total Degree	Reciprocity	0.730**
In-Degree Centralization	Out-Degree Centralization	-0.046
Ratio	Out-Degree Centralization	-0.870***
Total Degree	Out-Degree Centralization	0.040
Reciprocity	Out-Degree Centralization	-0.270
In-Degree Centralization	Total Centralization	-0.077
Ratio	Total Centralization	-0.530*
Total Degree	Total Centralization	0.260
Reciprocity	Total Centralization	0.230
Out-Degree Centralization	Total Centralization	0.580*

Table 3.6: Correlation matrix for within-teams analysis

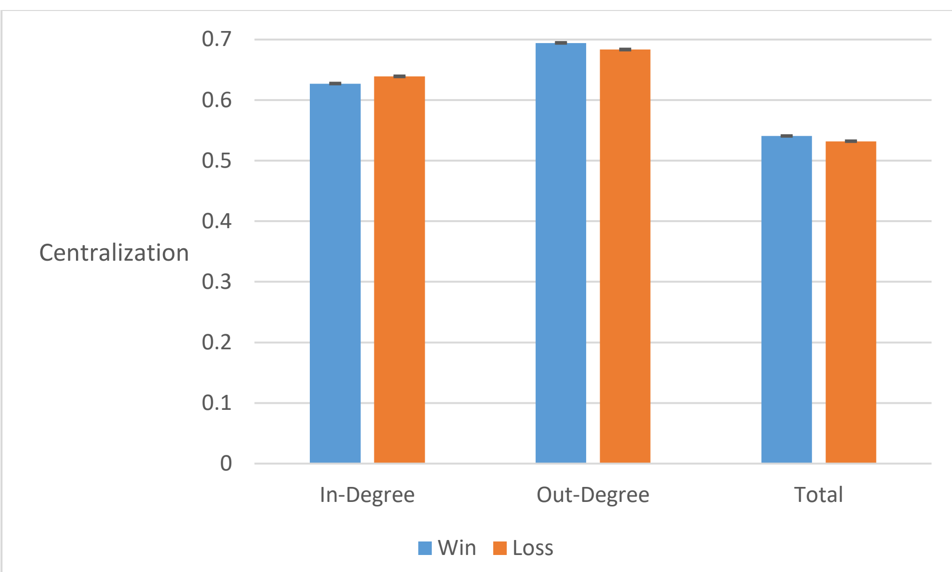

Figure 3.1: Between-teams centralizations. Error bars represent +/-1 SE

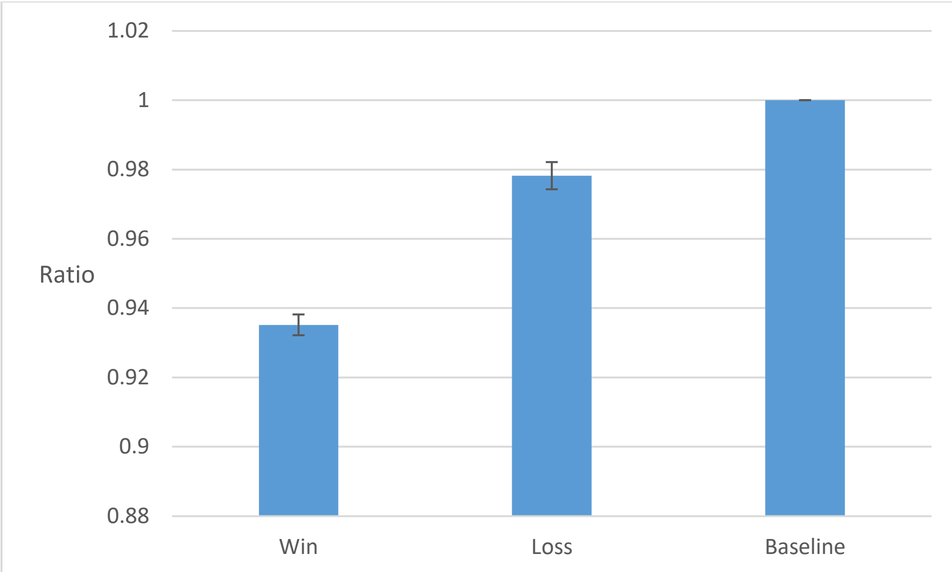

Figure 3.2: In-/out-degree centralization ratio for winning and losing teams compared to baseline

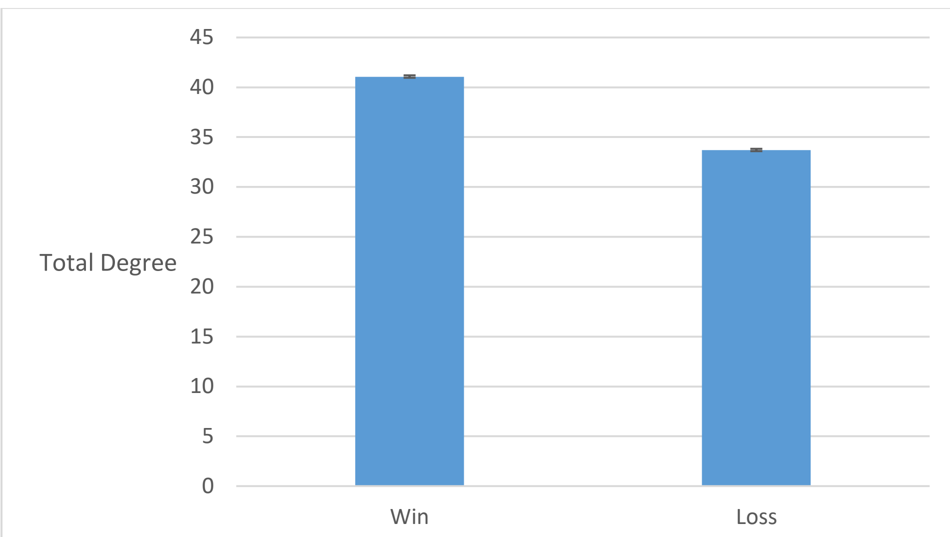

Figure 3.3: Between-teams degree totals

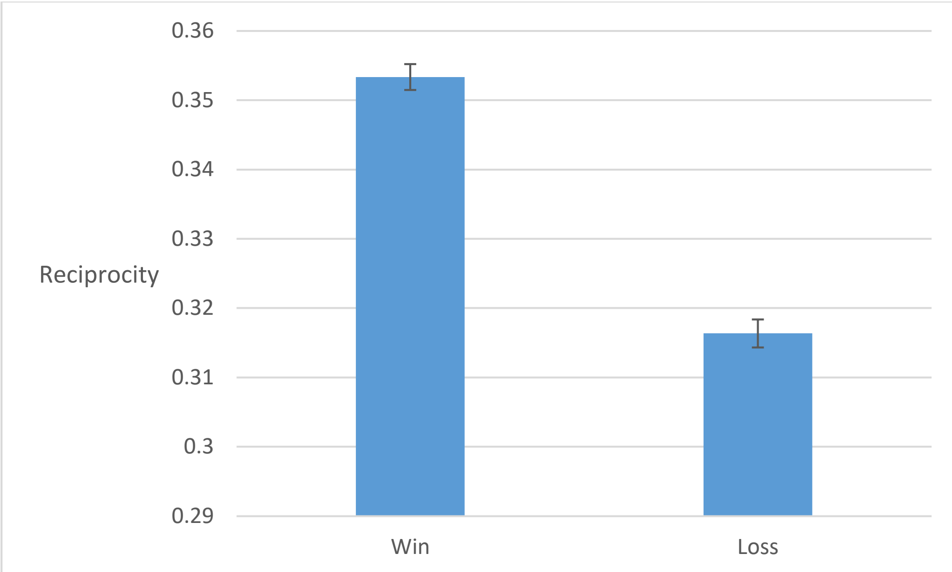

Figure 3.4: Between-teams network reciprocity

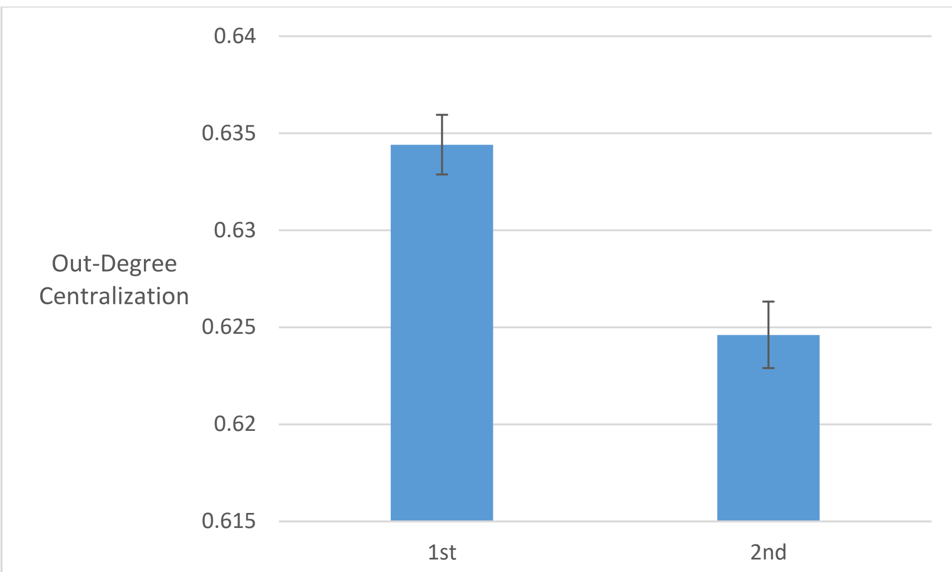

Figure 3.5: Out-degree centralization for first and second halves of each game

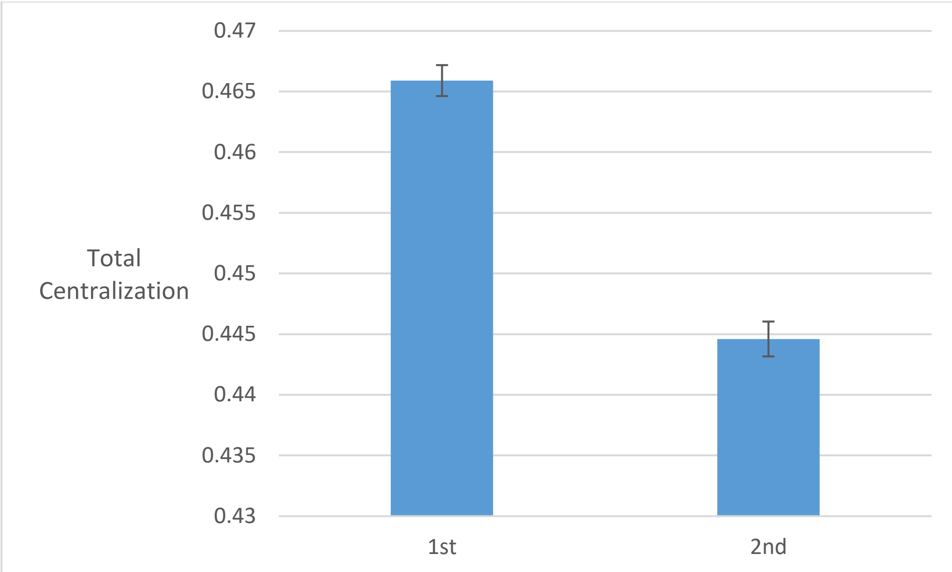

Figure 3.6: Total degree centralization for first and second halves of each game

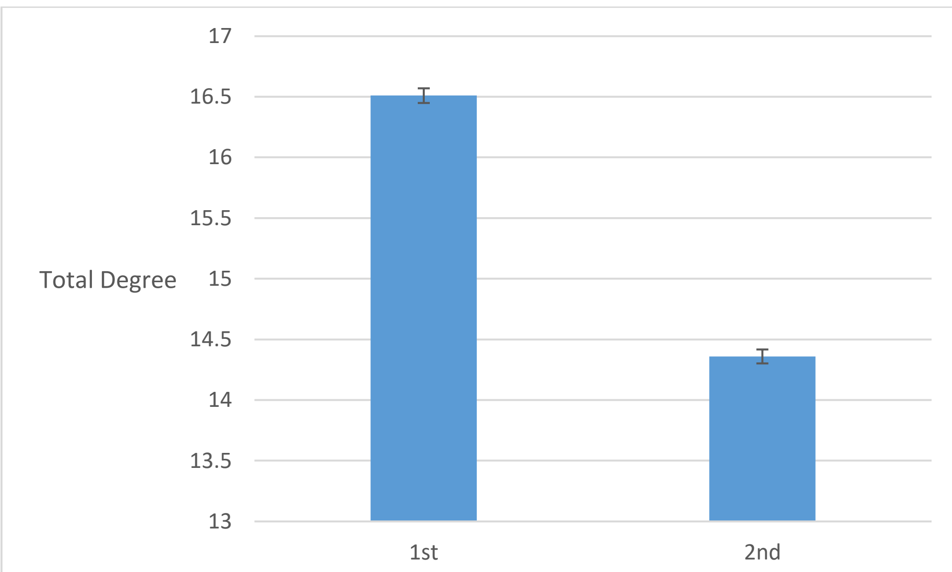

Figure 3.7: Total degree for first and second halves of each game

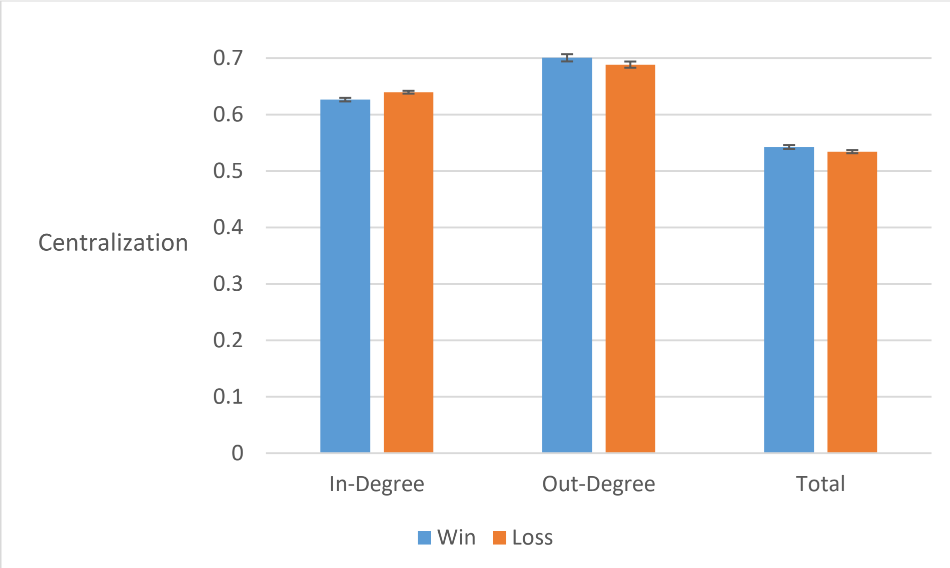

Figure 3.8: Within-teams degree centralizations

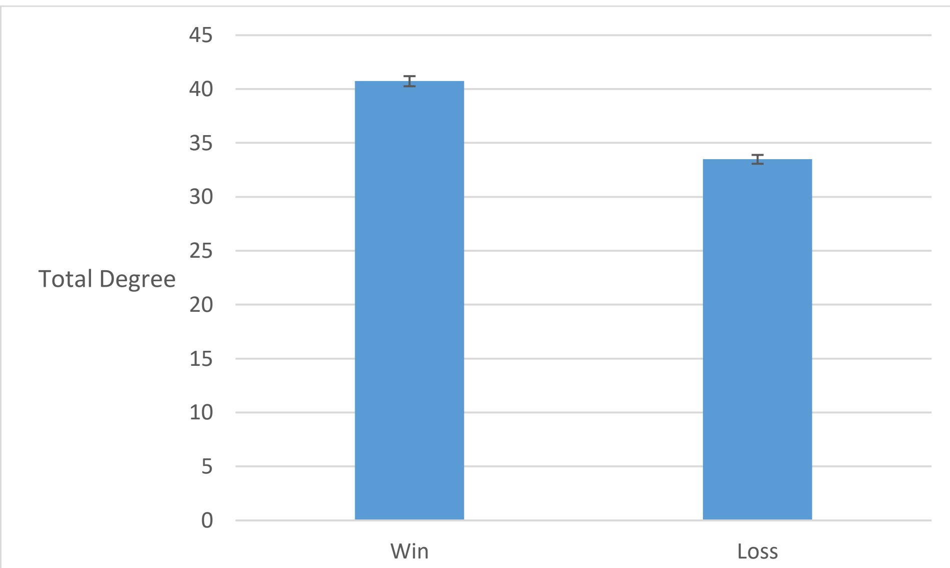

Figure 3.9: Within-teams total degree

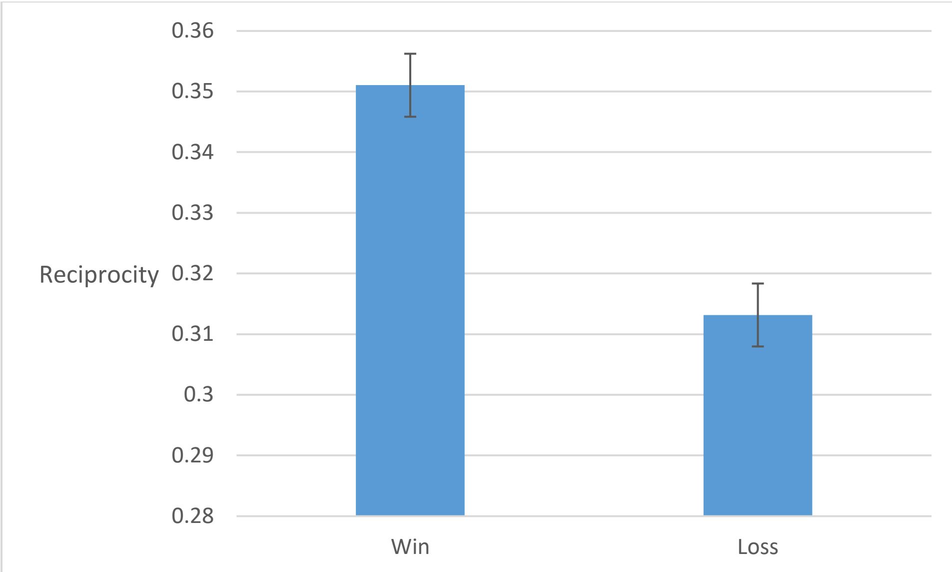

Figure 3.10: Within-teams network reciprocity

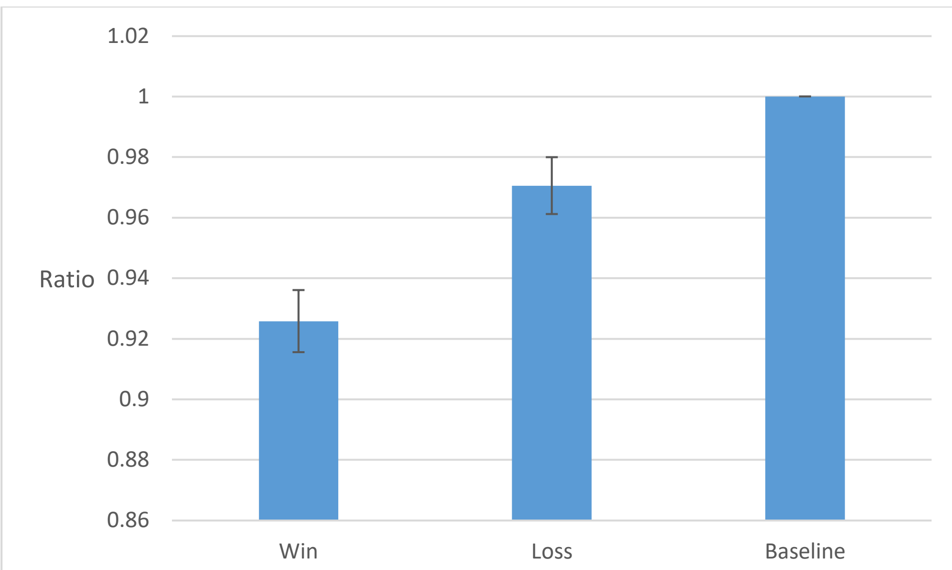

Figure 3.11: Within-teams ratios

Chapter 4. CONCLUSIONS AND FUTURE DIRECTIONS

Team sports such as basketball and soccer rely on the physical prowess and skill of their players, but there is also an underlying strategy to each team and match. Social network analysis (SNA) can help us understand, quantify, and improve upon such strategies. Games can be easily translated into networks, with players coded as nodes and passes as edges, which allows us to not only track and display each team's passing patterns, but compare them with each other as well. In doing so, we can uncover whether some patterns are more effective at winning matches than others, and why. Below, we sum up our research into the efficacy of some of these strategies, as reflected through network metrics.

4.1 DEGREE CENTRALIZATION

Contrary to several previous studies (Grund, 2012; Pena & Touchette, 2012), we found no evidence that less centralized teams perform better than more centralized ones in either our NCAA women's or NBA data when looked at as a whole. In other words, teams with more evenly distributed passing and receiving did not on average outscore their opponents. But when splitting centralization into in- and out-degree centralization specifically, we see significant differences in winning and losing teams in terms of their out-degree centralization. For the NCAA data, out-degree centralization negatively predicted success when compared within teams (see Table 4.1). The University of Washington (UW) women's team performed better when their players passed along the ball more evenly, regardless of who the ball was passed to. These results agree with and refine Grund (2012) and Pena and Touchette's (2012) findings by focusing separately on in- and out-degree centralizations as well as combining them into an overall centralization.

In contrast, in the NBA we found the opposite effect in our between-teams analysis: out-degree centralization is positively associated with success (Table 4.2). More successful teams relied on just a few players receiving the bulk of their teammates' passing. This disparity between the NCAA and NBA is likely due to differences in playing styles at the college vs. professional level. Since lower out-degree centralization is associated with more success in NCAA teams, a sound strategy within college basketball could be to keep outgoing passes unpredictable and thus prevent the opposing team from identifying and double-teaming players that get passed to preferentially. On the other hand, NBA teams, as noted by Rose (2016), are made of players with disparate skillsets; thus not everyone who receives a pass is equally likely to keep the ball in play successfully. Hence, it is sensible to allow these particular players to direct the majority of a possession's out-going passes to ensure a successful play.

Despite finding significant results in terms of out-degree centralization, we did not see similar effects in in-degree or total centralization like previous studies. However these studies were conducted on men's soccer teams rather than basketball. When Fewell et al. (2012) studied the 2010 NBA play-offs, they too concluded that there was no significant difference between winning and losing teams' degree centralizations. Thus, we can assume that lower centralization leads to success in soccer but not basketball, probably due to inherent differences between the two types of sports and their strategies or playing styles. For example, soccer teams consist of 11 players while basketball teams only have five; a larger number of nodes in the network can lead to less centralization more easily because it is easier to evenly distribute edges among more individuals. Fewell and her colleagues did find that teams with higher "entropy", or more unpredictability in their passing, win more games. It's possible that the best strategy in basketball is to vary the path the ball travels through passes between teammates to make it more difficult for the opposing team to successfully interrupt a play. Relying simply on lower degree centralization through distributing passes equally among all teammates may in fact be detrimental if it leads to less-skilled players receiving the ball more often and then either losing it or missing a shot. Therefore, in basketball at least, success relies less on who touches the ball than how the ball moves between teammates.

Outside the realm of sports, previous studies conducted on task-oriented groups such as Bavelas, (1950); Bavelas & Barrett, (1951); and Leavitt, (1951) have reported that decentralized groups perform better than centralized ones when asked to solve puzzles. Our results likely differ due to the physical and strategic differences between performing group tasks in the laboratory and playing sports on a court. Decentralization in a network might be more useful when members are merely seeking advice from their "teammates" on puzzle-solving rather than physically engaging in coordinated activity against a rival team. Furthermore, these studies found decentralized groups to be more effective only when a lot of "noise" is introduced to distract from the task at hand. Under favorable conditions, centralization does not affect success (Bavelas, 1950). It's possible that the games in our data set were not "noisy" enough to induce the need for decentralization; given more difficult matches, perhaps between evenly-matched teams or during the final rounds of a championship, we might see less centralized teams scoring better. In addition, Sparrowe, Liden, Wayne, & Kraimer (2001) found only marginal support for less centralized advice networks performing better in workplace settings. They did discover that in agonistic but not advice networks, having a higher network degree density lowers workplace efficiency, suggesting that degree exerts a stronger effect on negative vs. positive interactions between individuals. Since basketball by its nature as a team sport is a cooperative endeavor, we would not expect the same patterns as reflected in agonistic networks. All in all, we see that lower centralization in a network does not automatically guarantee success; other factors such as the setting/environment and type of task must be taken into consideration as well.

4.2 IN-/OUT-DEGREE CENTRALIZATION RATIO

When compared to each other, our results for the NCAA women's and NBA's average in-degree to out-degree centralization ratio were directly opposite to each other. For both, we initially expected winning teams to not possess a ratio significantly different from 1, since lower centralization in both in- and out-degree were predicted to lead to more wins (see above section). Instead, we found that for the NCAA women's games, winning teams possessed a higher ratio than 1, Figure 1, while in the NBA winners trended toward a ratio lower than 1, Figure 4.2. This ratio served as a measure of a team's variation in receiving over passing rates: a higher ratio reflects a more uneven distribution among teammates of receiving the ball compared to passing it on, and vice versa for a lower ratio.

The NCAA's higher ratio reflects more variation among teammates in who receives the ball compared to who passes. Winning teams thus concentrate their passes toward one or two "stars" who then presumably make the call on who else to pass to, given changing circumstances as the game progresses. On the other hand, within the NBA, successful teams had lower ratios than 1, meaning teammates receive passes at relatively equal rates but only one or a couple players are monopolizing out-going passes. Unlike college teams in which the best players receive a disproportionate amount of passes but are equally likely to receive those passes from any of their teammates, NBA players receive about equal amounts of passes, but their best players are more likely to be making those passes. In other words, the NCAA appears to rely on their star players receiving most of the passes during a possession while the NBA relies on star players making most of the passes. A possible explanation for this difference is that at the college level, teams with the best shooters win more games (since we did not code for the outcome of each possession, attempted and successful shots were not included in a player's out-degree, which would increase their in-/out-degree ratio), but at the professional level it's better to have the strongest players assisting or otherwise directing passes rather than attempting shots. There is probably a higher discrepancy in skill levels among college teammates than professionals, especially in shooting. If most players on professional teams can shoot equally well as their teammates, then a more effective strategy would be to focus on ensuring the ball can safely reach someone who is in a good position to attempt to score. Thus, the strongest players may actually be the ones entrusted to judge how to effectively control ball flow and who to pass to for a shot, rather than shooting themselves.

Of course, we cannot completely separate professional and gender effects between our two datasets, as we were comparing men's professional basketball to women's college basketball. The difference in their results could be due instead to gender rather than level of play. Perhaps women, due to their shorter average height and vertical jumping ability (Hoffman, 2006), find it more difficult to complete a successful shot than men do. If so, it would make sense to concentrate the tallest, strongest players in the scorer position. On the other hand, since it is less physically challenging for most male players to score, the emphasis shifts to ensuring that shooters have a clear shot at the basket. This depends on their teammates successfully controlling ball flow and passing to an individual in a good position to score. Thus, women's teams would be concentrating on increasing their best shooters' in-degrees by passing preferentially to them; the best players on men's teams are instead the ones with higher out-degrees who decide the best shooter to pass to. Or there could be an interaction between both gender and profession level effects in our results. Further analyses holding one variable constant while comparing the other would be very useful in isolating individual effects.

4.3 INDIVIDUAL CENTRALITY

In both the NCAA and NBA, players who scored the highest per game were not usually the ones involved in the most passes, Tables 4.1 and 4.2. In other words, network centrality did not increase individual success. Although previous studies have found a positive association between individual centrality and success (Sparrowe, Liden, Wayne, and Kraimer, 2001; Baldwin, Bedell, and Johnson, 1997), their work focused on advice and information networks rather than the more tangible edges of ball paths during a basketball game. In a physical contest such as team sports, each player has his/her own unique set of skills that is harder for others to emulate, unlike business and classroom networks where these roles are less physically defined and easier to overlap. As Rose (2016) points out, only focusing on a single skill such as shooting can cause a team to perform worse, as they are unable to mount effective defense and offense against their opponents. Hence, it makes sense for some players on a team to focus on trying to score points while others focus on assisting or directing the flow of passes, instead of one player trying to do everything. And again, this division of roles prevents the other team from singling out and disabling key players, thereby disrupting the focal team's plays.

4.4 TOTAL DEGREE

Total degree for a team per game increases both NCAA and NBA teams' chances of success, Figures 4.3 and 4.4, even after taking the total number of possessions into account. Simply going through more possessions and chances to score aren't enough to guarantee success; teams that engage in more passes per possession are the ones more likely to win. Doing so creates multiple paths to the basket and makes it harder for the other team to intercept the ball at any one path. It also helps decrease the predictability of a team's passing behavior, which as stated above improves the odds of scoring (Fewell et al., 2012). Finally, more passes per possession can "eat up the clock" and deny the opposing team more time for their own possessions and plays, especially if a team is already ahead and wishes to hold onto that lead. Within the NBA, we see that all teams decrease their total numbers of passes from the first to second halves of a game. One explanation is a shift in focus from offense to defense, which again points to less of a drive to score themselves and more to prevent the other team from scoring. Another reason for the decrease is to save players from fatigue and injury by lowering participation for the entire team. This should be further explored by separately analyzing the changes in passes between specific positions from first to second halves: While the number of passes may decrease overall, they may be increased between the point guard and small forward, for instance, indicating a more defensive strategy centered on drawing fouls from the other team rather than scoring.

4.5 RECIPROCITY

Among NBA teams, higher reciprocity contributes to success, indicating that winning teams are less likely to rely on a just a couple players to shoot or assist. This reflects that more of their players are skilled in both passing and scoring, and would keep the opposing team from simply blocking a team's best shooter each possession. It also indicates less selfish teams, with players more willing to share the ball with their teammates, which might help increase team cohesion and trust. It could be informative to investigate whether teams with higher reciprocity in passing behavior also display stronger friendship networks off the court. Higher levels of cooperation within groups are associated with increased reciprocity scores (Bowles & Gintis, 2004); reciprocity then is a good measure of cooperation. Given that basketball is inherently a collaborative game, reciprocity could provide coaches with a useful metric to track cooperation between their players, or even to increase it through increasing reciprocity in plays.

4.6 APPLICATIONS

Understanding how passing behavior affect a team's success through social network metrics can help coaches and players improve their strategies and reveal vulnerabilities within their own teams. Since we've seen that basketball appears to follow more of a ball-path-based than -player-based strategy, coaches can use this knowledge to adjust their plays accordingly by placing emphasis on creating a clear path for the ball to reach the basket rather than focusing most passes to a few star players. Doing so would make it more difficult for opponents to identify and double-team the best players. Our results also indicate a difference in the roles these star players assume between college and professional basketball. In the NCAA, the strongest players receive the majority of passes, while in the NBA they make the majority of passes. Knowing that these particular strategies' efficacy differ depending on the level of professionalism can help coaches pick the most effective style of play for their own team's setting.

Because reciprocity improves a team's success, coaches should foster more of it among teammates. Creating a less selfish team atmosphere through both formal plays and informal social ties between players should increase trust and better performance during a match. And again, sharing more passes between players makes it harder for the opposing team to successfully defend against a star player while creating more opportunities for one's own team to score. Additionally, these results can be applied to areas outside of sports, such as businesses and other task-oriented groups, community networks, and more. If increasing reciprocity in these other types of networks also leads to improved efficacy, this would be useful information for managers or community leaders, etc., who wish to increase their networks' success. Further research on how to raise a network's reciprocity levels would be very beneficial.

In general, there is much we can learn from comparing sports networks' metrics to other types of networks, in both directions. For example, we can apply our results from the sports realm to business networks by asking if the strict rules of organized sports allow networks to benefit more from unpredictability in edge formation than less formally-organized networks of coworkers in a company. Or conversely, we can ask if homophily, which has been found to foster preferential tie-formation between individuals in many cases (McPherson, Smith-Lovin, & Cook, 2001), is also in play between teammates. Do players tend to pass to those who are similar in age, race, team tenure, or status, to themselves? If so, could this affect the success of a team through personal biases? These cross-disciplinary questions can lead us to a greater understanding of how networks function under a wide range of circumstances and settings, and be applied across many areas of study.

4.7 FUTURE DIRECTIONS

As stated above, subsequent studies should run separate analyses on women's and men's games while holding the level of professionalism constant, or vice versa, to isolate gender or play level differences. Doing so may reveal whether the success of various types of networks is affected by these factors, which in turn could prove useful to coaches and players in choosing which network patterns to apply to their own teams. Such analyses can be applied to settings outside of sports as well. For example, highlighting the differences between men's and women's teams' networks may contribute to research on gender differences in general. There is evidence that women and men structure their networks and partnerships differently. Szell and Thurner's (2012) study found that females are more prone to homophily (forming stronger ties with similar others) and form more partnerships, while men are heterophiles and have fewer partners, but these partners are more well-connected than women's partners. In the medical field, SNA has been used to explain sex differences in marijuana use (Wister & Avison, 2009). By understanding these underlying differences in male vs. female network formation, researchers can apply results to areas as varied as sports, sociology, and medicine. Looking further into whether professional basketball requires different network structures than college basketball is useful for coaches who wish to tailor their strategy to the type of team they are managing (professional, amateur, etc.). Given how the differences in resources, pay and skill levels, and prestige likely have an effect on how college basketball networks vary from their professional counterparts, other people in leadership positions would benefit from these analyses too – for example, business supervisors at different levels or companies (i.e. middle-management vs. senior position, startup vs. a long-established company, profit vs. non-profit).

Studying the success rate of different types of networks against varying defensive strategies employed by opposing teams would provide us with a more holistic view of the game, and further help coaches decide which plays to apply against specific opponents. Most if not all research into sports networks currently only focuses on a single team, without taking their opponents into consideration, but this is of course only half of the game. Players and coaches much adjust their strategies not only against different opponents, but also within a single game to adapt to the other team changing their own strategy during play. This is also true of many other goal competitive networks in the world; companies, for example, must be expected to track their rivals' strategies and respond accordingly. Thus, studying how not only how a network's structure affects its success internally but also how it responds externally to other competing networks can be widely applied.

Another question that should be investigated further is how well players' off-court social networks with their teammates map onto their in-game passing patterns. If a team's friendship network correlates significantly with its passing network, it may indicate preferential passing between close social associates, which of course would have an effect on the efficiency of a team's on-court behavior. Players could be biased toward passing to their friends or preferred social partners regardless of whether those individuals are in the best position to receive a pass. Or in the opposite direction, teammates' social interactions outside of a game may be influenced by who they pass to or train with the most while on the court. If so, this may further isolate cliques within the team from each other, which is detrimental to a cooperative undertaking such as team sports. Research suggests that informal relationships like friendship and communication lines affect group success as much as formal interactions (Shaw, 1964; Baldwin, 1997). If coaches want to improve passing efficiency between teammates on the court, they would be well-advised to increase affiliative interactions and foster friendships between their players off the court as well.

Gyamarti, Kwak, and Rodriguez (2014) were able identify unique passing styles in European men's soccer teams through analyzing flow motifs. This analysis should also be applied to basketball, which would allow us to identify whether the most successful teams possess their own unique passing structure, and whether this structure, if adopted by other teams, would also increase success. We can also use the information to look at whether a team's composition affects its passing styles, allowing coaches to choose the most effect passing style for their kind of team. Again, these passing styles can be matched up against defensive strategies employed by opponents as a test of their effectiveness under different circumstances, which would not only allow the offensive side to tailor their passing to combat their opponents' defense, but also let defending teams pick the best strategy or formation to use against rivals which rely predominantly on a few specific passing patterns. Applying SNA more fully within a single game by mapping networks for each possession can help with identifying unique passing styles and create a dynamic view of how passing behavior in the focal team changes throughout the game, allowing for a closer level of analysis.

Being able to factor in coaching effects in future studies would provide more information on the origins of a team's network structure. If teams through multiple seasons under the same coach retain similar network structures, we can conclude that coaches exert more of an influence on their teams' passing behaviors. However, if the networks seem to fluctuate with different compositions of players under the same coach, it's more likely that the players themselves are deciding how to pass during matches. Passing structure could also be affected by the type of match (a conference game vs. a championship, for example) a team is playing within a season, or even across multiple seasons, which should also be taken into account.

Finally as several studies (Sparrowe et al., 2001; Baldwin, Bedell, & Johnson, 1997) point out, individuals that are more central within their networks are more likely to be successful, compared to their peers on the peripheries of the network. We did not find any relation between players with high degree centrality and how many points or assists they scored per game, but it's possible there are other measures of success that are affected by centrality. Are central players who receive or make the most passes more likely to be rewarded with raises, sponsorships, or awards, for instance? And if so, is it because they truly are more skilled, more visible since they have the ball so often, or both? These questions are important to answer from a management point of view, as it allows owners and organizations to better identify who the most successful players on a team are, and how to best allocate their resources among all their players.

4.8 TABLES AND FIGURES

	Estimate	*p*-value	Odds ratio
Intercept	2.317e+00	0.1884	10.14519
Out-Degree Centralization	-6.684e+00	0.0227 *	0.001250765
RPI	2.942e-03	0.0517	1.002946
Total Centralization	4.726e+00	0.0663	112.8433
Individual Centrality	4.057e-02	0.9038	1.041404
Total Degree	6.714e-05	0.9542	1.000067

Table 4.1: Logistic regression of 2006-2007 University of Washington games

	Estimate	*p*-value	Odds ratio
In-degree Centralization	-0.211142	0.17171	0.80965940
Out-degree Centralization	0.503957	0.00425 **	1.65525840
Total degree Centralization	-0.357718	0.05482	0.69926993
Total Degree	0.062946	< 2e-16 ***	1.06496970
Reciprocity	0.471769	7.88e-06 ***	1.60282724
Centrality	-0.112272	0.00350 **	0.89380092

Table 4.2: NBA between-team GLM results

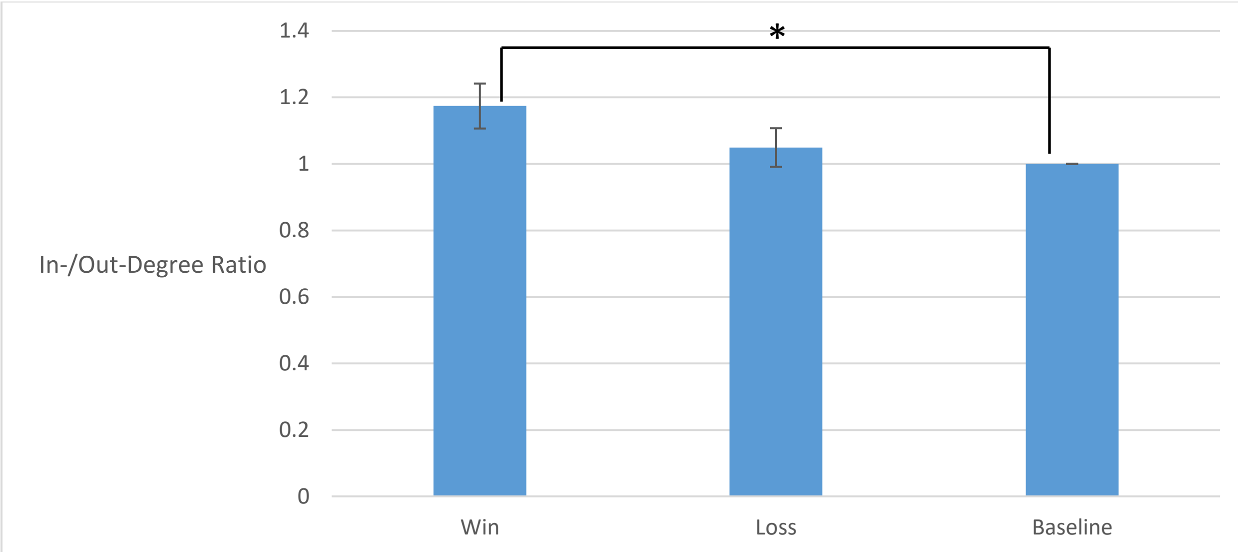

Figure 4.1: Average in-degree centralization/out-degree centralization ratios for winning and losing NCAA teams, compared to baseline (1:1). Error bars represent +/- SE.

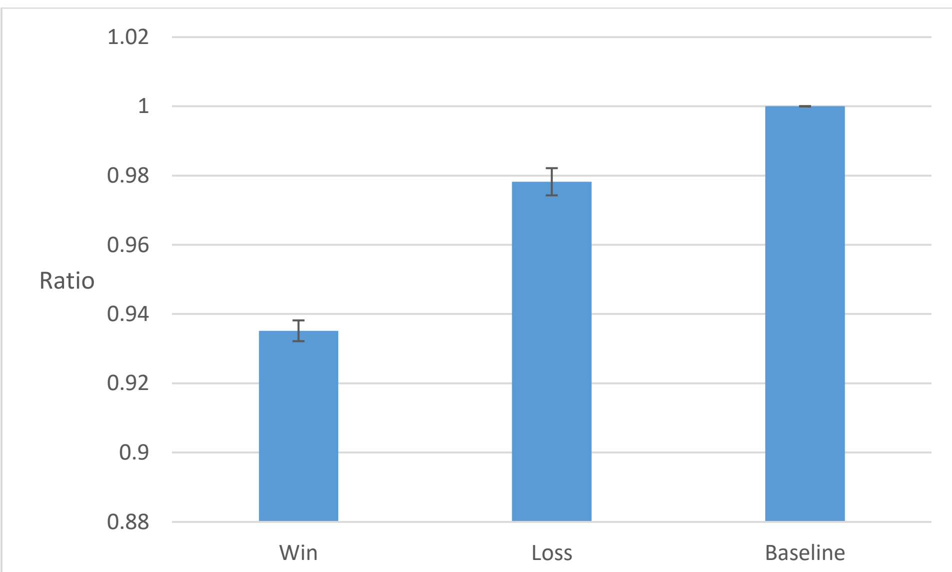

Figure 4.2: In-/out-degree centralization ratio for winning and losing NBA teams compared to baseline

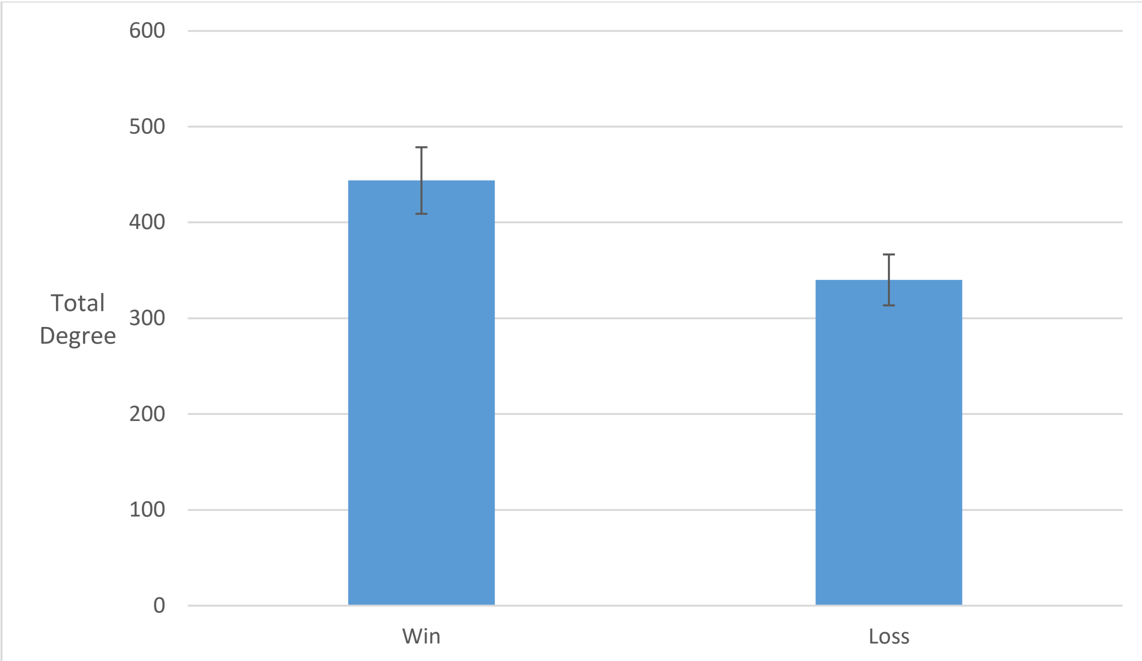

Figure 4.3: Total degrees for UW wins and losses

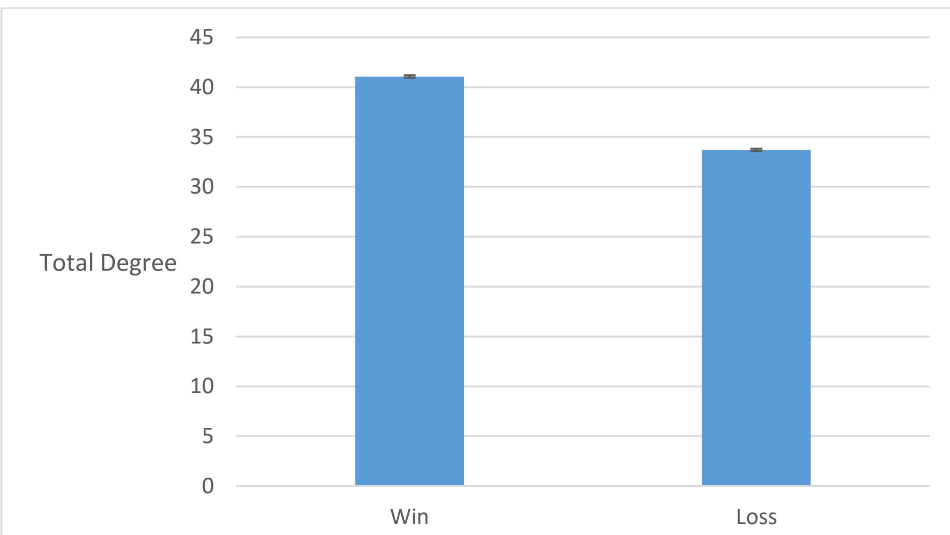

Figure 4.4: Total degrees for NBA wins and losses

REFERENCES

Apicella, C. L., Marlowe, F. W., Fowler, J. H., & Christakis, N. A. (2012). Social networks and cooperation in hunter-gatherers. *Nature*, *481*(7382), 497-501.

Baldwin, T. T., Bedell, M. D., & Johnson, J. L. (1997). The social fabric of a team-based MBA program: Network effects on student satisfaction and performance. *Academy of Management Journal*, *40*(6), 1369-1397.

Bavelas, A. (1948). A mathematical model for group structures. *Human organization*, *7*(3), 16-30.

Bavelas, A. (1950). Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America*, *22*(6), 725-730.

Bavelas, A., & Barrett, M. (1951). An experimental approach to organisational communication. *Personnel*, *27*, 386-397.

Bienaymé, I. J. (1845). De la loi de multiplication et de la durée des familles. *Soc. Philomat. Paris Extraits, Sér*, *5*, 37-39.

Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure* (Vol. 7). New York: Free Press.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (1992). UCINET IV version 1.0 reference manual. Columbia, SC: Analytic Technologies.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(5916), 892-895.

Bourbousson, J., Poizat, G., Saury, J., & Seve, C. (2010). Team coordination in basketball: Description of the cognitive connections among teammates. *Journal of Applied Sport Psychology*, *22*, 150-166.

Bowles, S., & Gintis, H. (2004). The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical Population Biology*, *65*(1), 17-28.

Brass, D. J. (1981). Structural relationships, job characteristics, and worker satisfaction and performance. *Administrative Science Quarterly*, 331-348.

Carolan, B. V. (2014). Social network analysis and education: theory, methods & applications. Los Angeles: Sage.

Cartwright, D., & Harary, F. (1956). Structural balance: a generalization of Heider's theory. *Psychological review*, *63*(5), 277.

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, *19*(22), 3127-3131.

Clemente, F. M., Couceiro M. S., Martins F. M. L., & Mendes, R. S. (2015). Using network metrics in soccer: A macro-analysis. *Journal of Human Kinetics*, *45*, 7-10. doi: 10.1515/hukin-2015-0013

Coleman, J., Katz, E., & Menzel, H. (1957). The diffusion of an innovation among physicians. *Sociometry*, *20*(4), 253-270.

Cooley Charles, H. (1909). Social organization: A study of the larger mind. *Charles Scribner's Sons, New York*.

Cross, R., Borgatti, S. P., & Parker, A. (2002). Making invisible work visible: Using social network analysis to support strategic collaboration. *California Management Review*, *44*(2), 25-46.

Davis, J. A. (1967). Clustering and structural balance in graphs. *Human relations*, *20*(2), 181-187.

de Sola Pool, I., & Kochen, M. (1978). Contacts and influence. *Social networks*, *1*(1), 5-51.

Deutsch, K. W. (1953). *Nationalism and social communication: An inquiry into the foundations of nationality* (pp. 209-222). Cambridge, MA: MIT press.

Durkheim, E. (1893). 1984. *The division of labor in society*.

Durkheim, E., & Halls, W. D. (1984). *The division of labor in society*. New York: Free Press.

Edwards, G. (2010). *Mixed methods approaches to social networks analysis. ESRC National Centre for Research Methods*. NCRM/015.

Emerson, R. M. (1976). Social exchange theory. *Annual review of sociology*, *2*(1), 335-362.

Eubank, S., Guclu, H., Kumar, V. A., Marathe, M. V., Srinivasan, A., Toroczkai, Z., & Wang, N. (2004). Modelling disease outbreaks in realistic urban social networks. *Nature*, *429*(6988), 180-184.

Fagiolo, G., Reyes, J., & Schiavo, S. (2008). On the topological properties of the world trade web: A weighted network analysis. *Physica A: Statistical Mechanics and its Applications*, *387*(15), 3868-3873.

Fewell, J., Armbruster, D., Ingraham, J., Petersen, A., Waters, J. (2012). Basketball teams as strategic networks. *PLoS One*, *7*, 1-9. doi: 10.1371/journal.pone.0047445

Flament, C. (1963). *Applications of graph theory to group structure*. Englewood Cliffs, NJ: Prentice-Hall.

Fonti, F., & Maoret, M. (2015). The direct and indirect effects of core and peripheral social capital on organizational performance. *Strategic Management Journal*, *37*(8), 1765-1786.

Freeman, L. C. (1989). Social networks and the structure experiment. *Research methods in social network analysis*, 11-40.

Freeman, L. C. (2004). The development of social network analysis. Vancouver, BC, Canada: Empirical Press.

Garlaschelli, D., & Loffredo, M. I. (2004). Patterns of link reciprocity in directed networks. *Physical Review Letters*, *93*(26), 268701.

Grund, T. U. (2012). Network structure and team performance: The case of English Premier League soccer teams. *Social Networks, 34*, 682-690.

Guidetti, M., Cavazza, N., & Graziani, A. R. (2016). Perceived disagreement and heterogeneity in social networks: distinct effects on political participation. *The Journal of Social Psychology, 156*(2), 222-242.

Guimera, R., & Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature*, *433*(7028), 895-900.

Gyarmati, L., Kwak, H., & Rodriguez, P. (2014). Searching for a unique style in soccer. *arXiv preprint arXiv:1409.0308*.

Heider, F. (1946) Attitudes and cognitive organization. *Journal of Psychology*, *21*, 107–112. http://dx.doi.org/10.1080/00223980.1946.9917275

Heider, F. (1958) The Psychology of Interpersonal Relations. New York: Wiley. http://dx.doi.org/10.1037/10628-000

Hidalgo, C. A. (2016). Disconnected, fragmented, or united? a trans-disciplinary review of network science. *Applied Network Science, 1*(1), 6.

Hobson, J. A. (1919). *The evolution of modern capitalism: a study of machine production*. The Walter Scott publishing co., ltd.

Hoffman, J. (2006). *Norms for fitness, performance, and health*. Human Kinetics.

Jamali, M., & Abolhassani, H. (2006). Different aspects of social network analysis. *Proceedings of the 2006 IEEE/WIC/ACM International Conference on* Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. *Journal of Abnormal and Social Psychology*, *46*, 38-50.

Web Intelligence, 66–72.

Kadushin, C. (1966). The friends and supporters of psychotherapy: on social circles in urban life. *American Sociological Review*, 786-802.

Katz, L. (1947). On the matric analysis of sociometric data. *Sociometry*, *10*(3), 233-241.

Kooij, R., Jamakovic, A., van Kesteren, F., de Koning, T., Theisler, I., & Veldhoven, P. (2009). The Dutch soccer team as a social network. *Connections*, *29*, 4-14.

Krause, J., James, R., Franks, D. W., & Croft, D. P. (Eds.). (2015). *Animal social networks*. Oxford University Press, USA.

Laumann, E. O., Gagnon, J. H., & Michael, R. T. R. & Michaels, S.(1994). *The social organization of sexuality*.

Le Bon, G. (1897). *The crowd: A study of the popular mind*. Fischer.

Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. *Journal of Abnormal and Social Psychology*, *46*, 38-50.

Leider, S., Möbius, M. M., Rosenblat, T., & Do, Q. A. (2009). Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics*, *124*(4), 1815-1851.

Lévi-Strauss, C. (1949). 1969. *The elementary structures of kinship*.

Loomis, C. P. (1946). Political and occupational cleavages in a Hanoverian village, Germany: A sociometric study. *Sociometry*, *9*(4), 316-333.

Lorrain, F., & White, H. C. (1971). Structural equivalence of individuals in social networks. *The Journal of mathematical sociology*, *1*(1), 49-80.

Luce, R. D., & Perry, A. D. (1949). A method of matrix analysis of group structure. *Psychometrika*, *14*(2), 95-116.

Macfarlane, A. (1883). Analysis of relationships of consanguinity and affinity. *The Journal of the Anthropological Institute of Great Britain and Ireland*, *12*, 46-63.

Malinowski, B. (1922). Argonauts of the Western Pacific: An account of native enterprise and adventure in Melanesian New Guinea.

Martineau, H. (1985). The Positive Philosophy of Auguste Comte. Kitchner, Ontario: Batoche.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, *27*(1), 415-444.

Molm, L. D. (1994). Dependence and risk: Transforming the structure of social exchange. *Social Psychology Quarterly*, 163-176.

Moreno, J. L., & Jennings, H. H. (1934). *Who shall survive?* (Vol. 58). Washington, DC: Nervous and mental disease publishing company.

Moreno, J. L., & Jennings, H. H. (1938). Statistics of social configurations. *Sociometry*, 342-374.

Morgan, L. H. (1851). *League of the Ho-de-no-sau-nee, Or*. Sage & Brother.

Moynihan, D. P., & Pandey, S. K. (2007). The ties that bind: Social networks, person-organization value fit, and turnover intention. *Journal of Public Administration Research and Theory*, *18*(2), 205-227.

Ohtsuki, H., Hauert, C., Lieberman, E., & Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, *441*(7092), 502-505.

Patnoe, S. (1988). *A narrative history of experimental social psychology: The Lewin tradition*. Springer Science & Business Media.

Pena, J. L., & Touchette, H. (2012). A network theory analysis of football strategies. *arXiv preprint arXiv:1206.6904*.

Radcliffe-Brown, A. R. (1957). *A natural science of society*. Free Press of Glencoe.

Rand, D. G., Arbesman, S., & Christakis, N. A. (2011). Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences*, *108*(48), 19193-19198.

Rose, T. (2016). *The end of average: How to succeed in a world that values sameness*. Penguin UK.

Scott, J. (2012). *Social network analysis*. Sage.

Serrano, M. A., & Boguná, M. (2003). Topology of the world trade web. *Physical Review E*, *68*(1), 015101.

Shaw, M. E. (1964). Communication networks. *Advances in experimental social psychology*, *1*, 111-147.

Simmel, G. (1908). Domination. *Georg Simmel on Individuality and social forms. Selected writings. The University of Chicago Press: Chicago*, 96-120.

Sparrowe, R. T., Liden, R. C., Wayne, S. J., & Kraimer, M. L. (2001). Social networks and the performance of individuals and groups. *Academy of management journal*, *44*(2), 316-325.

Spencer, H. (1897). *Principles of sociology* (Vol. 2). D. Appleton and Company.

Szell, M., & Thurner, S. (2013). How women organize social networks different from men. *Scientific reports*, *3*.

Tichy, N. M., Tushman, M. L., & Fombrun, C. (1979). Social network analysis for organizations. *Academy of management review*, *4*(4), 507-519.

Tönnies, F. (1855). *Community and civil society*. Cambridge Univ Pr.

Travers, J., & Milgram, S. (1967). The small world problem. *Psychology Today*, *1*, 61-67.

Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American journal of sociology*, *111*(2), 447-504.

Vaz de Melo, P. O. S., Almeida V. A. F., Loureiro A. A. F. (2008). Can complex network metrics predict the behavior of NBA teams? *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, 695-703. doi: 10.1145/1401890.1401974

Vendruscolo, M., Dokholyan, N. V., Paci, E., & Karplus, M. (2002). Small-world view of the amino acids that play a key role in protein folding. *Physical Review E*, *65*(6), 061910.

Wagner, J. A. (1995). Studies of individualism-collectivism: Effects on cooperation in groups. *Academy of Management Journal*, *38*(1), 152-173.

Watson, H. W., & Galton, F. (1875). On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, *4*, 138-144.

Wellman, B., Wong, R. Y. L., Tindall, D., & Nazer, N. (1997). A decade of network change: Turnover, persistence and stability in personal communities. *Social Networks*, *19*(1), 27-50.

Wey, T., Blumstein, D. T., Shen, W., & Jordán, F. (2008). Social network analysis of animal behaviour: a promising tool for the study of sociality. *Animal behaviour*, *75*(2), 333-344.

Wilson, David S. & Sober, Elliot (1996). (1994) Reintroducing group selection to the human behavioral sciences. BBS 17: 585-654. *Behavioral and Brain Sciences* 19 (4):777.

Wister, A. V., & Avison, W. R. (1982). “Friendly persuasion”: A social network analysis of sex differences in marijuana use. *International journal of the Addictions*, *17*(3), 523-541.