An Interdisciplinary Approach to Sports Analytics in a University Setting
(Technical Paper)

The Ethics of Sports Analytics in Olympic Track and Field
(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Maryanna Lansing
Spring, 2020

Technical Project Team Members
Aniket Chandra
Jacqueline Hoege
Rishap Iyer
Rachel Kreitzer
Jacob Leonard
Ben Metzger
Sarah Nelson
Carl Rhodes
Daniel Ungerleider
Peter Worcester

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

Introduction

In the early 2000s, Oakland Athletics revolutionized the sport of baseball through the use of analytics, finding undervalued players who could win games. Through the application of analytics and uniquely developed variables, they went to the playoffs every year from 2000 to 2003 (Candaele and Dreier, 2019). Since then, sports analytics has been a growing industry that has not only shaped the sport of baseball, but throughout all sports, giving teams a competitive edge. Coaches and athletes look at the athlete's biometric data to improve on field performance, identifying both strengths and weaknesses to work on. Owners look at analytical and biometric models to calculate athlete's value and negotiate salaries (Candaele and Dreier, 2019).

Now, more than ever, professional teams and colleges are looking for ways to incorporate sports analytics into their programs. Professional teams are hiring analysts to develop and run statistical models in house or consulting out to analytics groups. The New England Patriots use the Kraft Analytics Group, owned by the Patriots owner Robert Kraft, as their "secret weapon" (Merrimack College, 2017). Colleges and universities have more options. The University of Southern California has a performance science institute that offers research and classes, "to help students, entrepreneurs and organizations compete and excel in any pursuit through the applied practices of performance science" (USC, 2019). The University of Michigan has implemented sports analytics through the Exercise and Sport Science Initiative (Manuel and Cunningham, 2019). Another option has been to develop an in-house sports analytics center with hired analysts. The technical proposal will address what is the best way to implement sports analytics into the UVA athletics program and more specifically the track and field team.

In implementing analytics, any program must protect the data of their athletes, using them for their best interest. Some athletes feel "constantly scrutinized" (Henne, 2017). While

others feel “that the additional surveillance reflects greater investment in them and their continued growth” (Henne, 2017). The STS proposal will address the ethics behind biometric research in professional athletes, specifically Olympic track athletes, by studying two cases: the Russian doping case, and the case of Caster Semenya.

Technical Topic (Capstone)

Sports and Performance Analytics Center

The 2018-2019 collegiate athletics season for the Virginia Cavaliers was considered one of the most successful years for any college sports program (McElroy, 2019). The Virginia Cavaliers were awarded “National Champion” status in two sports: NCAA Division 1 Men’s Basketball and Lacrosse. However, to continue to operate at the highest level, the Virginia Athletics department must adapt their approach to current trends. Bradley Smith, a sports performance analyst at Northwestern University, stated “I think that very soon we’re going to see that the teams that don’t use analytics — they’re going to be left behind” (2019). Many teams in the Virginia Athletics department that have engaged in data exploration and analysis seeking a competitive advantage. For example, engineering students developed analytics tools to support decision-making on and off the field for the Virginia Cavaliers football team (Hayhurst, 2019, para. 14). Other sports in the program have collected little or no data. As of now, all the teams that have integrated data analytics into their operations have done so independently of the Athletics department.

Ted White, the Assistant Athletic Director and the capstone project's sponsor, argued: "We need a structure that would allow all the teams to collect, process and analyze the data efficiently with dedicated resources to provide a competitive advantage" (2019). The accepted ideal framework for sport analytics is illustrated in Figure 1. This structure could be supported by in house analysts that work with each team or students having more opportunities to do research with each team.

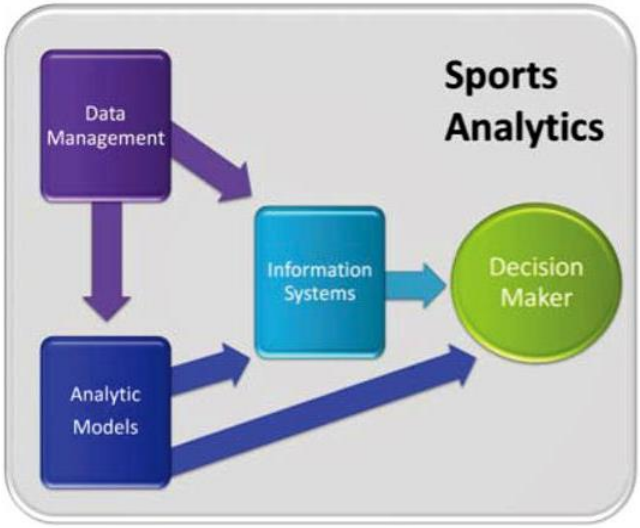

Figure 1: A framework of Sports Analytics (Alamar & Mehrotra, 2011)

One team in particular that is interested in incorporating sports analytics is the track and field team. So far, the Track and Field team has engaged in some injury analysis with a kinesiology grad student. This student performed a gait analysis on each member of the men's and women's cross-country team and has since been tracking distance, steps, and other variables each run. However, the team is looking to collect and incorporate data across the entire team to improve individual performance but also limit the number of injuries. There has been some research, but it is limited. For example, Yuanlong Liu and Robert W. Schutz, researchers in the School of Human Kinetics at the University of British Columbia, have developed mathematical

models to evaluate and predict future world records in track and field (Yuanlong and Schutz, 1998).

The objective of this technical project is to design and develop a proposal for a Center of Performance Analytics Research at the University of Virginia, while also providing more specific information on how this center could be used by the track and field team. There has not been much research into modeling for track and field. However, with the sources that are available, the best way to implement programs from the potential center will be developed. This project is a two-semester long research capstone that will cover a range of areas starting with initial out scoping and ending with a presentation to the Office of the President. Several stakeholders will be interviewed and the project will research both established programs at U.Va's peer and competitor schools. Some theoretical modeling will be performed for the track and field team.

While working with the Athletics Department on designing the sports analytics aspect of the project, the team will collaborate with pertinent academic departments on their engagement and their ideas. As with most academic and athletic initiatives, funding for the project is a critical factor in development (Groves, 2019). The various fundraising mechanisms the University already has at its disposal are being explored, including the donation programs and alumni outreach. The final deliverables will be a group paper in the form of a proposal, with the findings of a feasibility study.

STS Topic

The Ethics of Sports Analytics in Track and Field

As sports develop and athletes become more skilled and talented at all levels, teams, athletes, and owners look for those tools that give them the competitive edge. This need led to the developing business and research topic involving sports analytics. Some sports developed the analytics tool more than other sports. Still, it is rare to find a sport now without at least the beginnings of the use of sports analytics.

Researchers and teams are excited to research, collect data, and develop models. However, the consequences of collecting this data and developing these models have not always included the stakeholders affected including individuals' athletes, coaches, athletic trainers, team owners, nutritionists, hired sports analyst, teams as a whole, and fans. This proposal focuses on athletes and the effect of using biometric data relative to them.

The ethics of collecting biometric data and use of sports analytics needs further research. The result should balance an athlete's benefit and their privacy protected. The paper studies Olympic track and field athletes specifically, the two examples used to better understand the ethics and legality of collecting biometric data from athletes are: (1) doping violations in Russia and (2) the case of Caster Semenya. This project will develop over two-semesters through research ending with a formal paper.

In 2016, the Russian track athletes engaged in a massive doping scheme, causing them to be banned from competing in the Rio Olympics (Zorthian, 2016). The Russian national government had a whole division that implemented and enforced a doping strategy for its Olympic athletes, "These include micro dosing small amounts of erythropoietin, employing novel synthetic anabolic steroids not known to testing agencies, and simply avoiding out-of-

competition drug testing through inaccurate disclosure of their location. Other athletes used autologous blood transfusions, which still cannot be directly identified in an antidoping test" (Gleaves, 2017). Could biometric data help to prevent doping and catch violations earlier? There are new biometric data methods under development to enforce the anti-doping policies. However, most of these methods provide suspicion not proof, "This is an unfair position for athletes, whose reputation is significant for deriving their living. They must defend themselves against shadows of suspicion without any real tools at their disposal (Gleaves, 2017)." Still, these methods could deter athletes from using steroids or catch athletes beating other methods of drug testing.

The other example is the case of Caster Semenya, who was banned from competing in international competitions because she has higher than normal testosterone levels (Longman, 2019). Her case is at the center of debate over sex testing, "Proponents argue that sex testing is a practical (though imperfect) way to ensure the safety, competitiveness, and fairness of women's sport, by "accurately" sorting competitors into appropriate categories according to convention, science, and humanitarian values" (Wells and Darnell, 2014). However, her case also involves whether this type of testing violates the athlete's right to privacy, and whether it is too complex to set the standards of what constitutes fair in each gender.

The main theories that will be used to support the research in the coming thesis will be actor network theory and technological determinism. Actor-network theory is the interactions between actors (who can be humans or non-humans), intermediaries, and the network in which they exist (Cressman, 2009). This theory addresses and explains the relationship between the developing technology, athletes, coaches, and owners. Critics say that this theory is too broad and hard to define, but knowing the network and connections clearly, it can be a useful theory.

For this proposal, the network is known and stated above. Technological determinism is based on the premise that society is shaped by technology. (Smith, 1994) Right now with new and developing technologies, sports are being shaped by these technologies and the data they provide, rather than, the sports shaping the technology.

Research Question and Methods

"What are the ethics behind sports analytics and biometric data collection in track and field, and how are these ethics established?"

The methods that will be used to research this issue include network analysis, interviews, and historical case studies. The network analysis method describes the hierarchy of the Olympic track and field teams, and who had access to athlete data and the interactions between those in this hierarchy. The members of the network that will be analyzed are coaches, athletes, and the sensors that they are using. The interviews provide a personal perspective of an athlete on the use of biometric data. The interviews predominantly seek to be on the professional runners (the Reebok group) and their perspective. Around 2 or 3 interviews should take place early Spring, 2020. The athletes will be asked what type of data collection they have already participated in and their thoughts on the use of their biometric data. The historical case studies will be collected using the two examples: the Russian doping incident and the case of Mokgadi Caster Semenya. They will be studied in chronological order: the doping incident first, then the case of Caster.

Conclusion

This capstone project will, through systems analysis, develop alternatives to best incorporate sports and performance analytics across the University. These alternatives will be

developed through research on the processes of UVA and those at other universities. The STS research paper will explore the ethics of collecting biometric data on Olympic track athletes and the incorporation of sports analytics in professional track and field. The paper seeks to identify and resolve ethical issues related to the collecting of athletes' biometric data. The paper will provide specific examples of instances in which biometric data can both help the sport and the athletes, and be destructive.

Bibliography

Armstrong, S., Jovanov, E., & Kerwin, D. (2007, April). Wireless connectivity for health and sports monitoring: a review. British Journal of Sports Medicine. Retrieved from https://bjsm.bmj.com/content/41/5/285

Cressman, D. (2009). A Brief Overview of Actor-Network Theory: Punctualization, Heterogeneous Engineering & Translation, 1–17.

Groves, A. (2019, October). Personal interview

Gleaves, J. (2017). Biometrics and Antidoping Enforcement in Professional Sport. *American Journal of Bioethics*, *17*(1), 77–79. https://doi.org/10.1080/15265161.2016.1251644

Hayhurst, C. (2019, August). Data Analytics Helps College Coaches and Athletes Optimize Training and Performance. *EdTech Magazine.* Retrieved from: https://www.richmond.com/

Henne, K. (2017). "I Felt Like a Lab Rat": The Importance of Power and Context in Understanding Biometric Technologies. *American Journal of Bioethics*, *17*(1), 63–65. https://doi.org/10.1080/15265161.2016.1251659

Is Data Analytics the Secret Weapon of the New England Patriots? Maybe. (2017, October 27). Retrieved from https://onlinedsa.merrimack.edu/new-england-patriots-data-analytics/.

Lindsey, G. R. (1959, April). Statistical Data Useful for the Operation of a Baseball Team. *Operations Research*, Vol. 7, 197-207.

Longman, J., & Macur, J. (2019, May 1). Caster Semenya Loses Case to Compete as a Woman in All Races. *New York Times*. Retrieved from https://www.nytimes.com/2019/05/01/sports/caster-semenya-loses.html

Manuel, W., & Cunningham, R. (n.d.). Transforming the world of exercise and sport. Retrieved

from https://essi.umich.edu/.

McElroy, W. (2019, June). UVA has had the best athletic year of any college sports program. *Richmond Times-Dispatch.* Retrieved from: https://www.richmond.com/

Roberts, J. L., Cohen, I. G., Deubert, C. R., & Lynch, H. F. (2017). The Legality of Biometric Screening of Professional Athletes. *American Journal of Bioethics*, *17*(1), 65–67. https://doi.org/10.1080/15265161.2016.1251647

Russell, J. S., & Browne, A. (2018). Performance-enhancing drugs as a collective action problem. *Journal of the Philosophy of Sport*, *45*(2), 109–127. https://doi.org/10.1080/00948705.2018.1440355

Scherer, W. (2019, October). Personal interview

Shuster, S., & Hartog, E. (2016). Can the Games Get Clean? *TIME Magazine*, *188*(6), 63. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=117072401&site=ehost-live&scope=site

Siemens, G., Dawson, S. & Lynch, G. (2013, December). Improving the Quality and Productivity of the Higher Education Sector. Society for Learning Analytics Research.

Smith, B. (2019, August). A Conversation with Chris Hayhurst. *EdTech Magazine*. Retrieved from: https://edtechmagazine.com/higher/

Smith, M. R. (1994). Technological Determinism in American Culture, 1–35.

The USC Performance Science Institute: USC Marshall. (n.d.). Retrieved from https://www.marshall.usc.edu/faculty-research/labs-and-institutes/usc-performance-science-institute.

University of Virginia: Office of Communications. (2019). *Great and Good: The 2030 Plan*.

Charlottesville, VA: Author.

Wells, C., & Darnell, S. C. (2014). Caster Semenya, Gender Verification and the Politics of Fairness in an Online Track & Field Community. *Sociology of Sport Journal*, *31*(1), 44–65. https://doi.org/10.1123/ssj.2012-0173

White, T. (2019, September). Personal interview

Yuanlong Liu, & Schutz, R. W. (1998). Prediction Models for Track and Field Performances. *Measurement in Physical Education & Exercise Science*, *2*(4), 205. https://doi.org/10.1207/s15327841mpee0204_2

Zorthian, J. (2016). Russian Doping. *TIME Magazine*, *188*(5), 16. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=116950092&site=ehost-live&scope=site