RECOMMENDATION SYSTEM FOR GOLF TRAINING

BLACK AMERICANS AND TRAUMATIC HEAD INJURIES

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Systems Engineering

By
Jacob Ziller

October 31, 2019

Technical Project Team Members
Alanna Flores
Orlando Jimenez
Christopher Kaylor
Kelly Rohrer

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: Jacob Ziller Date 10/31/2019

Approved: Catherine D. Baritaud Date: Dec. 5, 2019
Catherine D. Baritaud, STS Division, Department of Engineering and Society

Approved: William T. Scherer Date: 10/31/19
William T. Scherer, Department of Engineering Systems and Environment

According to John Meyer (2019) of *The Denver Post Online*, there are twenty-four million golfers, both recreational and professional, in the United States. This fact, coupled with the reality that golf has not changed much in the past hundred years, has led to golf becoming an oft-cited example of a sport where there is vast potential for growth in the amount of data analysis performed (Dusek, 2019). This does not mean that companies have not already begun to collect data on the game of golf, however. GameForge is the leading company in the golf data analysis industry, with roughly two thousand rounds being entered into their database every month, and ShotLink has been used by the Professional Golfers' Association (PGA) Tour to provide data on each tournament stroke since 2003 (Burke, 2012). The data is there, it just is not being used in any meaningful way at the moment.

Also, a recent trend is the seemingly sudden outbreak of a particular brain injury, chronic traumatic encephalopathy (CTE), in retired athletes. This disease overwhelmingly affects professional football players, and as such former National Football League (NFL) players have been suing the league in recent years. This outbreak was so widespread that in 2018 the NFL paid former players who suffered traumatic brain injuries over a billion dollars (Garcia-Roberts & Murphy). Currently, there is no cure for CTE, and symptoms take years to appear in patients (Bailes, Fitzsimmons, Hammers, & Omalu, 2010). Additionally, a recent survey of 50,000 middle and high school-aged children done by Monitoring The Future (Hoyle, 2019) found that tackle football is mainly played by black children. This creates a troubling societal issue; one which a particular racial group is more prone to a traumatic injury than other racial groups. The Science, Technology, and Society (STS) research topic will look at how concussions affect black families differently than families of other races.

These two topics both deal with sports, but the STS topic does not deal directly with the technical topic. Because of this, the topics are loosely coupled. The technical research will draw from GameForge's expertise gathered in weekly meetings with Brian Bailie and Mark Sweeney, of GameForge, in order to custom-tailor a training recommendation based on practice drill and golf round performance. The STS topic will draw mainly from sociological journals concerning the overrepresentation of minorities in the NFL and the adverse effects CTE has on specific racial groups, as well as legal articles concerning the settlements made with respect to traumatic brain injury. These two topics are of great importance in the world today, albeit for very different reasons. Current-day clients require consultant-given recommendations to be backed by data, and this project is no different. It requires data analysis, problem solving, and computer programming skills to be able to create a final deliverable. All of these skills come together in the systems thought process and in the careers that most engineers will graduate into. That is the motivation behind the research in the golf analytics project; combining two industries gaining momentum: golf and data science. In terms of the STS research project, the issue of racial inequality has been a topical subject in the news in the United States for the country's entire history.

There is plenty of research done in the concussions space, and plenty of research done in the racial inequality and sociology spaces, but the intersection of the two has been mostly vacant in recent years. This opportunity to do some new research in an interesting field was the motivation for the choosing of the STS topic. The technical project is slated to be concluded by the time the members of the capstone group graduate in spring of 2020. September 2019 was spent cleaning up the golf round, drill, and user data dump the capstone group received from GameForge into a digestible form so it was able to be analyzed. The capstone group spent the weekly meetings during this month mainly asking questions to Bailie and Sweeney, trying to

make sense of the data and deliver basic data walkthrough slides to the client to ensure proper interpretation of the data. The next two months, October and November 2019, will be spent on accomplishing two main goals: clustering equal-scoring players based on play type and determining the amount of time that can pass after doing a drill until that drill no longer has any effect on round scoring. Once we get those two tasks done, the rest of the year will be focused on designing a recommendation system that has a few qualities. First, it should have the ability to determine what skills are most important in scoring, as well as the ability to rank the skills themselves. Next, it should be able to take the clusters we come up with using the K-means process and give each cluster a unique recommendation on how to improve. Finally, it should be able to forecast into the future what a golfer's progress should look like based on where they start and what drills GameForge recommends to them.

PERFORMANCE-BASED GOLF TRAINING RECOMMENDATION SYSTEM

The technical project seeks to design and implement, into the GameForge product, a training recommendation system which considers round scoring data, training drill performance data, and golf course trait data to provide both a general training regimen as well as a tournament-specific regimen. If a golfer has an upcoming tournament at a particularly hard course, the training system should take that into account and recommend drills specifically tailored for that course. According to Forbes' Mike Buteau (2016), almost eighty percent of golfers who consider score important have never taken a lesson. This is proof enough that, of the twenty-four million golfers currently in the United States (Meyer, 2019), there is a market out there for people who have not taken lessons but want to improve their golf skills in an alternative

way to lessons. The golf data analysis market is growing, and GameForge actively capturing this market.

This project will be undertaken as part of the University of Virginia Systems Engineering two-semester capstone experience. By the end of the academic year, the capstone group is expected to create a fully implementable drill recommendation system designed for golfers in the GameForge system to use, as well as to write a paper and give a presentation in the annual Systems & Information Design Symposium (SIEDS) conference. All of the capstone group's work has been done on personal computers, using R and Microsoft Excel. The project personnel include students, University faculty, clients, and a sponsor. The students involved are all fourth-year Systems engineers: Alanna Flores, Orlando Jimenez, Christopher Kaylor, Kelly Rohrer, and Jacob Ziller. The capstone group also is assisted by Ben Lenox, local high school student, who has been doing much of the coding in the language R that the group had been using at the beginning. In terms of faculty, Professor William T. Scherer and Senior Research Scientist Stephen C. Adams, both of the Systems Engineering department, are the advisors the capstone group turn to when looking for expertise in the data side of the project.

When looking for expertise in the golf subject matter side of the project, the capstone team turns to the clients. Brian Bailie, formerly a Women's Golf Coach at the University of Virginia, and Mark Sweeney, founder of AimPoint Golf, are both currently involved with GameForge. Any questions about golf, the GameForge website, or general steps to take when creating the training recommendation system are answered by Bailie and Sweeney. Finally, there is the financial sponsor, the consulting firm Ankura. While the capstone group does not expect any additional costs to be incurred, if there are any Ankura will cover them. Matt Burkett is a Senior Director at Ankura, and as a former student of Professor Scherer, has been coming to

select meetings on the behalf of Ankura to offer a different perspective on our research process and presentations.

The capstone group takes a statistics-based approach in creating this personalized training recommendation system. Mark Broadie is a well-respected name in the golf analytics industry due to his invention of the "strokes gained" group of statistics. In a 2018 interview with Walter Lis, of the *Chicago Golf Report*, he pointed out that in order to make any improvement in golf training, a golfer first needs to know what he needs to improve. In other words, look at statistics to show what weaknesses exist, and work to improve on said weaknesses. Whether strokes gained or GameForge's proprietary statistics, a golfer needs to have some report to show to a coach to allow for specialized training. An example of a relevant for improving scoring by using GameForge drills is a player's scoring performance on holes with a par value of five, displayed in Figure 1 (pg. 6). Since par five performance is so critical to overall round scoring, if a player struggles to make below par consistently on these holes in tournaments, GameForge suggests drills to mitigate that inefficiency. ShotLink, an industry leader in real-time, on course golf statistics for thousands of tournaments, both PGA and non-PGA (Burke, 2012), is a natural source GameForge pulls from when gathering shot data. Another benefit of using ShotLink data is that its database gets richer every week, so while there is not too much usable data in the ShotLink database now, that will not be the case for very much longer (Dusek, 2018). The capstone group solution will improve on existing training recommendation systems because of the personalization of the GameForge platform. The ability to tweak, in the middle of a golf tournament, a player's training regime based on how the golfer performed in the tournament is not offered in the space today.

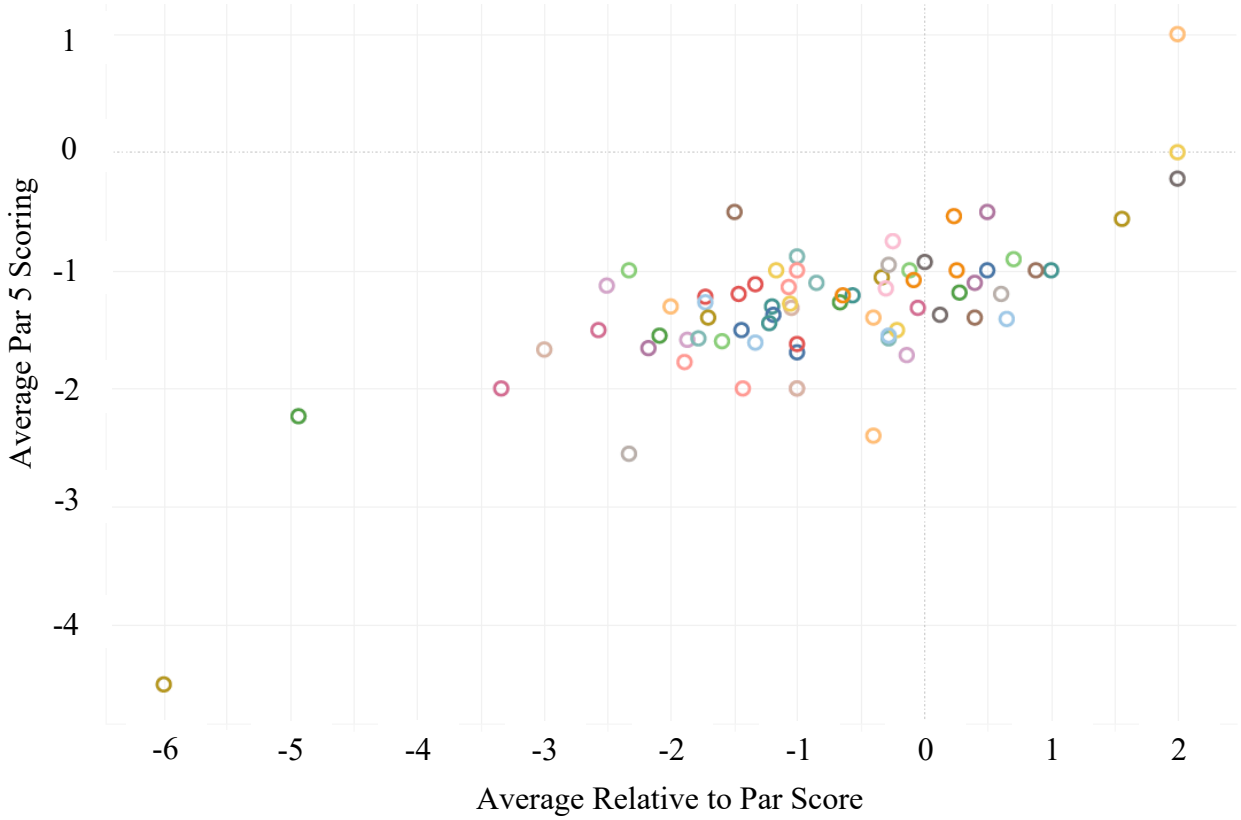

Figure 1: Par 5 Scoring Criticality: There is a strong positive correlation between how well, on average, a golfer performs on holes with a par value of five and how well that same golfer performs, on average, in the whole round. This is an example of a trend that GameForge can take advantage of when suggesting drills to a golfer in the system (Ziller, 2019).

A conclusion the capstone group has come to so far in research is that there really are two ways to be a good golfer: hit greens at an impressive rate or be a great scrambler. Scrambling can best be summed up as making par when the golfer fails to hit a green. The capstone group's K-means testing, a data science method intended to break data into discrete subsamples with similar traits, came up with only two clusters, and this was confirmed by Michael Fry and Jeffrey Ohlmann in their 2012 *INFORMS Journal on Applied Analytics* article. Fry and Ohlmann (2012) point out that tournament winners on the PGA tour have similar strokes gained characteristics week after week. In other words, they fall into one of the two clusters the capstone group laid

out. The ideal scenario for the capstone group is to be able to provide distinct training recommendations for each type of player, based on that cluster's relative strengths and weaknesses, as well as provide even more tailored recommendations for the various types of players that may exist within each cluster.

SPORTS SOCIOLOGY: HEAD INJURIES AND THEIR RACIAL EFFECTS

Among all CTE sufferers, there is a preponderance of former pro football players, of which seventy percent are black (Sonnad, 2018). Figure 2 provides a breakdown, by race, of the NFL, but it is clear that black men over index in pro football, as only six percent of the United States is made up of black males (Moore, 2015). Black men are overrepresented in the NFL for a number of reasons. First, low-income parents use sports, namely football, as an outlet to prevent their children from joining gangs while growing up (Hoyle, 2019). According to the United States Census Bureau (2019), race and low-income status are correlated, with blacks earning significantly lower median household incomes.

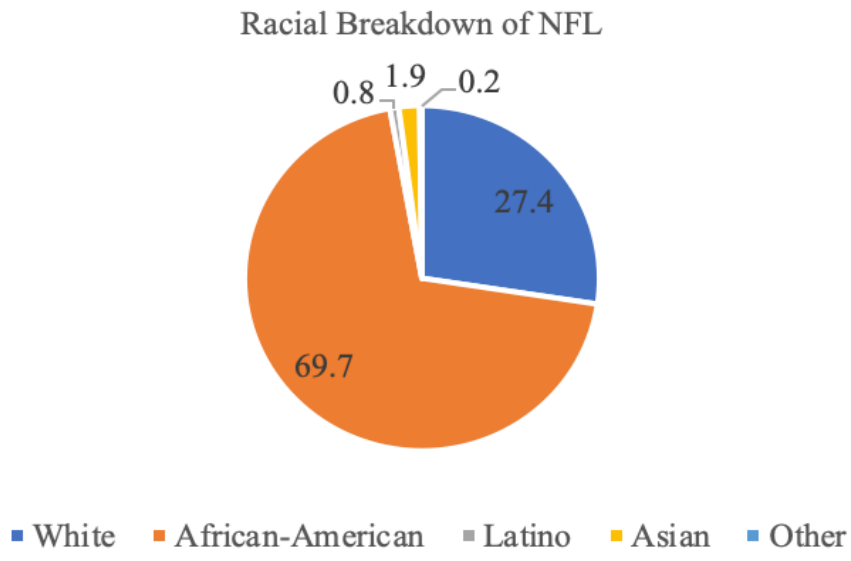

Figure 2: Racial Breakdown of NFL: This chart shows the stark black overrepresentation in the NFL. (Ziller, 2019).

Therefore, black parents participate in this activity at a higher rate than parents of other races. Also, in her 2019 article "The White Flight From Football", Alana Semuels shows that for many low-income families, football can be the only way to be admitted into a college. It is clear that depending on where a child grows up, his or her susceptibleness to CTE can vary dramatically (Semuels, 2019). Even among football-playing children, those in low income cities are often not fortunate enough to own helmets with the latest technology, much of it designed to prevent concussions and by extension CTE.

There is even evidence that within football there is a race problem, though. The positions most affected by concussions are played overwhelmingly by black athletes, reinforcing a narrative that black men put their whole life into a game that does not give them anything in return (Moore, 2015). In his book *Social Issues in Sport*, Ron Woods (2016) outlines the steady decline in participation that football has been experiencing in relation to other sports popular among youths. Interestingly, it seems that head injuries are a chief reason for this, as shown by youth leagues cutting down on the amount of tackling allowed in practice in response to the declined participation rates (Woods, 2016). All of these facts explain that, while football is still popular among children, especially those who grow up in low-income households and those who are black, there have been some negative stories written about football in the past few years and participation will continue to decline if nothing is done to combat the recent outbreak of head injuries.

Research on this topic will be conducted primarily by Jacob Ziller, with oversight by his STS advisor Professor Catherine D. Baritaud of the STS division of the Department of Engineering and Society. This problem needs to be explored more because of the sheer number of parties that have stock in a child's future. While it may not seem like it at the time, every

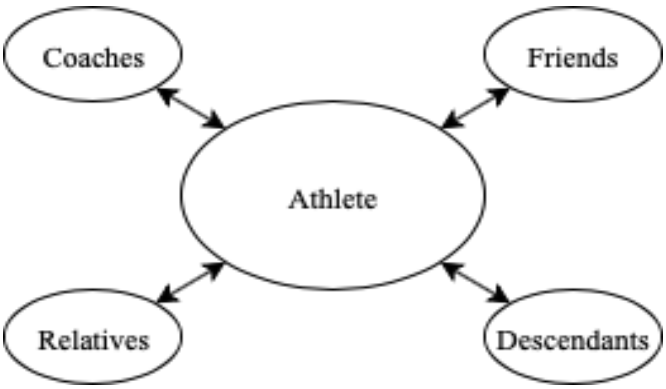

Figure 3: Effect of an Athlete's Decisions on Others: An athlete's decision to continue playing a sport that causes harm to his or her brain has a trickle-down effect on many other parties. Similar to Carlson's Social Relationships model (Ziller, adapted from Carlson, 2019).

decision a person makes effects multiple other people down the line. Figure 3 shows a few of the parties who would take interest in an athlete making the choice either to play or not to play football. This graphic is similar to a common STS framework, W. Bernard Carlson's Social Relationships model, adapted from the Social Construction of Technology model. The relationships are two-way, meaning that the athlete has an influence on each of those parties, but each of those parties has an influence on the athlete as well. For example, a relative of an athlete growing up at a lower socioeconomic level, such as a parent, may try to push an athlete to choose football in order to dissuade them from joining gangs, as Hoyle (2019) pointed out. On the other side, though, the athlete may want to pursue football so they can provide for their family on the rare case that they make the NFL. Each of these relationships has a two-way nature to it, which adds a level of complexity to the problem. For this reason, this problem of the sociological impact of head injuries is an interesting one for an STS topic.

Adding to the intrigue of this topic is that experts in this field do not agree on how to gather data to support a hypothesis. An example of this is how to represent socioeconomic status in an experiment. In their 2017 experiment, researchers Thomas Aicher, Trevor Bopp, Robert Turick, and Joshua Vadeboncoeur examine, through a survey of 451 college students, fourteen sports and the differences in racial makeup of youth participation in each sport. They assume that

race is a fine proxy for socioeconomic status. Just a year before, researchers Asken, et al. (2016) wrote about what makes one athlete more susceptible to CTE than another. In contrast to Aicher, et al., these authors caution against using simply race to mean socioeconomic status. Rather, they argue that the developmental environment a youth is in matters much more than the race they belong to (Asken, et al., 2016). All of these researchers are credible and affiliated with universities, they just have differing opinions on this topic. Therefore, the objective of this research is to be able to draw things from all sides, for example, both Aicher, et al. and Asken, et al., and synthesize findings to present a coherent argument for the true effect that head injuries and CTE has on different races in an academic paper. In conclusion, a combination of academic, social, and technological factors has led to black Americans getting CTE at a higher rate, while the football participation rate for that subset of Americans is not declining at the same rate as other athletes' rates.

WORKS CITED

Aicher, T., Bopp, T., Turick, R., & Vadeboncoeur, J. (2017). Are you welcomed? A racial and ethnic comparison of perceived welcomeness in sport participation. *International Journal of Exercise Science*, *10*(6), 833-844.

Asken, B., Sullan, M., Snyder, A., Houck, Z., Bryant, V., Hizel, L., … Bauer, R. (2016). Factors influencing clinical correlates of chronic traumatic encephalopathy (CTE): a review. *Neuropsychology Review I, 26*(4), 340-363.

Bailes, J., Fitzsimmons, R., Hammers, J., Omalu, B. (2010). Chronic traumatic encephalopathy, suicides and parasuicides in professional american athletes: the role of the forensic pathologist. *The American Journal of Forensic Medicine and Pathology*, *31*(2), 130-132. (2010, June).

Broadie, M. (2018, October 16). Mark Broadie – the godfather of golf analytics [Interview by Walter Lis]. *Chicago Golf Report*. Retrieved from https://www.chicagogolfreport.com/

Burke, M. (2012, April 5). ShotLink is making golf easier for hacks and harder for pros. *Forbes.* Retrieved from https://www.forbes.com/

Buteau, M. (2016, January 27). The solution to golf's stalled participation might be found in a lesson. *Forbes.* Retrieved from https://www.forbes.com/

Dusek, D. (2018, March 4). By the numbers: analytics become an increasingly important part of golf. *Golfweek Digital Edition*. Retrieved from https://golfweek.com/

Fry, M. & Ohlmann, J. (2012, April 1). Introduction to the special issue on analytics in sports, part I: general sports applications. *Informs Journal on Applied Analytics*, *42*(2), 105-228

Garcia-Roberts, G., Murphy, B. (2018, October 23). NFL players with brain trauma receive notice of settlements stripped to nothing. *USA Today*. Retrieved from https://www.usatoday.com/

Hoyle, B. (2019, February 6). *American football and CTE: is a racial divide inevitable?*. Retrieved from https://www.mdedge.com/psychiatry/article/194020/mixed-topics/american-football-and-cte-racial-divide-inevitable

Meyer, J. (2019, April 12). The state of golf in 2019 by the numbers. *The Denver Post*. Retrieved from https://www.denverpost.com/

Moore, A. (2015, December 24). *Football's war on the minds of black men*. Retrieved from https://www.vice.com/en_us/article/eze4gj/footballs-war-on-the-minds-of-black-men

Semuels, A. (2019, February 1). The white flight from football. *The Atlantic*. Retrieved from https://www.theatlantic.com/

Sonnad, N. (2018, May 24). The NFL's racial divide, in one chart. *Quartz*. Retrieved from https://www.qz.com

United States Census. (2019) *Current Population Survey.* Retrieved from https://www.census.gov/programs-surveys/cps.html

Woods, R. (2016). *Social Issues In Sport* (3rd ed.). Champaign, Illinois: Human Kinetics.

Ziller, J. (2019). *Par 5 scoring vs. round scoring.* [Figure 1]. *Prospectus* (Unpublished undergraduate thesis). School of Engineering and Applied Science, University of Virginia. Charlottesville, VA.

Ziller, J. (2019). *Racial breakdown of NFL.* [Figure 2]. *Prospectus* (Unpublished undergraduate thesis). School of Engineering and Applied Science, University of Virginia. Charlottesville, VA.

Ziller, J., Carlson, W.B. (2019). *Effect of an athlete's decision on others.* [Figure 3]. *Prospectus* (Unpublished undergraduate thesis). School of Engineering and Applied Science, University of Virginia. Charlottesville, VA. Adapted from Carlson.