UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

A Culture of Competition: Sport's Historical Contribution to Datafication

Stauff, M.

DOI
10.18146/2213-7653.2018.365
Publication date
2018
Document Version
Final published version
Published in
Tijdschrift voor Mediageschiedenis

License
CC BY-SA

Link to publication

Citation for published version (APA):
Stauff, M. (2018). A Culture of Competition: Sport's Historical Contribution to Datafication. *Tijdschrift voor Mediageschiedenis*, *21*(2), 30-51. https://doi.org/10.18146/2213-7653.2018.365

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:08 Sep 2026

Markus Stauff

A Culture of Competition: Sport's Historical Contribution to Datafication[1]

Abstract

This article considers the contribution of sports to the emergence of a contemporary big data culture. Why and how did media sports become so entangled with big data? How do media sports impact on the popularisation of big data as a cultural practice and as a cultural imaginary? In the first part of the article, I demonstrate how, as early as the nineteenth century, the standardisation and serialisation of sport competitions went hand in hand with a growing relevance of quantified evaluation of performances. Sports contributed to the modern 'avalanche of numbers' and thus became an important symbolic resource for the broader implementation of a data-based, 'normalistic' regulation of social practices. In the second part, I focus on the implementation of a public representation of actual big data practices in professional sports. Starting with a short overview of an initially slow-moving, but eventually comprehensive, appropriation of advanced statistical calculations and other big data practices since the 1970s, I analyse examples that illustrate the controversies around and the public legitimisation of metrics and data visualisation. My main claim is that sport, because of its historically long and close entanglement with numbers, both stimulates a naturalisation of datafied competition and fuels an ongoing debate about the quality and implications of different forms of metrics.

Keywords: Big data; media sports; quantification; evaluation; public understanding of technology

Public communication about sports – in journalism, on television, and on social media – increasingly incorporates data to evaluate athletic achievements: for instance, tables comparing the average performances of two contenders throughout a season, or 'heat maps' highlighting zones of the field where players were active most often during a game. Arguably, media sports is the cultural practice and institutional context which offers the earliest and most extensive examples of the broader popularisation of 'big data' – the now ubiquitous utilisation of computers to search for hidden patterns and correlations in massive sets of heterogeneous, often automatically collected, data. In the following, I analyse how media sports have contributed to the emergence of contemporary data culture and to the familiarisation with otherwise abstract big-data practices and their cultural imaginary, i.e. their publicly discussed possibilities and dangers.

Brett Hutchins has analysed in detail how datafication encroaches upon all aspects of contemporary sports – scouting, coaching, fan culture, mass media's representation of sports, and much more. Sports thus contribute to the cultural imaginary (and legitimisation) of big data as a source for innovative knowledge, and naturalises data monopolies and the increasing divide between data-rich and data-poor sports organisations.[2] Focusing on (European) football, David Beer has similarly argued that the application of big data in professional sports is one of the key contributions to 'everyday neoliberalism' – the wide acceptance of allegedly objective decisions based on quantitative measures and their visualisations.[3]

Building on this research – and more general reflections on big data's cultural impact[4] – I want to focus on the following issues: Why and how did media sports become so entangled with big data? Which characteristics and traditions of sport make it so attractive for big-data applications? How do certain features of media sports shape the practices and cultural imaginary of big data? To answer these questions, this article will focus on two distinct historical developments, each offering specific insights into the dynamics and cultural impact of sports' data culture.

The first three sections (part I) outline the emergence of sports as a 'data-rich environment' in the nineteenth century. While not discussing it in detail, Hutchins and Beer both point to how the recent mushrooming of big data in sports was preceded by a long historical entanglement between sports and data. A closer look at this historical interrelationship not only offers greater insight into some key features of modern sports that shape today's application of big data, it also enables us to approach big data not as a completely new phenomenon but as 'a chapter in a longer history (or, rather, histories) of observation, quantification, statistical methods, models, and computing technologies.'[5] I will show how and why modern sport has spawned a systematic collection of data and earned some of its key characteristics through such data. In this context, numbers should be singled out as a common and constitutive element not only of sport itself but also of sport's representation in popular media, which contributed to the modern 'avalanche of numbers' and became an important symbolic resource for the broader implementation of a data-based, 'normalistic' regulation of social practices.

The final two sections (part II) take a big leap forward, focusing on the actual rise of big data in sports from the 1970s onwards to show how the historical entanglement of sports and data feeds into today's data culture. After a short overview of the initially slow, but eventually comprehensive, appropriation of advanced statistical calculations and other big-data practices since the 1970s, this section enhances the historical perspective by analysing contemporary controversies and discussions that accompany the public legitimisation of metrics and data visualisation in sports. Adding to Beer and Hutchins, I will argue that media sports, aside from supporting 'everyday neoliberalism', also enable cultural debate on the workings, advantages, and disadvantages of big data. So, while parts I and II of the piece partly stand on their own, they are conceptually and historically interrelated.

My arguments are mainly based on a re-reading of existing historical research, augmented with my own analysis of symptomatic examples. Instead of offering a comprehensive historical account, this article analyses a number of exemplary historical constellations

(from the nineteenth, twentieth, and twenty-first centuries) to distil the conceptual insights they can provide into the data–sports entanglement. The current conspicuous presence of big data in media sports – while also being used to attract viewers[6] – actually feeds on an older, structural interplay between data and sports. The manner in which sports contribute to the cultural imaginary of big data is shaped by particular features of sports culture (for example, its seriality, its partisanship, its competitiveness). Sport's historically long and dense entanglement with data presents us with both a naturalisation of datafied competition *and* an ongoing debate about the quality and implications of different forms of metrics.

Dutch national athletics championships 1974 at Papendal: a happy Ciska Janssen at the scoreboard. Photo: Bert Verhoeff. Nationaal Archief.

I – The Emergence of Modern Sport as a 'Data-rich Environment'

Gambling, Breeding and the Datafication of Performance

In the following three sections, I want to argue that sport's conspicuous entanglement with big data in recent years has been enabled and shaped by some characteristic features of modern sports that were established in the nineteenth century. It was in this period, indeed, that the production and public circulation of different kinds of data first became a constitutive element of competitive spectator sports. The first section uses examples from sports publications to give an initial impression of the increasing relevance of data throughout the nineteenth century. The second section adds a more conceptual perspective, elaborating on how the modern organisation of sports and the public circulation of data support each other. The third, in turn, discusses the significance of sport's data culture, over the years, for the emergence of a new concept of competition: in sports and beyond, performance became evaluated in comparison to calculated averages and thus became increasingly 'normalised'.

As I explain in more detail below, data (including sport data) is not always quantitative and does not always consist of numbers. Since numbers dominate the current application of big data, however, an older debate on the role of quantification in the history of sports might be helpful to understand this tendency. Many scholars (most famously, Allen Guttmann[7]) have argued that the emergence of modern sports in the nineteenth century was intertwined with the broader dynamics of quantification, rationalisation, and standardisation that characterise modernity. Others, however, have presented earlier examples of quantification in sports. Especially in ancient Greece, sport was an important pillar of a broader 'culture of competition', in which at least the male part of the population systematically contended for honour, wealth, or status.[8] The Olympic Games were only one among hundreds of similar contests regularly organised in Greece, Egypt, Rome, and Asia Minor between 700 BCE and 400 CE.[9] Even if the Greeks did not record the times and distances achieved in different events, there was a quest for all-time records. 'Recordmania'[10] often focused on the number of victories, 'employing a highly developed system that recorded which athlete was the first to win each event; or, who was the first to win a particular combination or number of victories.'[11]

Historians have offered examples from other eras as well. In the fifteenth and sixteenth centuries, for example, archery and jousting tournaments were based on scoring systems awarding points to a predetermined set of achievements.[12] Examples demonstrate that the competitive character of sport, throughout its heterogeneous history, was often highlighted by the translation of performances into quantified data. Yet, the data was mostly used to gauge achievements during an individual contest or to mark the exceptionality of one athlete in comparison to all others. Only a minority of performances became datafied, only a very limited repertoire of data forms was applied, and there was little public communication that harnessed the available data (or triggered the production of additional data) to compare performances across different contests.

The organisation of sports that eventually became favourable for the production and application of big data did not begin to emerge until the eighteenth century and was established solidly at the end of the nineteenth century. If sport, in the nineteenth century, did not yet produce 'big data', it co-emerged with a 'data-rich environment' in which measurements of system inputs and outputs became available more and more often.[13] Increasingly, characteristics of 'items' (for instance, the age, weight, nationality of athletes) or processes (for example, the length, place, result of a race) became classified, noted down, and visually organised for easy overview, retrieval, and comparison (tables, rankings, lists, graphics). Those examples show that data also has non-numerical aspects (names, nationalities), yet numbers play a particularly important role, and the regular and systematic collection of data allows for the calculated comparison of non-numerical features as well (such as the numbers of wins for different nations).

In this article, I am particularly interested in the driving forces and characteristic media forms of sport's emerging data-rich environment. Most historical accounts of sport in the nineteenth century include relevant examples; however, to my knowledge at least, there is no comprehensive account of the incremental datafication of sports throughout the nineteenth century.

A racing calendar listing the horses and jockeys of every English horse race was first used in 1727. According to Tony Collins, this calendar marks 'the beginnings of what would become known as sports statistics.'[14] Amongst other things, it offered race results from the previous years to ensure that fair competitions could be organised. The drive towards a more data-rich comparison of performance often came from commercial interests. Since organisers demanded entrance fees for horse racing, cricket, and boxing in the eighteenth century, they had to guarantee transparent, fair, and exciting competitions. Gambling was also one of the major impulses behind such efforts in record-keeping.[15] Thus, data was not only used to determine and praise the winner, but also to give an account of how the win was possible in the first place.

The above examples, taken from Tony Collins' *Sport in Capitalist Society* (2013), can be supplemented with others I found in popular magazines. Arguably, the latter help to highlight the increasing implicitness of data, and especially numbers, in the public discourse on sports. When the first edition of the *American Turf Register and Sporting Magazine* was published in 1829, the editorial, referring to existing British examples, aimed at serving 'as an authentic record of the performances and pedigrees of the *bred* horse.'[16] Here – and in horse racing more generally – the economics of breeding constitutes a further motivation for publishing regular and reliable data. The journal regularly published the exact winning times (in minutes and seconds) of a number of horse races. Interestingly, the first issue also included a detailed review of a 'New Time Keeper'. Costing 120 dollars, this instrument was recommended to 'all gentlemen who own running or trotting horses, and wish, in their private trainings, to determine the speed of their horses correctly. It divides time to the sixtieth part of a second.'[17] This is another example of how practices of data collection spread beyond the moment of competition and how the introduction of new, refined tools promised to measure – and thus to datafy – heretofore inaccessible aspects of performances.

Gambling, breeding, and other commercial interests thus provoked a more systematic collection and publication of easily comparable results from different competitions, which all gained their own, independent dynamics over time. In the second half of the nineteenth century, numbers and rankings were increasingly used to compare performances, condense past events, and create expectations for the future. In this period, the emergence of a data-rich environment also manifested in publications for a wider audience. In the 1880s and 1890s, Vienna's *Allgemeine Sport-Zeitung*, for example, published lists of records for one-kilometre races.[18] It also represented the shot-by-shot development of billiard games through number-filled tables which covered the larger part of a page.[19] In the process, nineteenth-century sports fans clearly became familiarised with huge amounts of numbers as appropriate representations of sporting events.

Again, it is important to recollect that numbers are only one form of data (even if perhaps the most conspicuous kind). The overall process of datafication is characterised by the fact that, increasingly, different (qualitative and quantitative) aspects of a competition become distinguished and can be used to analyse and compare performances. In 1881, the *Allgemeine Sport-Zeitung* announced 38 upcoming horse races, with six separate tables across two pages of the magazine. Each table sorts the races according to different criteria: the distance of the race, the age of the horses, the donor of the prizes, the weight rules, and the horses' countries of origin.[20]

In the following years, additional categories like prize money per horse, per jockey, and per breeder are added to such overviews. These tables allow the reader to observe otherwise invisible aspects of the individual competitions and make comparisons across different races. Whether involved in gambling or not, readers thus can make their own comparisons – for example, to find out whether one breeder's horses stand out in any one particular discipline.

If Greek athletics measured performance only in terms of the number of victories, and although the Roman circus also paid attention to the defeated (and dramatised losing),[21] the creation of a data-rich environment equips *all* performances with potentially interesting and relevant aspects, thereby spawning interest in the creation of even more data.

The Public and Universal Evaluation of Performances

The above examples have been chosen to give an impression of the conspicuous growth of sports data throughout the nineteenth century. In the case of horse races, the influence of breeding (and, partly, gambling) remained highly visible over time. However, the quest to compare performances in more and more aspects also achieved a dynamic of its own – a development that can be observed in a number of different sports around 1900. In this section, I offer a more conceptual account of why and how the entanglement of data and sports was so quickly and solidly established in the second half of the nineteenth century. The emergence of a wider – that is, mediated – public that extends the comparison of performance to an eventually universal scale is a key factor here.

So far, Tobias Werron has offered the most comprehensive analysis of this historical transformation. He argues that the connection between the popular press and telegraphy created a discourse that compared and evaluated sports performances – including performances from different places and from different moments in time.[22] Newspapers that print the results of competitions in several cities on one and the same page foster an interest in an extended comparison of performances – for instance, how would team A, which defeated all contenders in its home town, perform against team B, which also won all matches in its home town? On theone hand, this provoked a more comprehensive standardisation and universalisation of rules – especially in terms of each sport's spatial and temporal framework – and spawned the regular and systematic organisation of events. On the other hand, this 'gradual artificialisation of the sports environment'[23] and serialisation of contests triggered the production of data from different competitions and thus allowed a wider public (beyond the audience attending an individual match) to evaluate performances.

Of course, the ever-wider dissemination of sports as data culture was part of (and supported by) broader developments. Vanessa Ogle calls comparison the 'single most important intellectual device by which nineteenth-century observers gauged an interconnected and competitive world.'[24] The processes of standardisation that resulted from and enabled comparisons of sports performances went along with broader processes of temporal and spatial standardisation provoked by railroads, factories, and telegraphy, which together familiarised audiences with the use of data, and especially numbers, on a daily basis (for example, through railroad timetables).[25] The spread of sports' artificial environments was additionally fostered by both the 'idealistic

internationalism' of projects like the International Olympic Committee (founded in 1894) and certain imperialist endeavours (especially those of Great Britain), which often harnessed sports as 'soft power' to display their modernity or to transform 'the "native" into an athlete.'[26] Military forces, missionaries, educators, and of course organisations like the YMCA introduced the local elites to sports, while the culturally dominated groups – and especially the lower classes – often 'forced their unwelcome way into sports from which the dominant group desired to exclude them.'[27]

Sport, however, added its particular dynamic to these broader social and political processes, not least because its increasing artificialisation and the related data-based performance evaluations ultimately aimed at a 'universal horizon of comparison.'[28] This is most visible in the emergence of 'world records', comparing achievements from all over the world and throughout all of history. In other words, it is perhaps not rationalisation and quantification that are the key characteristics of spectator sports, but the 'continual communication about contests that creates a particularly modern relationship between single contests on the one hand and spatially, temporally and socially universalised comparison of contests on the other.'[29]

Not only was it the sports environment – the standardised fields with their related regulations – that became artificialised, so too did the evaluation of sports performances. The combination of serialised organisation of competitions with serialised observation of competitions eventually changed the definition of achievement. Ultimately, it was no longer the individual win that defined a champion but one's performance across a number of competitions and in relation to other competitions on a national or even global scale. The league system particularly 'defined the quality of the performance of football teams as a matter of long-term achievement (...), and thus redefined it in a basically statistical manner.'[30] Alongside other forms of ranking, the league system made this new form of competition visible for a wide audience and trained readers to adopt a more abstract, quantitative take on competition.

Quantification and records have always existed in sports, and the culture of gambling and breeding stimulated the publication of richer data on competitions and their participants. However, the emergence of modern competitive sports and their particular structure in the second half of the nineteenth century provided an incredible boost to the collection of data about sports performances. The ongoing, at least weekly, observation of dozens of similar events, all of which were optimised for contingent yet definite outcomes, unavoidably produced a data-rich environment, allowing for and fostered by a public evaluation of performances. Additionally, the comparison of performances from different places and different moments required that attention be paid to aspects of the performance beyond the mere end result.

Here too, it is important to note that data was not necessarily quantitative or numerical; narrative forms that summarised and hyperbolised individual situations also contributed to the public evaluation and comparison of performances. Inasmuch as the narrative forms were often based on repetitive scripts and schemata (as evidenced, for instance, by psychological portraits of athletes having 'heart' or being 'unstable'), they can be considered

data too. As I will argue later, however, the difference between quantitative and qualitative data gets even more explicit – and contested – with the transition to big data starting in the 1970s.

Historically speaking, the repetitiveness of standardised competitions allowed for the regular production and collection of numbers, and for reliable comparison of events across huge spatial and temporal distances, thus representing a significant innovation. Baseball is by far the most extreme example. Journalist Henry Chadwick introduced a standardised newspaper box score for the sport in 1859, recording five aspects for each player (runs, hits, putouts, assists, and errors). In 1861, he additionally published the first annual guide for baseball.[31] In his book *The Signal and the Noise* (2012), Nate Silver, who also runs the data journalism website *FiveThirtyEight*, describes baseball as:

> perhaps the world's richest data set: pretty much everything that has happened on a major-league playing field in the past 140 years has been dutifully and accurately recorded, and hundreds of players play in the big leagues every year. Meanwhile, although baseball is a team sport, it proceeds in a highly orderly way.[32]

In a sense, baseball might be considered an extreme case. However, I would argue that this is the case only in its historical context – in other words, it is a forerunner rather than an outlier. It demonstrates how sport creates endless numerical data, as long as the means are available to break down competitions into different parts and aspects. For some sports, this happened later than for others, but a consistently upward trend remains visible throughout the past 150 years. The serialised and standardised organisation of events, and the circulation of evermore detailed and standardised data, goes hand in hand with the emergence of a public that uses the data for a universally extended comparison of performances. This constellation offers a clear rationale for the application of big-data procedures, as soon as the required statistical tools and the computational power were available.

Sport's Contribution to 'Normalistic' Competition

Before taking a closer look at the actual take-off of big-data procedures in the late twentieth century, I would like to note that the data-rich environment of sports was already part of a broader cultural trend in the nineteenth century, which in turn contributed to the plausibility of new forms of competition. Since the main objective of these first three sections is to give a systematic, historical account of sport's current entanglement with big data, I once again focus, primarily, on quantitative data and on how sport contributed to the cultural imaginary of numbers and statistics.

Tony Collins claims that, in a more general sense, sport's historical emergence 'was not merely co-terminous with the expansion of capitalism but an integral part of that expansion, not only in economic organisation but also in ideological meaning.'[33] In Britain, this connection was even explicated in the public-school curricula of the early nineteenth century. Team sports like cricket, rugby, and football were supposed to teach pupils teamwork, leadership, physical

courage, and 'the importance of competition'[34] – and in this respect, one might argue, their rationale was not so different from that of 'rankings', 'talent', and 'achievement' in today's neoliberal ideology.[35] Nina Verheyen additionally shows how the growing visibility of sport and its recordmania undermined the conviction that the potential for human achievement had a natural limit.[36]

More specifically, sport was part of the nineteenth century's 'avalanche of printed numbers',[37] with its relatively new forms of quantified, statistical, and normalised competitions. Ian Hacking, in his book *Taming of Chance* (1990), describes how at the very beginning of the nineteenth century, London and Paris used statistics to debate which of the two cities was more 'suicidal'.[38] Jürgen Link analysed in detail how this avalanche of numbers contributed to a modern form of competition based on statistical averages and relates to what he calls 'normalism': the organisation of evermore social practices according to a discourse of normality based on statistical averages but also on narratives and metaphors distinguishing normal and abnormal behaviour.[39] In the late nineteenth century, calculated averages achieved a new and ambivalent function as benchmarks for 'normal' functioning and 'normal' performance (which are different from 'normative' benchmarks). On the one hand, the average was a positive reference point that could signal problematic extremes: results that are too high above or too low below the average are considered problematic (or, in medical terms, pathological or abnormal). On the other hand, the average becomes the negative reference point for competition: it signals boredom or lack of progress and threatens the recognition of individual achievements. Even more importantly, though, the exponential growth developments in the late eighteenth century and early nineteenth century (in population, industries, et cetera) forced a disconnection of the average from any kind of pre-established norm or ideal type; the average itself became a dynamic reference point provoking a broader, quantitative and qualitative, reflection on what was considered normal.[40]

In contrast to older forms of rivalry, modern, data-based forms of competition thus evince a strong diachronic dynamic: if any metrics of a body or of a country are below average at one point in time, this signals the possibility of and the need for improvement, so that an above-average result can be achieved the next time round – or in other words: it suggests that it can be normalised. Through the homogenisation of separate fields of societies, such as education, punishment, and healthcare, but also family life, sexuality, suicidal behaviour, or public transportation, they all can be regularly measured and thus observed in terms of how their performances are positioned with respect to the average. With each moment of measurement, though, the average itself changes; ideally it increases, thus driving the next step of the competition.

While this normalistic form of social (self-)regulation is based on the homogenisation and datafication of distinct practices, it requires collectively plausible symbols and narratives if it is to have an effect beyond circles of experts and professionals. In the second half of the nineteenth century, Francis Galton, the man who coined the concept and term of eugenics and created the statistical notion of correlation, referenced sports to model human improvement through competition.[41] Jürgen Link now argues that sport is one of the most important pillars

for the broader social plausibility and legitimacy of such normalistic competition. In his book *Versuch über den Normalismus* (2006), he suggests a number of key aspects of data-rich, normalistic competition which are most clearly modelled in sports.

Firstly, sports display competition as a social form: by creating a well-defined and homogenised set of rules – the proverbial 'level playing field' – modern competitive sports guarantee that everybody can participate under identical circumstances (consider here the aforementioned 'universal horizon of comparison' in sports). Serialised competition gives everybody a shot. While the datafication of the entire population makes it difficult for an individual to imagine his or her own position in the datafied social field, sports metonymically shows the select few publicly participating in a competition.[42]

Secondly, the universalisation of competition is enabled by sport's creation of layered hierarchies, which are all integrated into the same competition. This in turn makes competition realistic for everyone. While of course it makes no sense for us to compete with the elite players, we *can* participate in a lower-level form of the same competition. This is most clearly formalised in the system of hierarchised leagues. The Champions League, the Premier League, the B Leagues, or Youth Leagues all have their specific averages and their own rankings, which gives us all a realistic chance of participation. This system even gets applied in political discourse to classify individual countries as belonging to different 'leagues', each with their own specific normalities.[43] While exporting cricket and football to colonial territories supposedly helped to establish a particular form of 'rational' administration, similar to that applied more generally by the colonising country, the symbolic order of sports allowed for the perpetuation of a hierarchical distance between coloniser and colonised, who might play the same game but always at a different level.

The league system is one example of the final and perhaps most productive dynamic that sport contributes to normalistic competition: the flexible combination of continuity and discontinuity between different competitions. Leagues combine continuity with discontinuity in a vertical manner. Teams can be promoted or demoted between the different leagues of, say, English football – a logic that creates continuity. However, the level of performance expected (their 'normality') is different *within* each league, which in turn signals a discontinuity (it would be 'unfair' to compare a team from the third league to one in the first). In horizontal terms, we can think of the extended parallel existence of black and white leagues, or of the existence of separate male and female competitions, both of which normalise the discontinuity between the performance of certain groups of people. Sport thereby embodies the alleged need to create different, discontinuous classes of normality, but *also* the possibility to transition between these different classes or even of their eventual integration.[44]

Once again, non-quantified cultural procedures (such as narratives or heroic character types) are also of importance to all these aspects of sports competition. It might even be argued that the intense combination of quantitative data (for instance, records and averages) with qualitative forms (narratives of persistence, or cultural clichés about male and female attitudes) is what makes sport such a rich model for the broader adoption of artificialised, diachronic competition. The normalistic (self–)regulation of societies always involves adding symbolic resources

to numerical data in order to help individuals and organisations 'navigate' between averages and extremes and to equip quantified performances with excitement or fear.[45] (For instance, a 7–1 result in a game between two teams of the same league is 'not normal' and therefore perceived as a 'humiliation' for the losing team.) At the same time, however, the important symbolic functions sport performs for broader normalistic competition are based on a data-rich environment that creates homogeneity and continuity (for example, between all baseball games), which then allow for the measurement and marking of averages, differences, and hierarchies (between players, teams, countries, leagues, et cetera). It is this focus on quantitative, datafied competition that makes sport such a rich metaphorical source for naturalising rankings in politics and economics.[46]

II – Big Data in Contemporary Sports Culture (1970s–today)

The Take-off and Institutionalization of Big Data in Sports

The data-rich environment that allowed for the emergence of modern spectator sports in the nineteenth century also made possible the current boom of big data in sports. This was not a linear process and, as I will outline in this section and the next, big data is partly so visible in sports *because* it is contested. To show how the systematic entanglement of sports and data plays out and contributes to the cultural imaginary of big data today, I will take a big leap, discussing first how big data became established in contemporary sports, and then flagging which aspects of big data are highlighted in its connections with sports.

Throughout the twentieth century, the media representation of sports – especially television coverage, which dominated sports culture in the second half of the century – was dominated, arguably, by spectacle: heroic narratives, beautiful bodies, commercialised and nationalistic mega events.[47] Data, however, continuously remained an important and highly visible element of such spectacle. Reporters for radio and television always based their narrative accounts on score sheets, rolodexes, index cards, 'spotting boards' (to quickly identify all team members) and, of course, human assistants to facilitate the recording and quick retrieval of data.[48] Even if the numbers were often overshadowed by more qualitative data, they remained essential throughout.

In a seminal text from 1983, Margaret Morse shows how television uses graphics and statistics to 'underline the scientificity'[49] of the competition (and to legitimise the pleasures of watching male bodies). For a long time, though, these statistics remained pretty stable and only covered a very narrow set of features of the performances. While visual technologies (like slow-motion replay [50]) regularly changed the ways in which sports performances were evaluated and appreciated, big data's disruptive potential to redefine the understanding of performance and success was actually met with scepticism in the organisationally and hierarchically conservative world of professional sports.

Again, baseball is a case in point. While there had been efforts to this effect in the 1950s and 1960s, the use of statistical calculations to gain a refined understanding of baseball truly gained traction in the 1970s. It was mainly fans who collected data on past games and suggested

metrics in order to compare players, understand the impact of venues on performance, and debate what were actually the most significant metrics for the success of a team.[51] The most famous instance of this is the publication, in 1977, of Bill James' first *Baseball Abstract*, presenting '68 pages of in-depth statistics.'[52] In its first five years, the yearbook was self-published; afterwards, a commercial publisher took it on for seven more instalments. In the 1980s, James tried to engage a network of fans to collect play-by-play data for past games. While he did not succeed in doing so, a similar project – called *Retrosheet*, set up in 1989 – was eventually successful. It now offers the play-by-play data of every Major League game since 1871 on a non-commercial website.[53] Since the early 1990s, a lot of related activities were developing online, e.g. within the online discussion group rec.sport.baseball and on the commercially successful statistics and forecast site *Basketball Prospectus*.[54] These examples also illustrate Aronova's general observation that 'practices of observing, collecting, and sorting data are by nature collective endeavors.'[55]

Fans' interest in, and knowledgeability about, data has also been fostered by fantasy sports, an increasingly successful genre since the 1980s. Paper-based games were the precursor for online games, with both using data from actual leagues to let players compete as managers of self-selected teams.[56] This datafied approach transforms the practice of having fans identify with one athlete into fan engagement as 'vicarious management'.[57] Launched in 1991, *USA Today Baseball Weekly* contained extended box scores from games, specifically catered to fantasy baseball players.[58]

These developments of the 1970s and 1980s can be understood in terms of a transition to 'big data, as the volume, velocity, relationality, and variety of data production dramatically increased.'[59] After an early phase when things were still done with pen and paper, personal computing brought the promise of individual manageability of data collections – at which time the movement became part of what Kevin Driscoll has described as 'database populism'.[60] As Boyd and Crawford argue for big data in general, these dispersed practices were 'changing the objects of knowledge, while also having the power to inform how we understand human networks and community.'[61] Sparked by fans, such big-data practices initially showed interesting parallels to what one might call 'data activism': the re-use of institutional data and their re-appropriation within 'alternative epistemologies'.[62] Computational methods were used to identify patterns with 'the aura of truth, objectivity, and accuracy.'[63] Eventually, though, it turned out that fans' alternative evaluations of performance were not at all oppositional to the hegemonic sports discourse, and they quickly became a commodified contribution to the evaluation of performance (and its ongoing improvement).

It took until the late 1990s before such new models for data-based evaluations were systematically taken up by professional teams, which then started hiring specialists for data analytics and metrics. These specialists often came from the financial or gambling industries, and were hired to support the scouting of players and the development (and evaluation) of playing tactics.[64] This quickly led to an 'intensification of systems of measurement [and] the rise of powerful data infrastructures' which characterise the broader transition to big-data culture.[65] While baseball was the earliest example, similar developments took root in other sports in the

early 2000s, including European football, which was long considered much more resistant to datafication because of the character of the game and a lack of historical data.

In addition to the ever-novel ways of using the already available data (for instance, through the discovery of new correlations and patterns), the professionalisation of sports analytics also provoked technological innovations to collect more detailed data. More and more games, including games from lower leagues and from other countries, were videotaped to then be fractionised into data on passes, assists, shots, and so on. Even today, this work is often still done manually and sometimes outsourced to workers in third-world countries. Increasingly though, tracking systems are automated, producing further data on previously inaccessible aspects of performance. Since 2006, the speed, position, and break of every pitch in Major League baseball have been measured in real time.[66] In basketball, the NBA introduced a system that automatically tracks players, refs and the ball, capturing 72,000 coordinates per game for each tracked item.[67] A more recent trend is the introduction of wearable technology that offers real-time physiological data on players. Since 2017, the NFL has partnered with a data company to position RFID chips in all players' shoulder pads. The data is supposed to grant 'a deeper understanding of the game by accessing new visualizations, stats and fantasy recommendations never available before.' It is symptomatic of the interlacing of professional and popular endeavours in this field that the data is made available to the teams, but also to 'broadcasters as well as other media partners to enhance the fan experience.'[68] For European football, with its more flowing game situations, tracking was key to the eventual development of a big-data logic specific to the sport. Meanwhile, data visualisations have become staples in the evaluation of sports performances as well as their communication to broader audiences.

The multiplication of data also led to new, commercial data-collecting and metrics-offering intermediaries, who tend to sell their services to both professional teams and media companies.[69] One of the most interesting developments here is how the shared relevance of statistics and metrics to professional and fantasy sports has blurred prior boundaries and allowed fans and gamers to become actual experts whose knowledge is respected and sometimes even paid for by professional teams.[70] In contrast to citizen science, where the data collection of amateur volunteers often needs to be fostered and simultaneously carefully disciplined to make it productive,[71] sports fans are prolific sources whose insights can easily be tapped by the industry.

There is no need to recount all the details of this development here. The resulting organisational struggles and the changes to scouting and coaching have all been summarised in a number of hugely popular books.[72] One of the most famous among them is 2003's *Moneyball*, which was made into a film in 2011.[73] Since the 2010s, the popular press (from special-interest magazines to daily newspapers) has also, at least selectively, adopted the use of the kind of data- and metrics-based analytics which previously had a narrower audience through blogs and online forums (the latter of which in turn had taken over from the paper-based yearbooks of the 1970s and 1980s).[74]

In all of this, the 'disruptive' potential of big data was underlined from the start. One of the earliest accounts of 'data activism' in the popular sports press appeared, in 1964, under the title

"Baseball Is Played All Wrong", and similar titles announcing a revolution in sports still appear regularly today.[75] When procedures that had been developed by fans over the course of twenty years were finally taken up by professional teams in the late 1990s, they changed game strategies and tactics, and had an even greater effect on the ways in which players were evaluated and hired. They contributed to new conflicts and hierarchies in teams' organisation – not least because the qualifications and expertise of the sitting managers and scouts was put into question.

Taking advantage of the data-rich environment that shaped and stabilised the emergence of modern sports, big data's characteristic drive to produce evermore numbers for the disclosure of heretofore unrecognised patterns and correlations continuously questions what it is that 'makes a performance'. The journalistic accounts of big data, published since around 2010, take advantage of this surprising change of perspective too. They often tell stories of underappreciated players, who do not play spectacularly, but for whom the data shows that their presence makes the team play better – for instance, because they do not have to resort to tackling or because they prevent dangerous situations *before* they occur. In this way, the power of big data is made evident to a broader public.

Legitimising and Debating Big Data

The developments sketched above only intensify sport's symbolic function in a broader normalistic culture. The compilation of evermore data on evermore detailed aspects of sports performances allows for the near endless multiplication of diachronic competitions. It gets easier to add classifications – for example, by ranking the best defensive and offensive players, or the best left-handed or right-handed players. Data visualisations add accessibility and plausibility to data-based claims. Nowhere is the observation of competitions in terms of averages, progressions, correlations, and 'normal deviations' as regular and as popular as in sports (where suspicions of doping, for instance, can be raised, but also appeased, on the basis of the calculated 'normalcy' of performance data[76]). Further research is needed to better understand the data politics of contemporary sports and show whether and how certain tools and data visualisations lend plausibility to further divisions of the field, how they add new layers and hierarchies, and how they use calculated averages to argue for a discontinuity or continuity between performances.

Here, however, I want to focus on a different aspect: Instead of asking how big-data sport contributes to the broader cultural (re-)definition of competition and normality, I want to zoom in on the way contemporary sport offers a key arena where big data is put forward as a visible, but also contestable, form of knowledge. While sport unquestionably contributes to hegemonic ideas about performance, it also provides a field for open discussion of the 'correct' ways of evaluating performances. In this final section, I claim that this is another key contribution of sport to the cultural imaginary of big data.

The fan practices described in the previous section are one example of where the relevance of big data – and the 'right' way of using it – is explicitly discussed. Though most of the aforementioned popular accounts tend towards a rather teleological story, in which big data finally

won over the traditionalists and eventually proved its superiority, the actual big-data saga is still ongoing and very much contested. I already pointed out that quantitative and qualitative, and narrative and numerical, forms of accounting played an important role in sports from the beginning. While sport, at least since the end of the nineteenth century, is a necessarily data-rich cultural practice, this does not mean that all agree on which data are the most relevant and insightful. On the contrary: sports always foster discussions about which are the correct, essential, and most telling aspects of a performance – and which, therefore, are the most appropriate forms of observation.

The emergence of big data in sports provoked publicly visible conflicts between more traditional forms of scouting on the one hand and data-based selection of new players on the other. Big data analysis still needs to legitimise itself against general scepticism and alternative, well-established modes of evaluating performance, but it necessarily remains an experimental approach that is continuously exploited for making new pleas for more comprehensive, more precise, or simply additional metrics. Thus, there are at least two debates: one about the legitimacy of numbers in sports; the other about which numbers are the most appropriate. In the following, I will sketch examples of both kinds of discussion.

In sports media, the advantage of big data over *other* forms of evaluation needs to be made explicit over and over again. With its focus on the human body, sport most clearly articulates the tension between competition as an abstract mechanism and competition as physical rivalry. As recently as 2016, the German former football international Mehmet Scholl raged in a radio show against what has, half-jokingly, been called the generation of laptop coaches: he accused them, for example, of ignoring the human aspect of coaching and of rejecting the more edgy, physical players.[77] In the US, former basketball player Charles Barkley plays the role of the data sceptic just as stubbornly. In a notorious rant, he declared analytics to be a misconception of effeminate managers who overlook the relevance of individual talent: 'a bunch of guys who ain't never played the game [and who] never got the girls in high school.'[78] While there is a broad consensus that a competition is eventually condensed into a numerical account (e.g. the 2–1 of a football game), the ongoing performance itself thus becomes symbolically protected against extensive quantification. Such popular debates powerfully continue a much older and broader dichotomy that parallels narrative accounts with the human scale and quantitative accounts with inhuman, abstract rationality.[79] This tendency seems particularly strong in European football, but even in the 'numbers game' of baseball, complaints that the box score drains the life from the sport have been around since the early twentieth century.[80]

When popular media do opt for a data-based approach, they often invest extra effort to make its advantages explicit to their readers. In 2015, German weekly *Der Spiegel* introduced new metrics to evaluate players of the Bundesliga (Germany's primary male football competition), under the headline 'Enough with the Arbitrariness!' ('Schluss mit der Wilkür!'). The first issue explained that, contrary to the often intuitive and subjective grades that players are allocated in other publications, the grades in this publication are based on numbers only – a lot of numbers, and carefully weighted ones.[81] It mixes a form of transparency that

allows readers to understand which data are used and how, with the hegemonic ideology implying that:

> data can be used to inform decisions and that this is seen to be a more objective and analytically accurate approach to decision-making. In other words, there is a sense that reducing the need for human intuition, discretion and agency can lead to decision-making being more accurate and value-generating.[82]

This leads to the second debate. Even where data practices are broadly accepted, the chosen metrics have to continually prove their pertinence and virtue – not least through the ways in which they connect with and integrate other, not data-related, forms of evidence. The debates of the 1980s and 1990s often already 'concerned not *whether* statistics should be used, but *which* ones should be taken into account' (my emphasis).[83] Today, professional data analysts and mass media alike still have to continuously display the relevance and solidity of their algorithms and metrics.

In a short promotional clip, the sports-data company Opta presents itself as a group of hard-working and committed people who take the greatest pains to produce reliable data. They specifically underline that their data provides much more detailed insights than the mere (quantitative) final result of the game.[84] The mass media, in turn, use data-crunching technologies to uphold their original promise: that one can see more of a sport in front of one's television set than in the stadium. Here, big data feeds into the competition between different media and between different media companies; in the process, the promise of better analytics is an important asset, which helps to attract and retain audiences. However, quality is not a self-evident aspect of a given metric or data-visualisation; the evaluative power of each practice always has to be *made* plausible.

For the 2016 European Championship in men's football, a German public broadcaster presented a new model for evaluating individual performances, which was called 'packing'. Interestingly, this model was developed by two former professional players who were not satisfied with the way defensive players were evaluated. They claimed that their system, which takes as its key indexical point the number of opposing players a dribble or pass gets past, focused on quality, not quantity, thereby deflecting the above-mentioned concern that statistics abstract away the human aspects of a game.

Additionally, the system received praised for its simplicity. Online fanzine *Bundesliga Fanatic* reminds its readers not to 'forget that up to thirty million Germans watch "Die Mannschaft" play, so you need a metric that is easy to explain and understand, since grandmas and kids watch too.'[85] However, it is symptomatic of big data's cultural imaginary that German television references German triumphs to establish the superiority of the 'packing' metrics. It did so by claiming that the system could account for Germany's 7–1 win against Brazil during the 2014 World Cup much better than alternative approaches, and in boasting that it offered evidence that the German Player Toni Kroos delivered a more outstanding league season than Lionel Messi.

National soccer match on 3 March 1963. The Belgian team has beat the Dutch. Photo: Hugo van Gelderen, Nationaal Archief.

Conclusions

One should not be surprised by the extent to which big data produces evidence in sports, not through mere metrics, but also through its entanglement with cultural meaning – for instance, in the aforementioned case, nationalism. Inversely, the meanings produced by sports also lend extra visibility to the qualities and impact of big-data procedures. The data-rich environment that was built up over the course of more than a century allows for inventive applications and offers a general plausibility to statistics and metrics. At the same time, sport's focus on authentic physicality, individual talent, and spontaneity supports a kind of nostalgia for pre-datafied forms of competition. In this way, the stakes of commercialised sport, its partisanship, and the attendant 'forensic fandom' together guarantee a constant controversy regarding the most appropriate ways to evaluate performance – with or without data.[86]

Modern competitive sport, this article has argued, plays an important role in the emergence and shaping of contemporary big-data culture. Earlier research, especially Hutchins' and Beer's contributions, already offers a sorely needed critical perspective. It is symptomatic of the broader effects of datafication that the growing relevance of metrics is driven by the commodification of data, producing in turn new hierarchies (in terms of data rich vs data poor). In this way, the popularity of sport's datafied competition feeds into 'everyday neoliberalism'.[87]

On the one hand, I have tried here to expand these insights with a more historical and systematic analysis of the entanglement of sports and data. In the nineteenth century, the emergence of modern competitive sports and that of a data-rich environment mutually supported each other. Sport's recurring, rule-based competition, with its quantitative results, continuously produced data that could be applied in order to compare performances across time and space. The quantified records of performances simultaneously fostered a sense of a universalised, fair,

and objectively evaluated competition. Twenty-first-century sports, in other words, have not suddenly been 'taken over' by a big-data regime, and the rise of big data is not caused solely by computer technology. Prior to the spread of neoliberal rationalities, sports already offered important symbolic support for a normalistic culture in which the most variegated aspects of society became organised by compartmentalised competitions and metrified rankings that carefully balanced growths and averages, and outstanding and mediocre performances.

On the other hand, I have also sought to highlight some interesting ambivalences of the sports–data entanglement. Having emerged in, and contributed to, a data-rich environment in the nineteenth century, sport seems a most pertinent field for the application of big data. And indeed, after a phase of hesitation, professional sports took inspiration from the data activism of fans and have now become one of the most publicly visible examples of metrics, data visualisation, and big data's pattern-recognition dynamics. Even so, such temporary scepticism might be significant, as it shows how sport, not least because of its characteristic public comparison of performances, partisanship, traditionalism, and nostalgia, contributes to, and ignites, an ongoing debate about what the most appropriate evaluation criteria are. Therefore, the introduction of big data in sports has been, and continues to be, a matter of public discussion – more so than the introduction of big data in other fields.

In sports, of course, the discussion on big data remains highly depoliticised. Fed by either a nostalgic (and counterfactual) refusal of quantification and abstraction, or by an eagerness to present an even more advanced, more abstract but allegedly more significant metric, the debate rarely touches upon issues of data commodification or the naturalisation and normalisation of hierarchies. For those who value data literacy, it might be interesting to further research the circulation of metrics and data visualisations in popular sports culture and the discursive dynamics that, while also lending those plausibility, open them up for scrutiny and debate. Perhaps the multiplication of metrics in sports – a practice currently focused primarily on competition and rankings – might then even leave some room for insights not structured by the 'winner–loser culture'[88] so dominant in sports.

Notes

1. Earlier versions of this article were presented at the workshops 'Historische Soziologie der Rankings' (University Bielefeld, October 2017) and 'Global Culture of Competition' (University of Amsterdam, January 2018). The participants of these workshops, the two anonymous peer reviewers, and the editors of this special issue gave essential feedback to the arguments of the text. Florian Duijsens helped to make the final version more readable.

2. Brett Hutchins, "Tales of the Digital Sublime: Tracing the Relationship between Big Data and Professional Sport," *Convergence: The International Journal of Research into New Media Technologies*, 21 May 2015, https://doi.org/10.1177/1354856515587163.

3. David Beer, "Productive Measures: Culture and Measurement in the Context of Everyday Neoliberalism," *Big Data & Society* 2, no. 1 (10 June 2015) https://doi.org/10.1177/2053951715578951.

4. Danah Boyd and Kate Crawford, "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon," *Information, Communication & Society* 15, no. 5 (June 2012): 662–679, https://doi.org/10.1080/1369118X.2012.678878; Zeynep Tufekci, "Big Data: Pitfalls, Methods and Concepts for an Emergent Field," *SSRN Electronic Journal*, 2013, https://doi.org/10.2139/ssrn.2229952; Rob Kitchin and Tracey P. Lauriault,

"Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work," *The Programmable City Working Paper* 2 (29 July 2014), http://www.nuim.ie/progcity/.

5. Elena Aronova, Christine von Oertzen, and David Sepkoski, "Introduction: Historicizing Big Data," *Osiris* 32, no. 1 (1 September 2017), https://doi.org/10.1086/693399.
6. Television coverage of sports, for example, regularly refers the audience to accompanying websites or apps which are supposed to offer additional background info such as historical or real-time data on individual players or teams. This extends the number of sites where they can show commercials and keeps the – potentially distracted – audience within reach of the broadcaster.
7. Allen Guttmann, *From Ritual to Record: The Nature of Modern Sports* (New York: Columbia University Press, 1978).
8. Nick Fisher, "The Culture of Competition," in *A Companion to Archaic Greece*, ed. Kurt A. Raaflaub and Hans van Wees, Blackwell Companions to the Ancient World (Chichester, UK; Malden, MA: Wiley-Blackwell, 2009), 528.
9. Sofie Remijsen, *The End of Greek Athletics in Late Antiquity*, Greek Culture in the Roman World (Cambridge; New York: Cambridge University Press, 2015).
10. Dietrich Ramba, "Recordmania in Sports in Ancient Greece and Rome," in *Ritual and Record. Sports Records and Quantification in Pre-Modern Societies*, ed. John Marshall Carter and Arnd Krüger (Westport: Greenwood Press, 1992), 35.
11. David C. Young, *A Brief History of the Olympic Games* (Malden, MA: Blackwell Pub, 2004), 30.
12. Joachim Rühl, "Sports Quantification in Tudor and Elizabethan Tournaments," in *Ritual and Record. Sports Records and Quantification in Pre-Modern Societies*, ed. John Marshall Carter and Arnd Krüger (Westport: Greenwood Press, 1992), 65–86.
13. Howard Jay Chizeck, "Modelling, Simulation and Control in a Data Rich Environment," *Proceedings of the Annual Symposium on Computer Application in Medical Care*, 26 October 1986, 65.
14. Tony Collins, *Sport in Capitalist Society: A Short History* (New York: Routledge, 2013), 8.
15. Collins, *Sport in Capitalist Society*, 8.
16. *American Turf Register and Sporting Magazine* 1, no. 1 (September 1928), 1 (accessible at https://babel.hathitrust.org/cgi/pt?id=umn.31951000735810m;view=1up;seq=8); on the relevance of the magazine, see Gerald R. Gems, Linda J. Borish, and Gertrud Pfister, *Sports in American History: From Colonization to Globalization* (Champaign, IL: Human Kinetics, 2008), 86.
17. Gems, Borish and Pfister, *Sports in American History*, 42.
18. For example: *Allgemeine Sport-Zeitung*, 5 January 1896, 13.
19. For example: *Allgemeine Sport-Zeitung*, 27 January 1881, 46.
20. *Allgemeine Sport-Zeitung*, 13 January 1881, 15f.
21. Remijsen, *The End of Greek Athletics*, 336ff.
22. Tobias Werron, *Der Weltsport und sein Publikum: Zur Autonomie und Entstehung des modernen Sports*, (Weilerswist: Velbrück, 2009); Tobias Werron, "World Sport and Its Public. On Historical Relations of Modern Sport and the Media," in *Observing Sport. System-Theoretical Approaches to Sport as a Social Phenomenon*, ed. Ulrik Wagner and Rasmus Storm (Schorndorf: Hofmann, 2010), 33–59.
23. John Bale, *Sports Geography*, 2nd ed. (London; New York: Routledge, 2003), 129.
24. Vanessa Ogle, *The Global Transformation of Time: 1870–1950* (Cambridge, Massachusetts: Harvard University Press, 2015), 6.
25. Ogle, *The Global Transformation of Time*; Eviatar Zerubavel, "The Standardization of Time: A Sociohistorical Perspective," *American Journal of Sociology* 88, no. 1 (1982): 1–23.
26. For the former, see John Hoberman, "Toward a Theory of Olympic Internationalism," *Journal of Sport History* 22, no. 1 (1995): 37; Emily S. Rosenberg, "Transnational Currents in a Shrinking World," in *A World Connecting, 1870–1945, A History of the World*, ed. Idem (Cambridge, Mass: Belknap Press of Harvard University Press, 2012), 831–834; for the latter, Bale, *Sports Geography*, 53.
27. Allen Guttmann, *Games and Empires: Modern Sports and Cultural Imperialism* (New York: Columbia Univ. Press, 1994), 179; see also: S. W. Pope, "Imperialism," in *Routledge Companion to Sports History*, ed. S. W. Pope and John Nauright (London and New York: Routledge, 2010), 229–247.
28. Werron, "World Sport and Its Public," 38.
29. Tobias Werron, "How Are Football Games Remembered? Idioms of Memory in Modern Football," in *European Football and Collective Memory*, ed. Wolfram Pyta and Nils Havemann (Houndmills, Basingstoke, Hampshire: Palgrave Macmillan, 2015), 18–39.

30. Werron, "How Are Football Games Remembered?", 34.
31. Collins, *Sport in Capitalist Society*, 57. For a concise historical overview of some of the changes of the box score and how it related to changes of the rules of the game, see: Alan Schwarz, "Take Me Out to the Box Score," *The New York Times*, 2 April 2006, http://query.nytimes.com/gst/fullpage.html?res=9B01E7D71230F931A35757C0A9609C8B63.
32. Nate Silver, *The Signal and the Noise. Why so Many Predictions Fail – but Some Don't* (New York: Penguin Press, 2012), 80; for a critical evaluation of Silver's own data practice, see: Hutchins, "Tales of the Digital Sublime".
33. Collins, *Sport in Capitalist Society*, 13.
34. Ibid., 30.
35. William Davies, *The Limits of Neoliberalism: Authority, Sovereignty and the Logic of Competition* (Los Angeles: Sage, 2014).
36. Nina Verheyen, *Die Erfindung der Leistung* (München: Hanser Berlin, 2018), 168.
37. Ian Hacking, *The Taming of Chance* (Cambridge [England]; New York: Cambridge University Press, 1990).
38. Ibid., 64–80.
39. Jürgen Link, "On the Contribution of Normalism to Modernity and Postmodernity," *Cultural Critique*, no. 57 (2004): 33–46; Jürgen Link, *Versuch über den Normalismus. Wie Normalität produziert wird* (Göttingen: Vandenhoeck & Ruprecht, 2006).
40. Link, *Versuch über den Normalismus*, 323ff.
41. Ibid., 157.
42. Ibid., 325.
43. Ibid., 431.
44. Ibid., 335.
45. Jürgen Link, "On the Temporal Quality of the Normalistic Fun and Thrill Tape," *Cultural Critique*, no. 57 (2004): 69–90.
46. Eva Lavric and Josef Weidacher, "Rankings in Sports Discourse and Their Metaphors," in *The Discourse of Sport: Analyses from Social Linguistics*, ed. David Caldwell et al. (Taylor & Francis, 2016), 151–169.
47. Alan Tomlinson and Christopher Young, ed., *National Identity and Global Sports Events: Culture, Politics, and Spectacle in the Olympics and the Football World Cup* (State Univ. of New York Press, 2006); Michael Real, "Sport and Spectacle," in *Questioning the Media. A Critical Introduction*, ed. John Downing, Ali Mohammadi, and Annabelle Sreberny-Mohammadi (London/Thousand Oaks/New Delhi: Sage, 1990), 345–356; Michael Oriard, *Reading Football. How the Popular Press Created an American Spectacle* (Chapel Hill/London: Univ. of North Carolina Press, 1993); Richard Gruneau, "Making Spectacle: A Case Study in Television Sports Production," in *Media, Sports and Society*, ed. Lawrence A. Wenner (London: Sage, 1989), 134–154.
48. This is most visible in autobiographical accounts by reporters, for example: Bill Mercer, *Play-by-Play: Tales from a Sports Broadcasting Insider* (Lanham, MD: Taylor Trade Pub, 2007).
49. Margaret Morse, "Sport on Television: Replay and Display," in *Regarding Television: Critical Approaches – An Anthology*, ed. E. Ann Kaplan (Los Angeles: American Film Institute, 1983), 50.
50. Morse, "Sport on Television"; Markus Stauff, "The Accountability of Performance in Media Sports: Slow-Motion Replay, the 'Phantom Punch', and the Mediated Body," *Body Politics* 2, no. 3 (2014): 101–123.
51. Jeff Hildebrand, "Playing Moneyball. Math and Baseball," in *Mathematics in Popular Culture: Essays on Appearances in Film, Fiction, Games, Television and Other Media*, ed. Jessica K. Sklar and Elizabeth S. Sklar (McFarland, 2013), 114–122.
52. "Bill James," *Wikipedia*, 6 March 2018, https://en.wikipedia.org/w/index.php?title=Bill_James&oldid=829024040.
53. http://www.retrosheet.org; see also: Hildebrand, "Playing Moneyball. Math and Baseball," 118.
54. Hildebrand, "Playing Moneyball," 117.
55. Aronova, Oertzen, and Sepkoski, "Introduction," 6.
56. Richard G. Lomax, "Fantasy Sports: History, Game Types, and Research," in *Handbook of Sports and Media*, ed. Arthur A. Raney and Jennings Bryant (Mahwah, NJ: Lawrence Erlbaum Associates, 2006), 416–425.
57. Thomas Patrick Oates, "New Media and the Repackaging of NFL Fandom," *Sociology of Sport Journal* 26, no. 1 (March 2009): 31–49.
58. *The Washington Post*, January 29, 1991; https://www.highbeam.com/doc/1P2-1046769.html
59. Kitchin and Lauriault, "Towards Critical Data Studies".
60. Kevin Driscoll, "From Punched Cards to 'Big Data': A Social History of Database Populism," *Communication +1* 1, no. Article 4 (29 August 2012), http://scholarworks.umass.edu/cpo/vol1/iss1/4.
61. Boyd and Crawford, "Critical Questions," 665.

62. Stefania Milan and Lonneke van der Velden, "The Alternative Epistemologies of Data Activism," *Digital Culture & Society* 2, no. 2 (1 January 2016), https://doi.org/10.14361/dcs-2016-0205.
63. Boyd and Crawford, "Critical Questions," 663.
64. For examples, see: Christoph Biermann, *Matchplan. Die neue Fußball-Matrix* (Köln: Kiepenheuer & Witsch, 2018).
65. Quote from Beer, "Productive Measures," 4.
66. https://fivethirtyeight.com/features/baseballs-new-pitch-tracking-system-is-just-a-bit-outside/; 28 April 2017.
67. https://zenportfolios.ca/bcit-bsys-2050-2015/tag/basketball/; 27 February 2015.
68. https://www.zebra.com/us/en/about-zebra/newsroom/press-releases/2017/zebra-technologies-nfl-wilson-sporting-goods-ball-tracking.html; 7 September 2017.
69. For a critical analysis of their role, see: Hutchins, "Tales of the Digital Sublime".
70. https://www.theguardian.com/technology/2015/dec/05/how-football-manager-changed-the-game; 5 December 2005.
71. Etienne S. Benson, "A Centrifuge of Calculation: Managing Data and Enthusiasm in Early Twentieth-Century Bird Banding," *Osiris* 32, no. 1 (1 September 2017): 286–306, https://doi.org/10.1086/694172.
72. To name but a few examples: Simon Kuper and Stefan Szymanski, *Soccernomics: Why England Loses, Why Spain, Germany and Brazil Win, and Why the US, Japan, Australia, Turkey – and Even Iraq – Are Destined to Become the Kings of the World's Most Popular Sport* (New York: Nation Books, 2012); Dean Oliver, *Basketball on Paper. Rules and Tools for Performance Analysis* (Washington: Brassey's Inc., 2004); Tobias J. Moskowitz and L. Jon Wertheim, *Scorecasting: The Hidden Influences Behind How Sports Are Played and Games Are Won*, Reprint edition (New York: Three Rivers Press, 2012); Andy Glockner, *Chasing Perfection: A Behind-the-Scenes Look at the High-Stakes Game of Creating an NBA Champion* (Boston, MA: Da Capo Press, 2016); David Sally and Chris Anderson, *The Numbers Game: Why Everything You Know About Soccer Is Wrong* (Penguin Books, 2013).
73. Michael Lewis, *Moneyball: The Art of Winning an Unfair Game* (New York, NY: W. W. Norton & Company, 2004); Bennett Miller, *Moneyball*, 2011.
74. There is no space here for detailed discussion of the co-evolution of data journalism and data-based reporting on sports. It is characteristic of many similar developments that Dutch news website *De Correspondent* (http://decorrespondent.nl), which focuses on in-depth coverage, has regularly discussed metrics- and data-related developments in sports since 2014; symptomatic for the commercial interests in this development, *The Guardian* dealt with the topic in some of its 'sponsored features'.
75. In 2013, a book with the subtitle *Why Everything You Know About Soccer is Wrong* (see above); in 2015, *The Guardian* published an article on sports data under the headline "A Whole New Ball Game" (15 April 2015).
76. E.g.: http://www.cyclingnews.com/news/dumoulin-releases-vuelta-a-espana-power-data/; 19 September 2015.
77. https://www.merkur.de/sport/fussball/mehmet-scholl-neue-kritik-an-laptop-trainern-9432899.html; 8 December 2017.
78. https://www.si.com/nba/2015/02/11/charles-barkley-analytics-video-daryl-morey-houston-rockets-gm; 11 February 2015.
79. John Durham Peters, "'The Only Proper Scale of Representation': The Politics of Statistics and Stories," *Political Communication* 18, no. 4 (October 2001): 433–449, https://doi.org/10.1080/10584600152647137.
80. Schwarz, "Take Me Out to the Box Score".
81. http://www.spiegel.de/sport/fussball/bundesliga-spielernoten-spix-und-spon11-die-vermessung-der-bundesliga-a-1160253.html.
82. Beer, "Productive Measures," 5.
83. Silver, *The Signal and the Noise*, 95.
84. https://www.youtube.com/watch?v=Gy-qpVfOWA.
85. http://bundesligafanatic.com/impect-packing-the-future-of-football-analytics-is-here/.
86. The notion of forensic fandom is borrowed from Jason Mittell, "Forensic Fandom and the Drillable Text," *Spreadable Media* (blog), 17 December 2012, http://spreadablemedia.org/essays/mittell/.
87. Hutchins, "Tales of the Digital Sublime"; Beer, "Productive Measures".
88. Claudio Colaguori, *Agon Culture: Competition, Conflict and the Problem of Domination* (Whitby, Ont: de Sitter Publications, 2012), 3.

Biography

Markus Stauff teaches Media Studies at the University of Amsterdam. His main research interests are television and digital media, governmentality, and the visual culture of media sports. Recent publications include: "Slow-Motion Replay, the 'Phantom Punch', and the Mediated Body," *Body Politics* 2, no. 3 (2014): 101–123; "Taming Distraction: The Second-Screen Assemblage, Television and the Classroom," *Media and Communication* 4, no. 3 (2016); "Non-Fiction Transmedia: Seriality and Forensics in Media Sports," *M/C Journal*, March (2018).