

Big Data and Analytics in Sport Management

Nicholas M. Watanabe and Stephen Shapiro

University of South Carolina

Joris Drayer

Temple University

Big data and analytics have become an essential component of organizational operations. The ability to collect and interpret significantly large data sets has provided a wealth of knowledge to guide decision makers in all facets of society. This is no different in sport management where big data has been used on and off the field to guide decision making across the industry. As big data evolves, there are concerns regarding the use of enhanced analytic techniques and their advancement of knowledge and theory. This special issue addresses these concerns by advancing our understanding of the use of big data in sport management research and how it can be used to further scholarship in the sport industry. The six articles in this special issue each play a role in advancing sport analytics theory, producing new knowledge, and developing new inquiries. The implications discussed in these articles provide a foundation for future research on this evolving area within the field of sport management.

Keywords: consumer behavior, econometrics, performance, social media, quantitative analysis

In 2003, Michael Lewis ([2003](#)) published *Moneyball: The Art of Winning an Unfair Game*, which forever changed the role of data in decision making in sport. However, the concepts presented in the book were not new, even in the context of baseball player evaluation. According to the Society for American Baseball Research, the creation of the box score in the mid-19th century was perhaps the first known example of systematically gathering player performance data. Since then, examples of the tracking of player performance data were common, culminating in the 1980s when Bill James began publishing a series of books on potential predictive power of certain player statistics. These books were largely credited with the creation of the field of sabermetrics, defined in 1980 by James as “the search for objective knowledge about baseball” ([Society for American baseball research, 2021](#)).

However, the phenomenon that followed “Moneyball” was not purely about sabermetrics. More broadly, the book was about identifying inefficiencies and trusting the objectivity of data. For the first time, a team was willing to transfer the balance of decision-making authority to data. Subsequently, advances in the collection, storage, transfer, and analysis of data have advanced the methods used in “Moneyball” to new heights and created more confidence in data’s ability to inform and predict. Indeed, big data analytics has created a surge in innovation throughout the sport industry, spawning many products and services that fundamentally changed how organizations operate.

A quick Internet search for “impact of Moneyball” yields numerous articles, blog posts, law reviews, podcasts, videos, and more about the influence of data-based decision making on business, law, criminal justice, and even faculty evaluation. Within the sport industry, the use of data in decision making is omnipresent with the bulk of attention on player performance systems such as those built by Catapult, Krossover, Second Spectrum, and others. The introduction of these systems has forever altered athletic training, game strategy, and player personnel management.

In much the same way that data have altered how sport organizations run their operations on the field/court/ice, they have also changed how organizations run the business behind the game. For example, Ticketmaster launched LiveAnalytics in 2011 with the aim of helping organizations in the live event industry better communicate with fans and, ultimately, to sell more tickets ([Ticketmaster, 2011](#)). Quantcast (San Francisco, CA) and other similar organizations also aid in identifying viable market segments for more efficient and effective marketing strategies. Salesforce is a popular internal tool used by many organizations to manage growing customer databases filled with data from a variety of sources, including social media. Although these products have a role enhancing profitability, academic research is needed to explore additional uses of data in effective segmentation, content creation, and consumer decision making.

We now live in a world where big data and analytics are pervasive within society ([George, Haas, & Pentland, 2014](#)). The diffusion of big data and analytics has impacted the operations and management of business organizations ([Colbert, Yee, & George, 2016](#)), transformed the measurement and methods used by marketers ([Erevelles, Fukawa, & Swayne, 2016](#)), shaped consumer purchasing decisions, and even changed the ways we communicate and live our daily lives ([Kitchin, 2014](#)). This is certainly the case in sport where professional and collegiate sport organizations have adopted the use of big data and analytics to improve performance and decision making on and off the field ([Fried & Mumcu, 2016](#); [Mondello & Kamke, 2014](#)). There is a natural alignment between sports and analytics as data have been a foundation for both the popularity and evolution of the industry. This growth of the use of big data and analytics by sport organizations has led to an increase in the number of academic programs creating degrees, courses, and even academic centers in this area ([Zaharia, Kaburakis, & Pierce, 2016](#)).

In this special issue, we address the evolution of big data and analytics in sport, highlighting concerns within this sector and providing articles that extend theory on big data while offering insight for future direction in sport. Within this special issue introduction, we present the importance of big data and analytics in the discipline and within the sport management literature, offer a

Watanabe and Shapiro are with the University of South Carolina, Columbia, SC, USA. Drayer is with Temple University, Philadelphia, PA, USA. Watanabe ([nmwata@mailbox.sc.edu](mailto:nmwatana@mailbox.sc.edu)) is corresponding author.

summary of each article included in the special issue, highlight directions for future research, and identify challenges and barriers in this diverse sector.

The Importance of Big Data to Sport Management

The term “big data” is often defined as a data set that is so large and complex that it cannot be managed or analyzed through traditional computing methods (Jin, Wah, Cheng, & Wang, 2015). However, management scholars (George et al., 2014) argue that the number of observations or the size of the data set should not be the focus of big data research. As such, when discussing big data, many refer to its unique characteristics in regard to volume, velocity, and variety (Zikopoulos & Eaton, 2011). Volume indicates the quantity of the data, velocity refers to the speed at which data are produced and collected, and variety suggests that the data can come from several different sources (McAfee & Brynjolfsson, 2017).

Although these characteristics can provide a better definition of big data, there are two specific concepts that emphasize the importance of big data for research in the social sciences. First, when discussing volume, the concept of data scope suggests that when a large volume of data is collected, it provides a more comprehensive manner through which to examine behaviors. As such, the scope of big data highlights its importance for building better understanding of phenomena as it allows the examination of populations rather than samples. Within the context of sport, data scope can be observed in studies of demand that utilize the full population of individuals who attend matches (e.g., Sung & Mills, 2018) or even the complete responses from social media users to marketing messages from sport organizations (e.g., Yan, Watanabe, Shapiro, Naraine, & Hull, 2019). Volume can also include longitudinal studies of fan or organizational behaviors, such as those using annual national surveys (Aizawa, Wu, Inoue, & Sato, 2018; Wicker & Frick, 2015).

Second, building upon the characteristics of volume and variety, data granularity is the ability to collect data with a greater level of detail and, thus, have units of analysis that are closer approximations to phenomena than traditional data sources (George, Osinga, Lavie, & Scott, 2016). In the realm of sport management, one example of the use of granular data can be found in social network analysis studies that measure the exact moments when individuals interact with one another on digital platforms (Abeza, O'Reilly, Séguin, & Nzindukiyimana, 2015; Yan, Pegoraro, & Watanabe, 2018).

In summary, big data and analytics have numerous applications and opportunities to be used in sport research, which will expand the theoretical and empirical understanding of phenomena and behaviors. However, the sport literature that uses big data to test theory and answer research questions is lacking, and the scholarly discussion on the theoretical nature of big data and the social, legal, and ethical concerns is extremely limited. Thus, the call for research in this area is warranted.

Summary of Big Data in Sport Literature and Special Issue Articles

Although there have been multiple articles that have used large data sets and have incorporated “big data” analytic techniques, the theoretical foundations of analytics and the practical implications

of its use within the context of sport have not been thoroughly investigated. Most of the articles in this area are published within industry outlets, discussing the increased use of analytics in sport and the need for more data-driven decision making within sport organizations (Szymanski, 2020).

The academic literature in this area has focused primarily on big data and on-field performance (Goes et al., 2020; Morgulev, Azar, & Lidor, 2018; Van den Berg, Coetzee, & Mearns, 2020) with less attention dedicated to big data and sport business (Hutchins, 2016; Millington & Millington, 2015) and overarching discussions of analytics in the field (Spaaij & Thiel, 2017; Szymanski, 2020). As big data has evolved in sport, this literature has become critical to our understanding of the role of sport analytics and its impact on various aspects of the industry. However, an area of deficiency in many big data articles is the lack of theoretical discussion and connection between individual cases of big data usage and a greater understanding of sport analytics as a tool in the organizational decision-making process.

According to Szymanski (2020), there are three main concerns within the literature. First, most sport analytics studies lack sufficient theory, which is how we make sense of large data sets. Second, the focus should be on the pursuit of knowledge rather than profit, which leads to secrecy and stunts the advancement of knowledge. Finally, a critical component of sport analytics is the ability to predict. The development of models to analyze large data sets must be predictive and public to advance our understanding of these analytic techniques. These points are significant, and addressing these concerns can lead to a more comprehensive understanding of the role that analytics plays in sport management theory and practical application.

The six articles in this special issue address some of these concerns by providing work in the sport analytics space that is grounded in theory, is focused on extending knowledge, and offers predictive qualities. The articles are diverse, including studies on the impact of unexpected outcomes on online reviews, athlete misconduct and sponsor stock prices, consumer demand as it relates to tanking, season ticket holder (STH) attendance, and hotels on gameday, and methodological investigations regarding big data and social media engagement. The following section highlights each individual article and its impact on our greater understanding of big data and analytics in sport.

Unexpected Outcomes and Online Reviews

Matti (2021) examined the role of emotional cues developed from unexpected game outcomes on perceptions of local businesses. This study extended knowledge on loss-aversion and reference-dependent preferences through an investigation of close to one million Yelp reviews. Consistent with loss aversion, Matti found that unexpected losses led to more negative business reviews. As online reviews have become a significant factor in driving consumer interest, this study extended our understanding of loss aversion within an online big data context. In addition, this examination extended sport analytics knowledge through volume and velocity of data, which are key aspects of big data.

Athlete Misconduct and Sponsor Stock Prices

Ge and Humphreys (2021) examined the impact of athlete endorsers who engaged in inappropriate conduct on sponsor stock prices. Previous theory has suggested homogeneity regarding the

negative impact of athlete misconduct on stock price. This investigation extended this work by examining potential heterogeneity based on type of misconduct. Through an examination of 863 examples of misconduct and their impact on millions of shares of stock traded daily, this study was able to expand the vertical and horizontal reach of the data, further enhancing its use from a big data perspective. Results showed that both the type of misbehavior (self-harm vs. harming others) and the level of media attention had differing impacts on sponsor stock price. Misbehavior that harmed others and increased media coverage led to greater declines in stock price. This investigation extended theory on the impact of endorser behavior on corporate stock prices while developing a model that could be used to predict the influence of future endorser misbehavior. These findings extended our theoretical and practical knowledge of this evolving relationship between athlete and sponsor.

Consumer Demand

There are three articles in this special issue focused on various aspects of consumer demand. All three studies advance theory while providing predictive models. DeSchriver, Webb, Tainsky, and Simion (2021) examined the impact of collegiate football games on local hotel demand. The authors leveraged a big data set with volume and variety as they enhanced both the scope and granularity of data. The investigation used multiple markets, multiple years, and multiple data sources. Over 1,000 hotels with over 4 million daily performance metrics were assessed for 1,249 home college football games. Findings showed that team performance, opponent, school events, and hotel class positively influenced demand. This study advanced demand literature in sport and hospitality using a large and diverse data set.

Karg, Nguyen, and McDonald (2021) examined “no show” behavior for STH of Australian Rules Football. Taking advantage of another large and diverse data set, the authors focused on why certain STH chose to attend games. Through an examination of over 59,000 observations of individual STH behavior, findings showed that individual STH factors, game viewing quality, and sporting contest quality stood out as important determinants of STH attendance. This investigation made significant analytic contributions by extending knowledge on interaction effects between game and individual characteristics as well as the impacts within consumption smoothing.

Gong, Watanabe, Soebbing, Brown, and Nagel (2021) investigated the impact of tanking in the National Basketball Association through sentiment analysis. The authors examined big data through a multidimensional lens by emphasizing the concept of variety within sport analytics research. Not only did this study examine patterns of behavior within the data itself but also merged it with other data to provide further levels of insight. Data included 166,000 tweets over the course of 5,945 unique National Basketball Association games. A predictive model for National Basketball Association home game attendance was developed, and results showed that the volume of discussions for the home team and sentiment toward tanking by the away team impacted game attendance. This investigation used big data to advance knowledge on tanking theory in sport and provided a model to predict tanking effects in the future.

Game Outcomes and Social Media Impact

Weimar, Soebbing, and Wicker (2021) examined the effect of game outcomes on change rates of social media followers. The

authors extended knowledge on methodology related to big data within the context of social media. The volume (over 900,000 observations) and variety (across three social media platforms) of the data in this study provided a more comprehensive understanding of social media impacts. The authors extracted the relative importance of wins, draws, and losses through dominance analysis, indicating that a victory yielded the highest increase in followers. Overall, this study expanded existing research on social media analytics within sport and advanced our understanding of methodological approaches when using big data sets.

The six articles in this special issue demonstrate the impact that big data research can have on our understanding of managerial concepts within sport. Collectively, these studies extend theory within their respective areas and across sport analytics. They also focus on the advancement of knowledge within the discipline and offer models that can predict future outcomes in the field. The evolution of data science has provided tools for extending knowledge beyond the scope of what has been done previously, providing a wealth of opportunity for future scholarly work that can build on the research within this special issue.

The Future of Big Data and Analytics in the Sport Management Literature

There are many areas of sport management that could benefit from the inclusion and integration of big data and analytics. The wealth of information that is now available to researchers presents a unique opportunity to advance research into existing topics while also allowing researchers to consider new questions that previously could not be examined with traditional data sets and methods (George et al., 2014, 2016).

The first topic area in sport management that could benefit from the application of big data and analytics is in segmentation—the examination of specific communities or user groups divided by certain characteristics (Fried & Mumcu, 2016). Although existing studies have certainly examined consumer behaviors of a wide range of groups based on various demographic factors, the granular nature of big data allows researchers to move beyond traditional classifications that have been used to divide groups. For example, computer scientists have used advanced statistical techniques to develop density models to segment hospitality industry consumers not by examining factors such as income or gender but, rather, by exploring nontraditional factors that group together these individuals, such as their likelihood to search for certain terms on the Internet (Bose, Munir, & Shabani, 2020). Although this certainly provides benefits to organizations looking to market and sell products/services to their customers, segmentation will also likely continue to grow in importance within consumer-based research as it will advance our theoretical and empirical understanding by providing more detail and nuance in relation to behaviors.

Another area that will likely continue to advance in the sport management literature is in research focused on social media, which has witnessed significant growth over the last decade with the proliferation of social networking sites (Abeza et al., 2015; Filo, Lock, & Karg, 2015). Notably, several studies have utilized social media data to examine a variety of phenomena and behaviors in sport, including the formation of digital networks (Naraine & Parent, 2016; Yan et al., 2019), how organizations and consumers use digital sites (Clavio & Kian, 2010; Watanabe, Yan, & Soebbing, 2015), and the communicative patterns displayed on social media (Wakefield & Bennett, 2018). In addition, two

articles included in this special issue examine big data in social media. Indeed, social media is one of the areas of sport management that has seen the greatest number of studies using big data and analytics precisely because social networking sites present locations wherein large masses of users are constantly creating incredible volumes of content (Van Dijck & Poell, 2013). For example, during the Super Bowl, users on various social media sites such as Twitter, Facebook, Instagram, and so forth, create millions of posts that include textual and visual information. At the same time, there is need for sport scholars to exercise caution in examining social media data. There is such a large volume of information that many studies will only be capturing a small snapshot of the actual volume of information and behaviors that are being displayed on these platforms.

Another area that will likely see advancement with the growth of big data and analytics in sport is pricing. Researchers have long been concerned with pricing in sport as it is one of the core determinants of the consumer decision to attend games and purchase products (Drayer & Rascher, 2013) and is a topic that extends across multiple disciplines including marketing, economics, and consumer behavior. Big data presents an opportunity for researchers focused on pricing as it allows an evolution to being able to examine consumer responses to price at a more granular level. That is, the traditional approach to examining pricing in sport has often considered price as being static with prices for tickets often being set well in advance of a season beginning (Fort, 2004; Soebbing, 2008) and, thus, not having significant fluctuations during a season. In this manner, studies that have incorporated information about ticket prices have generally used the lowest price or average price to consider the relationship between consumer interest in sport and price (Coates & Humphreys, 2007). However, with advances in ticket pricing methods used by organizations, including variable ticket pricing (Rascher, McEvoy, Nagel, & Brown, 2007), price dispersion (Soebbing & Watanabe, 2014), and dynamic ticket pricing (Drayer, Shapiro, & Lee, 2012; Shapiro & Drayer, 2012), understanding how consumers react to price has become a more dynamic and complicated process. As such, recording a single price for a ticket is not sufficient. Prices constantly fluctuate because of instantaneous changes of prices and reselling of tickets, and thus, there is need for data on the actual prices of transactions to improve the analysis of consumer response to pricing.

Finally, big data and analytics have already transformed the work environment inside and outside of sport with employers looking for new ways to be able to gather data about employees and evaluate their performance (Lopez, 2020). However, although player performance is seen as the source of the big data revolution in the sport industry, there is still ample room for growth, especially considering the complexity of human and organizational behaviors in sport. At the same time, there is also a need to acknowledge that beyond the topic areas mentioned in this introduction, numerous opportunities exist for the integration of big data and analytics in various areas of sport management. For example, sport ecology is one such area that could benefit from the use of big data, especially in measuring the impact that sport has on pollution (Locke, 2019) as well as how poor environmental conditions can impact fans (Watanabe, Yan, Soebbing, & Fu, 2019), players, and staff (Archsmith, Heyes, & Saberian, 2018).

To utilize big data, there is a need for scholars to access large data sets. There are several approaches, which can be used depending on the technical skills and resources that are available to a research team. One common approach to gathering large data sets is to use software designed to scrape large volumes of data from

websites and other digital sources. For example, the papers by Matti (2021) and Gong et al. (2021) published in this special issue both utilized scraping routines to collect textual information created by users on Yelp and Twitter, respectively. Typically, these scraping routines are either programmed by the user or a third party who then releases them for public use in programs such as R Studio or Python (R Studio, Boston, MA) (Landers, Brusso, Cavanaugh, & Collmus, 2016). Although scraping data from public websites has now been embraced by most disciplines and journals as a legitimate method of data collection, the need to understand programming language to gather data using this method is one potential barrier that prevents some scholars from utilizing big data.

Another option is to pay programmers or companies that specialize in scraping to collect the data. However, it should be noted that although the scrapers may be successful in accessing some data, the use of paid data collection to do big data is often prohibitively expensive. Furthermore, because the researcher is often not the one who is collecting and/or analyzing the data, they must give up control over their study, and this raises the question as to whether using third-party data collection and analysis is an acceptable practice in academic research.

Outside of scraping, there are other methods that can be used to gather large data sets. One common approach is to utilize data sets that are provided free of charge to either researchers or the general public. Such data sets are provided both by governmental entities, such as the Bureau of Economic Analysis in the United States, as well as corporations. Notably, corporations such as Yelp, which deal with large subsets of human behavior in the form of business reviews, have started the practice of releasing large data sets every year so that researchers can study this information as well as provide recommendations on how Yelp can improve its digital site for all users. In the case of data provided by corporations, there is some need for scholars to exercise caution as it may not always be the case that the data have been collected and cleaned to the standards expected in academic research.

This brings us to the final way in which researchers can access data, which is either through partnering with companies that have the tools to collect data or working directly with a sport organization that is producing the data. Although such partnerships between academics and the industry are certainly desired, there is also need for scholars to be careful when entering such partnerships due to publication restrictions and consistent data quality.

Conclusion

In summary, the six articles presented in this special issue further extend our knowledge of both theory and practice related to big data and analytics in sports. The authors examine the role of big data in sport and how it furthers our understanding of managerial concepts in the field. The scope of this special issue is broad and diverse, covering sport management in various contexts with large volumes of data through a variety of sources. As big data continues to evolve and permeate the sport landscape, this special issue can provide the groundwork for the advancement of sport analytics theory, methodology, and industry application. Future research can extend this work as technological advancements allow us to access, organize, and interpret large data sets throughout the sport industry, providing a more comprehensive understanding of ever-changing phenomena in our discipline.

References

- Abeza, G., O'Reilly, N., Séguin, B., & Nzindukiyimana, O. (2015). Social media scholarship in sport management research: A critical review. *Journal of Sport Management*, 29(6), 601–618. doi:[10.1123/JSM.2014-0296](https://doi.org/10.1123/JSM.2014-0296)
- Aizawa, K., Wu, J., Inoue, Y., & Sato, M. (2018). Long-term impact of the Tokyo 1964 Olympic Games on sport participation: A cohort analysis. *Sport Management Review*, 21(1), 86–97. doi:[10.1016/j.smr.2017.05.001](https://doi.org/10.1016/j.smr.2017.05.001)
- Archsmith, J., Heyes, A., & Saberian, S. (2018). Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation. *Journal of the Association of Environmental and Resource Economists*, 5(4), 827–863. doi:[10.1086/698728](https://doi.org/10.1086/698728)
- Bose, A., Munir, A., & Shabani, N. (2020). A quantitative analysis of big data clustering algorithms for market segmentation in the hospitality industry. In *2020 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1–6). Dublin, Ireland: IEEE.
- Clavio, G., & Kian, T.M. (2010). Uses and gratifications of a retired female athlete's Twitter followers. *International Journal of Sport Communication*, 3(4), 485–500. doi:[10.1123/ijsc.3.4.485](https://doi.org/10.1123/ijsc.3.4.485)
- Coates, D., & Humphreys, B.R. (2007). Ticket prices, concessions and attendance at professional sporting events. *International Journal of Sport Finance*, 2(3), 161.
- Colbert, A., Yee, N., & George, G. (2016). The digital workforce and the workplace of the future. *Academy of Management*, 59(3), 731–739. doi:[10.5465/amj.2016.4003](https://doi.org/10.5465/amj.2016.4003)
- DeSchriver, T., Webb, T., Tainsky, S., & Simion, A. (2021). Sporting events and the derived demand for hotels: Evidence from Southeastern Conference football games. *Journal of Sport Management*, 35(3). doi:[10.1123/JSM.2020-0268](https://doi.org/10.1123/JSM.2020-0268)
- Drayer, J., & Rascher, D. (2013). Guest editors' introduction: Sport pricing research: Past, present, and future. *Sport Marketing Quarterly*, 22(3), 123.
- Drayer, J., Shapiro, S.L., & Lee, S. (2012). Dynamic ticket pricing in sport: An agenda for research and practice. *Sport Marketing Quarterly*, 21(3), 184–194.
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897–904. doi:[10.1016/j.jbusres.2015.07.001](https://doi.org/10.1016/j.jbusres.2015.07.001)
- Filo, K., Lock, D., & Karg, A. (2015). Sport and social media research: A review. *Sport Management Review*, 18(2), 166–181. doi:[10.1016/j.smr.2014.11.001](https://doi.org/10.1016/j.smr.2014.11.001)
- Fort, R. (2004). Inelastic sports pricing. *Managerial and Decision Economics*, 25(2), 87–94. doi:[10.1002/mde.1108](https://doi.org/10.1002/mde.1108)
- Fried, G., & Mumcu, C. (Eds.). (2016). *Sport analytics: A data-driven approach to sport business and management*. New York, NY: Taylor & Francis.
- Ge, Q., & Humphreys, B. (2021). Athlete misconduct and team sponsor stock prices: The role of incident type and media coverage. *Journal of Sport Management*, 35(3). doi:[10.1123/JSM.2020-0106](https://doi.org/10.1123/JSM.2020-0106)
- George, G., Haas, M.R., Pentland, A. (2014). Big data and management. *Academy of Management Journal*, 57(2), 321–326. doi:[10.5465/amj.2014.4002](https://doi.org/10.5465/amj.2014.4002)
- George, G., Osinga, E., Lavie, D., & Scott, B. (2016). Big data and data science methods for management research. *Academy of Management Journal*, 59(5), 1493–1507. doi:[10.5465/amj.2016.4005](https://doi.org/10.5465/amj.2016.4005)
- Goes, F.R., Meerhoff, L.A., Bueno, M.J.O., Rodrigues, D.M., Moura, F.A., Brink, M.S., ... Lemmink, K.A.P.M. (2020). Unlocking the potential of big data to support tactical performance analysis in professional soccer: A systematic review. *European Journal of Sport Science*. Advance online publication. doi:[10.1080/17461391.2020.1747552](https://doi.org/10.1080/17461391.2020.1747552)
- Gong, H., Watanabe, N., Soebbing, B., Brown, M., & Nagel, M. (2021). Do consumer perceptions of tanking impact attendance at National Basketball Association games? A sentiment analysis approach. *Journal of Sport Management*, 35(3). doi:[10.1123/JSM.2020-0274](https://doi.org/10.1123/JSM.2020-0274)
- Hutchins, B. (2016). Tales of the digital sublime: Tracing the relationship between big data and professional sport. *Convergence*, 22(5), 494–509. doi:[10.1177/1354856515587163](https://doi.org/10.1177/1354856515587163)
- Jin, X., Wah, B.W., Cheng, X., & Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research*, 2(2), 59–64. doi:[10.1016/j.bdr.2015.01.006](https://doi.org/10.1016/j.bdr.2015.01.006)
- Karg, A., Nguyen, J., & McDonald, H. (2021). Understanding season ticket holder attendance decisions. *Journal of Sport Management*, 35(3). doi:[10.1123/JSM.2020-0284](https://doi.org/10.1123/JSM.2020-0284)
- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 1–12. doi:[10.1177/2053951714528481](https://doi.org/10.1177/2053951714528481)
- Landers, R.N., Brusso, R.C., Cavanaugh, K.J., & Collmus, A.B. (2016). A primer on theory-driven web scraping: Automatic extraction of big data from the Internet for use in psychological research. *Psychological Methods*, 21(4), 475. PubMed ID: [27213980](https://pubmed.ncbi.nlm.nih.gov/27213980/) doi:[10.1037/met0000081](https://doi.org/10.1037/met0000081)
- Lewis, M. (2003). *Moneyball: The art of winning an unfair game*. New York, NY: Norton.
- Locke, S.L. (2019). Estimating the impact of Major League Baseball games on local air pollution. *Contemporary Economic Policy*, 37(2), 236–244. doi:[10.1111/coep.12404](https://doi.org/10.1111/coep.12404)
- Lopez, M.J. (2020). Bigger data, better questions, and a return to fourth down behavior: An introduction to a special issue on tracking data in the National Football League. *Journal of Quantitative Analysis in Sports*, 16(2), 73–79. doi:[10.1515/jqas-2020-0057](https://doi.org/10.1515/jqas-2020-0057)
- Matti, J. (2021). Frustrated customers: The effect of unexpected emotional cues on Yelp reviews. *Journal of Sport Management*, 35(3). doi:[10.1123/JSM.2020-0147](https://doi.org/10.1123/JSM.2020-0147)
- McAfee, A., & Brynjolfsson, E. (2017). *Machine, platform, crowd: Harnessing our digital future*. New York, NY: WW Norton & Company.
- Millington, B., & Millington, R. (2015). 'The datafication of everything': Toward a sociology of sport and big data. *Sociology of Sport Journal*, 32(2), 140–160. doi:[10.1123/ssj.2014-0069](https://doi.org/10.1123/ssj.2014-0069)
- Mondello, M., & Kamke, C. (2014). The introduction and application of sports analytics in professional sport organizations. *Journal of Applied Sport Management*, 6(2), 11.
- Morgulev, E., Azar, O.H., & Lidor, R. (2018). Sports analytics and the big-data era. *International Journal of Data Science and Analytics*, 5(4), 213–222. doi:[10.1007/s41060-017-0093-7](https://doi.org/10.1007/s41060-017-0093-7)
- Naraine, M.L., & Parent, M.M. (2016). Illuminating centralized users in the social media ego network of two national sport organizations. *Journal of Sport Management*, 30(6), 689–701. doi:[10.1123/jsm.2016-0067](https://doi.org/10.1123/jsm.2016-0067)
- Rascher, D.A., McEvoy, C.D., Nagel, M.S., & Brown, M.T. (2007). Variable ticket pricing in major league baseball. *Journal of Sport Management*, 21(3), 407–437. doi:[10.1123/jsm.21.3.407](https://doi.org/10.1123/jsm.21.3.407)
- Shapiro, S.L., & Drayer, J. (2012). A new age of demand-based pricing: An examination of dynamic ticket pricing and secondary market prices in Major League Baseball. *Journal of Sport Management*, 26(6), 532–546. doi:[10.1123/jsm.26.6.532](https://doi.org/10.1123/jsm.26.6.532)
- Society for American baseball research (2021). A guide to sabermetric research. Retrieved from <https://sabr.org/sabermetrics>
- Soebbing, B.P. (2008). Competitive balance and attendance in Major League Baseball: An empirical test of the uncertainty of outcome hypothesis. *International Journal of Sport Finance*, 3(2), 119.

- Soebbing, B.P., & Watanabe, N.M. (2014). The effect of price dispersion on Major League Baseball team attendance. *Journal of Sport Management*, 28(4), 433–446. doi:[10.1123/jsm.2013-0024](https://doi.org/10.1123/jsm.2013-0024)
- Spaaij, R., & Thiel, A. (2017). Big data: Critical questions for sport and society. *European Journal for Sport and Society*, 14(1), 1–4. doi:[10.1080/16138171.2017.1288374](https://doi.org/10.1080/16138171.2017.1288374)
- Sung, H., & Mills, B.M. (2018). Estimation of game-level attendance in major league soccer: Outcome uncertainty and absolute quality considerations. *Sport Management Review*, 21(5), 519–532. doi:[10.1016/j.smr.2017.12.002](https://doi.org/10.1016/j.smr.2017.12.002)
- Szymanski, S. (2020). Sport analytics: Science or alchemy? *Kinesiology Review*, 9(1), 57–63. doi:[10.1123/kr.2019-0066](https://doi.org/10.1123/kr.2019-0066)
- Ticketmaster. (2011, March 15). LiveAnalytics launches to deliver rich fan insights in live entertainment for venues, teams, artists and sponsors. *Cision PR Newswire*. Retrieved from <https://www.prnewswire.com/news-releases/liveanalytics-launches-to-deliver-rich-fan-insights-in-live-entertainment-for-venues-teams-artists-and-sponsors-117992829.html>
- Van den Berg, L., Coetzee, B., & Mearns, M. (2020). Establishing competitive intelligence process elements in sport performance analysis and coaching: A comparative systematic literature review. *International Journal of Information Management*, 52, 102071. doi:[10.1016/j.ijinfomgt.2020.102071](https://doi.org/10.1016/j.ijinfomgt.2020.102071)
- Van Dijck, J., & Poell, T. (2013). Understanding social media logic. *Media and Communication*, 1(1), 2–14. doi:[10.17645/mac.v1i1.70](https://doi.org/10.17645/mac.v1i1.70)
- Wakefield, L.T., & Bennett, G. (2018). Sports fan experience: Electronic word-of-mouth in ephemeral social media. *Sport Management Review*, 21(2), 147–159. doi:[10.1016/j.smr.2017.06.003](https://doi.org/10.1016/j.smr.2017.06.003)
- Watanabe, N., Yan, G., & Soebbing, B.P. (2015). Major League Baseball and Twitter usage: The economics of social media use. *Journal of Sport Management*, 29(6), 619–632. doi:[10.1123/JSM.2014-0229](https://doi.org/10.1123/JSM.2014-0229)
- Watanabe, N.M., Yan, G., Soebbing, B.P., & Fu, W. (2019). Air pollution and attendance in the Chinese Super League: Environmental economics and the demand for sport. *Journal of Sport Management*, 33(4), 289–302. doi:[10.1123/jsm.2018-0214](https://doi.org/10.1123/jsm.2018-0214)
- Weimar, D., Soebbing, B., & Wicker, P. (2021). Dealing with statistical significance in big data: The social media value of game outcomes in professional football. *Journal of Sport Management*, 35(3). doi:[10.1123/JSM.2020-0275](https://doi.org/10.1123/JSM.2020-0275)
- Wicker, P., & Frick, B. (2015). The relationship between intensity and duration of physical activity and subjective well-being. *The European Journal of Public Health*, 25(5), 868–872. PubMed ID: [26142405](https://pubmed.ncbi.nlm.nih.gov/26142405/) doi:[10.1093/eurpub/ckv131](https://doi.org/10.1093/eurpub/ckv131)
- Yan, G., Pegoraro, A., & Watanabe, N.M. (2018). Student-athletes' organization of activism at the University of Missouri: Resource mobilization on Twitter. *Journal of Sport Management*, 32(1), 24–37. doi:[10.1123/jsm.2017-0031](https://doi.org/10.1123/jsm.2017-0031)
- Yan, G., Watanabe, N.M., Shapiro, S.L., Naraine, M.L., & Hull, K. (2019). Unfolding the Twitter scene of the 2017 UEFA Champions League Final: Social media networks and power dynamics. *European Sport Management Quarterly*, 19(4), 419–436. doi:[10.1080/16184742.2018.1517272](https://doi.org/10.1080/16184742.2018.1517272)
- Zaharia, N., Kaburakis, A., & Pierce, D. (2016). US sport management programs in business schools: Trends and key issues. *Sport Management Education Journal*, 10(1), 13–28. doi:[10.1123/SMEJ.2015-0007](https://doi.org/10.1123/SMEJ.2015-0007)
- Zikopoulos, P., & Eaton, C. (2011). *Understanding big data: Analytics for enterprise class hadoop and streaming data*. New York, NY: McGraw-Hill Osborne Media.