Data Everyday: Data Literacy Practices in a Division I Sports Context

Tamara Clegg[1,2], Daniel M. Greene[1], Nate Beard[1], Jasmine Brunson[1]
Information Studies[1], Education[2]
University of Maryland, College Park
{tclegg, dgreene1, nbeard}@umd.edu; jbrunso1@terpmail.umd.edu

ABSTRACT
Data analysis is central to sports training. Today, cutting-edge digital technologies are deployed to measure and improve athletes' performance. But too often researchers focus on the technology collecting performance data at the expense of understanding athletes' experiences with data. This is particularly the case in the understudied context of collegiate athletics, where competition is fierce, tools for data analysis abound, and the institution actively manages athletes' lives. By investigating how student-athletes analyze their performance data and are analyzed in turn, we can better understand the individual and institutional factors that make data literacy practices in athletics meaningful and productive—or not. Our pilot interview study of student-athletes at one Division I university reveals a set of opportunities for student-athletes to engage with and learn from data analytics practices. These opportunities come with a set of contextual tensions that should inform the design of new technologies for collegiate sports settings.

Author Keywords
HCI and Sports, Data Literacy, Personal Informatics

CSS Concepts
• **Human-centered computing~Human computer interaction (HCI)**;

INTRODUCTION
Sports analytics is a rapidly growing area of study and innovation. Popularized by baseball general manager Billy Beane, *sabermetrics* introduced the sports world to the sophisticated use of statistical analysis of sports records for evaluating player performances to make personnel and game-time decisions [16, 22, 30, 51]. Today, nearly every major sports team uses data analytics to inform major decisions [67]. In light of the rising popularity of data science and analytics, researchers from a variety of fields including business, athletics, computer science, and HCI have begun to explore the development of innovative tools for data collection and visualization in sports contexts.

But sports are necessarily a social experience. Too often, sports analytics researchers focus on the technology and its impact on performance at the expense of understanding athletes' experiences with data [52]. This is particularly the case in the largely understudied context of intercollegiate athletics, where competition is fierce, tools for data analysis are ubiquitous, and the institution actively manages athletes' personal and academic lives [69]. By investigating how student-athletes analyze their data and are analyzed in turn, we can better understand the individual and institutional factors that make data literacy in athletics meaningful and productive—or not. Such research is vital for designing tools and technologies for data analysis in sports contexts, and could inform the design of new data literacy learning experiences for student-athletes.

Here we present a pilot interview study at one (US) Division I institution with 11 student-athletes, and one strength and training coach, from 8 different intercollegiate sports. We seek to understand the data literacy practices of student-athletes, coaches, and trainers—emphasizing the student-athletes' engagement with and access to these data. Within this context we ask:

- What are the practices around data among student-athletes, coaches, trainers, and athletics staff members in a (US) Division I sports context?
- How does the social and organizational context of Division I sports structure these practices?

Our pilot interview study reveals a set of opportunities for student-athletes to engage with and learn from their experiences and interactions with data. These opportunities come with a set of contextual tensions that should inform the design of new technologies for collegiate sports settings. This work contributes to the growing body of work in HCI and sports, offering a detailed analysis of a largely understudied sports context: organized, high-level varsity college athletics. Additionally, we integrate frameworks of data literacy within sports contexts to shed new light on ways such athletic practices relate to data literacy.

BACKGROUND
In this section, we summarize three bodies of research that inform this study. First, we define data literacy and its

CHI 2020, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04...$15.00.
DOI: https://doi.org/10.1145/3313831.3376153

relevance to sports. Second, we situate our work in the context of US-based intercollegiate sports and underscore sport science research on student-athlete learning and well-being. Third, we consider prior research on technology for sports analytics from various fields as well as from HCI.

Defining Data Literacy: Two Perspectives

Our definition of data literacy synthesizes findings from two research contexts. First, data literacy has been defined by researchers in the library sciences as a subset of information literacy [39, 50, 60]. Information literacy is defined as recognizing when information is needed and then locating, evaluating, and using it effectively [60]. *Information* includes discovered knowledge, or processed data (i.e., from prior research, news sources, books, etc.); while *data* are representations of properties of objects and events [1]. Hence *data literacy* is defined as knowledge about how to access, interpret, critically assess, manage, handle, communicate, preserve, and ethically use data in order to gain new understanding of the world [39, 50, 60]. Library and Information Sciences researchers have developed a framework that breaks down the components of data literacy: (1) having an awareness of what data are and their role in society; (2) understanding how to find and obtain data (e.g., being aware of types and selecting the most relevant); (3) reading, interpreting, and evaluating data; (4) managing data (e.g., being aware of the need to save or archive data and understanding how to use tools to do so); (5) preparing data for analysis, synthesizing, and analyzing for specific questions; and (6) ethically using data and acknowledging data sources [39, 50].

Information and data literacy practitioners recognize the relevance of information literacy for athletes and have begun to launch efforts to expose them to sports-related library resources [57]. However, most efforts to engage student-athletes on college campuses in information literacy (e.g., bringing athletes to libraries or library resources for information course requirements) [57] have been decoupled from their experience as athletes. Additionally, the efforts have been in the broader context of information literacy, focusing on library resources as opposed to their day to day experiences with data in their sports play.

Second, our definition of data literacy includes the development of practices and skills foundational for *data science*. Research primarily from business and computer science on data science, defines data science as a set of fundamental principles that support and guide the principled extraction of information and knowledge from data. Data science involves identifying opportunities for automated data analysis as well as the innovative use and interpretation of such analytics for decision making [51]. Data science emphasizes big data or large data sets often containable only in specialized database systems [e.g., 2]. Whereas data literacy emphasizes general knowledge of data and its uses, data science focuses on data-driven decision making—practically applying insights from data to real-world decisions. Because of this, data science researchers emphasize the importance of carefully considering the context in which data is situated, having the sense to look deeper into the metrics to see when something is beneficial, what problems might arise, what's being counted and what's not. But it also includes having an eye for when and how data analytics, data mining, and big data sets would be helpful for new problems in new domains. Thus, we include this set of foundational practices and skills in our definition of data literacy.

Business, privacy, and information systems researchers have studied sports contexts as an application area for data science [16, 22, 67]. In doing so, they have identified social and ethical considerations for data science in sports settings. Specifically, they have shown the need for buy-in from leadership in athletics for data science practices to be effective [22] and they have recognized the need for awareness and privacy rights of individuals from whom data is being collected [16]. Yet the focus of data science research has been on types of data being collected, ways to store and mine the data, and ways organizational leaders use data science to make decisions. Less emphasis has been placed on the experience of those being monitored or studied.

We aim to bridge the gaps in data literacy and data science research to study the experiences of those from whom data is being collected—student-athletes in our case. We want to see how student-athletes are exposed to and engaged with data.

Division I Sports as a Context for Data Literacy

Our study is situated within the context of American intercollegiate sports. We focus on a Division I university in the National Collegiate Athletic Association (NCAA). The NCAA regulates and organizes most athletics programs of colleges and universities in the United States and Canada [69] and is divided into three divisions— Divisions I, II, and III. Divisions I and II offer athletics scholarships. Division I schools generally have the largest student bodies, manage the largest athletic budgets, offer the most athletic scholarships [40], and are considered the most elite athletic programs in the country [69].

The level of competition at Division I universities continuously increases, intensifying the need for and use of technology to enhance sports play. At the same time, the safety, well-being, and social and educational development of the young people in this ultra-competitive context has been called into question. For example, Division I schools have built expansive, athlete-only facilities that isolate athletes from the larger campus and student body [20, 69] and other forms of isolation include academic major clustering—especially for African-American athletes—into fields that lead to lower paying jobs after graduation [18, 19, 28, 47, 54, 59].

In light of these findings and recent occurrences at Division I schools, athletics teams have been called upon to foreground the athletes' experience and well-being and to consider ways athletes learn and develop through their athletic experience [68, 69]. Weight et al., [69] for example, compared health literacy surveys of collegiate athletes and non-athletes. They found that athletes had significantly higher health literacy scores than non-athletes, but both sets of scores were relatively low. Weight et al., [69] suggest the need for more opportunities for athletes to become aware of and reflect on what they are learning through their sports play. This work and others [34, 58, 62, 68, 69, 72] suggest there are myriad ways athletes (and non-athletes) learn and develop through their everyday sports practice and ways to harness sports play for learning. Zimmerman et al., [72] for example are leveraging youth soccer teams' sports play to introduce them to machine learning techniques, through modeling their sports moves. Although this body of work has looked at athletes' health literacy, collaboration, and machine learning, we specifically focus on student-athletes' practices around data. While we are not studying learning in this paper, it acts as a pilot study in which we can understand the context of data collection, use, and analysis in these team settings and identify opportunities for learning.

Technology Studies of Sports Analytics

Technological innovations have contributed a great deal to the range of data collection and analysis practices and tools available to sports teams. These technologies often leverage sophisticated analytics to enhance sports play. Business, computer science and sports science researchers have published on the wide array of technology used on elite, semi-professional, and professional sports teams. These technologies include systems for monitoring training loads (i.e., the level of intensity and effort of a player or team during practice or training sessions) [7, 9, 29, 56], biometric analyses to assist with injury prevention and recovery [53, 65], video analyses of sports plays [4, 70], and automated feedback for athletes on ways to adjust their technique and training for enhanced play [2].

Often, these technologies leverage or promote sophisticated data analysis techniques and tools. For example, video analysis tools allow teams to slow down, group, or annotate video data to analyze sports plays or techniques. Similarly, injury prevention tools analyze patterns within an athlete's training and performance data to determine when they are most at risk for injury [65]. While opportunities for data analytics abound in these new technologies, researchers in computer science, business, and sports science have primarily been focused on measuring the effects of these tools and techniques on athletes' sports play and/or fitness. Less work has considered the athlete's experience with these tools and data analytic techniques.

Similarly, there have been a small group of researchers in HCI who focus on sports analytics and are actively developing new devices and techniques to capture performance data (e.g., sensors, motion capture techniques) as well as new interactive visualization techniques for those data [e.g., 3,16,18,19]. Data visualization research in this area has specifically been focused on making the large amounts of data produced in sports contexts digestible to analysts, coaches, scouts, journalists, and fans to understand and compare games, plays, and trends [38, 45, 46, 48, 49, 61]. However, with the exception of several emerging studies [27, 31] little attention has been paid to athletes' experience with data collection, analysis, and interpretation, particularly as they are the objects of study and must carry out the plays or techniques analytics suggest. Herdal's work [2] does explore visualizations for youth athletes and suggests that data visualizations can help youth develop data literacy. However, few studies explore the influence of team and staff infrastructure on the design and use of sports visualizations. Furthermore, the widespread use of data in sports presents an opportunity for life-relevant learning [13], where athletes could learn through engaging in activities and pursuits deeply meaningful to their lives. However, we first need to understand athletes' relationship to data literacy and the role it plays on their teams.

We particularly build on the emerging body of work in HCI and sports that focuses on understanding the context of sports play and interactions in order to inform design. Indeed, a subset of these studies has looked at athletes' experiences with data in a variety of sports contexts, focusing specifically on the social experiences of athletes and how they collect and share data and information. Nylander & Tholander [44, 64] interviewed golfers, runners, and skiers to understand their experiences and data collection practices with wearable technologies. Their findings highlighted the importance of athletes' subjective feelings (e.g., of pain, enjoyment, ease, etc.) and advised that designers incorporate such subjective experiences into technologies, integrating them with quantitative measurements of performance. They also found that while athletes focus intensely on data collection, they engage in much less data analysis, often not leveraging a fraction of the analytic capabilities of their wearable technologies. Wozniak et al., [71] similarly conducted an interview study with trail runners, climbers, and skiers to understand how they share information within these contexts. Their study highlighted the importance of helping others navigate sharing data with different types of audiences (e.g., fellow athletes, supporters, public groups), helping athletes plan and manage sports activities to avoid risks (e.g., weather tracking, injuries, etc.), helping athletes collect and share data at different scales and complexity, and helping athletes navigate privacy concerns.

While these studies inform our understanding of data practices within sports contexts, the focus is largely on recreational contexts where participation is voluntary and driven by the athlete. Less work considers socio-cultural understandings of more formalized semi-professional or professional sports team contexts where motivation, organization, structure, and technology use are likely much different. There is one notable exception: Rapp & Tirabeni [52] compared the experiences of amateur and elite athletes across endurance and non-endurance sports. Seeking to understand athletes' experiences with personal informatics, they found that elite athletes think more critically about their data in part due to their extensive reliance on sensory perceptions of their bodies to adjust their sports performance and in part due to their close communication with knowledgeable coaches who work with them to contextualize both sensory and quantitative data to make informed decisions about their sports performance. They suggest the potential for personal informatics tools to help amateur athletes learn to better contextualize and think critically about their data in a manner similar to the way elites work with coaches to make more informed training and performance decisions.

Rapp & Tirabeni's [52] research builds on the extensive body of work on personal tracking and quantified-self efforts [e.g., 12, 21, 32–37, 55] which emphasize data collection and analysis practices via tracking devices to capture and analyze data for daily life goals (e.g., healthy living, attaining fitness goals). Our work is alternatively grounded in the perspective of data literacy because we want to understand the cognitive and process-oriented aspects of athletes' practices so that we can map to broader canonical learning—data literacy in our case—that could extend beyond sports contexts and into other areas of study and learning. Our focus thus includes but extends beyond the *personal* tracking athletes do to also consider which practices are prevalent and relevant to the *team* as a whole, how the team works together to carry out these practices, and how both individual and team practices are structured by the larger institutional setting of high-level campus athletics. Focusing on athletes' experiences from this perspective enables us, in a context of institutional learning (i.e., universities) to understand both the practices and ways they might promote life-relevant learning.

METHODS

Context/Setting

To understand student-athletes' experiences with data, we conducted 12 semi-structured interviews with athletes from a variety of sports—and one strength and conditioning trainer—about how they collect, analyze, and apply data and how these practices fit into social, athletic, academic, and health contexts. All athletes and the trainer were from State U, a Division I university in the northeastern United States. Table 1 lists the participant's pseudonym, sport, and gender. Our interpretivist approach is meant to inform the design of a larger, longitudinal study of how and what student-athletes learn from their individual and team data.

We recruited participants through snowball sampling, aligning our sample size with common practices in qualitative research, particularly in HCI [8, 26]. Sampling across sports provided researchers with a variety of data practices to compare in different social contexts (e.g., ball sports like American football or softball analyze trajectories, running sports like track analyze pace and weight). Sampling across State U's teams also provided researchers with a diverse participant set (5 male, 7 female; 2 African-American, 1 Arab-American, 2 Asian-American, 7 White). College sports are obviously separated by gender, but de facto demographic groupings by race and class also exist—often owing to the family resources required to participate in certain amateur sports prior to attending college. For example, black men make up 56% of NCAA Division I basketball players but only 3% of golfers [41]. Participants were compensated with a US $25 Amazon gift card. The study was approved by the researchers' Institutional Review Board and, for the sake of transparency and recruitment, the State U Athletics Department.

Interviews were split between the first two authors and conducted in closed conference rooms on the State U campus between April and July of 2019. Interviews were recorded and then transcribed. The hour-long interviews focused on three topics: data practices for sports play (e.g., "What sort of things do you measure for your sports play? Who does the measurement and how?"), organizational practices (e.g., "What sort of messages do your trainers give you about your data?"), and academic practices (e.g., "What is your major and how did you choose it?"). Open-ended questions allowed for a wide variety of responses about the different kinds of data participants found important in their lives. The first two authors wrote and shared journal reflections with each other immediately following interviews.

Data Analysis

Transcripts were coded in stages, following a constructivist grounded theory approach [14]. The first two authors developed thematic codes on separate transcripts, compared them, and consolidated the codes based on the overlap.

PIs then separately open coded a subset of interviews, compared for consistency, and repeated until all transcripts were coded. Our analytic approach was based in Critical Discourse Analysis (CDA), a mode of textual analysis founded in critical linguistics [17, 23] and later generalized to other fields [e.g., 19,28,50], which attempts to situate micro-sociological conversation (e.g., interviews) within macro-sociological contexts (e.g., Division I Campus Athletics). It is an inductive process for understanding how participants make meaning through social practice, and so thematic codes necessarily emerge from analysis rather than being established in advance [10]. The goal was to develop 'sensitizing concepts' [6] that explain how participants understand what data is collected from them, how, and why—in their own terms. Our approach is 'critical' insofar as it attends to the power differentials throughout the sample, exploring players' uneven capacity to make meaning and have that meaning accepted throughout the organization [11].

This method surfaced four core themes, each with a set of descriptive child codes: Data Analysis Practices (example child codes: Data Ranking, Data Cleaning), Personal and Team Technologies (example child codes: Counting Calories, GPS and Fitness Trackers), Social Interactions Around Data (example child codes: Athletes Withholding

Data, Coaches Holding Athletes Accountable), and Data Feelings (example child codes: Overthinking It, Freer Without Data). As a final review, two research assistants read the full transcripts and then the coded excerpts attached to each code, informing the lead authors of their agreement or disagreement with a particular code. Disagreements were discussed and codes were calibrated, negotiated, and finalized.

Pseudonym	Sport	Gender
Seth	Wrestling	Male
Caitlin	Track—Thrower	Female
Amanda	Golf	Female
Lei	Track—Distance Runner	Female
Nicole	Softball	Female
Ben	Football	Male
Raquel	Track—Thrower	Female
David	Golf	Male
Corinne	Track—Distance Runner	Female
Matthew	Cheer	Male
Jade	Softball	Female
Alex	Strength & Conditioning Coach	Male

Table 1. Summary of participants.

FINDINGS

Our findings are presented as a description of particular themes uncovered with respect to participants' data literacy practices, followed by a 'tension' that reveals how the context of Division I athletics structures those practices.

Tension 1: Staff Analytics Practices and Relevance to Athletes

State U staff gather an enormous amount of team- and individual-level data about student-athletes, but our findings suggest players really only attend to the latter. Data gathered and analyzed differ by role. The *strength and conditioning coach* is responsible for developing and overseeing the fitness regimes specific to different sports. *Head coaches* and, in some sports, *position coaches* work with athletes on their actual play of the game—developing strategies for performance in the sport (e.g., plays, techniques) for the overall team and specific player positions respectively. *Sports medicine trainers* keep track of players' overall health, focusing on injury prevention, diagnosis and rehabilitation. We identified two specific types of data literacy practices these professionals carry out.

Strength & Training Analyses. First, Alex, the strength trainer for two sports at State U spoke at length about ways he tracks athletes' training and performance data and uses this data to make decisions about student-athletes' training regimens. He is responsible for keeping track of how much athletes are working during games and practices. He uses some specialized technologies, such as GPS trackers. Less frequently, he also uses survey apps to gather more subjective data from athletes (e.g., pain, fatigue levels). However, the technology Alex uses most is Excel. He gathers GPS and other performance data into Excel; he develops general workout plans for groups of athletes using Excel formulas to individualize each athlete's plan (e.g., calculating specific numbers for weights, reps, etc.) on 'lift cards' given to athletes.

In determining athletes' workout plans, Alex refers back to performance and training data to consider which players need more conditioning (e.g., if their GPS trackers show lower overall mileage or average speeds) and to monitor athletes' training loads for spikes or peaks in performance to prevent injuries. He often shares this data with sports medicine coaches who monitor players' reports to prevent injury or to assess rehabilitation data with a players' pre-injury performance.

Team Level Analyses. Second, while it was not a primary concern for athletes themselves, coaches, strength trainers and other staff focused on team-level analyses. Softball player Jade talked about how a staff member records statistics for each play of their softball games. Her coach then calculates the data to see how the team measured up to their goal of being "92% routine on defense"—referring to the rate at which the defense successfully executed mundane plays. Similarly, Alex provides both team and individual reports of every game and practice to each of the team coaches with whom he works. These Excel reports provide team and positional averages for games and practices. For example, he looks at the average distance players covered in a game as well as the average speeds. He and the coaches use this data to make sure that athletes' speeds and distances during practice are on par with their speeds in the game that week. Additionally, coaches will sometimes request special analyses from the strength trainer to inform their coaching decisions (e.g., comparing players' performance metrics to see if a substitution can be made or to see if an individual player was working hard enough as compared to the team).

While coaches, trainers, and other staff use individual and team-level data to inform training and coaching decisions, athletes are much less engaged in these practices. Trainers and coaches make strategic decisions to hide or emphasize certain data, which are often quite personal. However, athletes' low engagement with team data also may be in large part because athletes were more drawn to data they perceived to be directly relevant to their individual efforts. For example, while the "routine-on-defense" metric is communicated to Jade and her teammates, it is not clear that she bought into its relevance, telling us, "[Our head coach] just loves that number for some reason. She's always like, '92%!' I don't really get it, but that's what she likes." Similarly, some teams (e.g., football) have staff that record all plays and practices for film review. However, Ben, a football player, said he only watches the film that directly involves him. Even then, he focuses on those particular plays in which he feels his role matters, ignoring, for example, receiving plays where his role is less impactful. These findings suggest that to expose student-athletes to a range of data literacy practices relevant to their sports play, athletes need to see how team-level analyses and coaching/training decisions are relevant to their own individual play.

Tension 2: Athletes' Agency vs. Team Needs

Interviews revealed that athletes' agency with respect to data collection and analysis was inversely proportional to the resources available to their team. That is, students in lower-revenue sports with less technological or staff support from the university had more freedom to choose what to measure, how, and why, compared to peers in higher-revenue sports with expensive equipment and extensive staff.

Participants' competitiveness drove data collection and analysis in their sports play and personal lives. Seth and his wrestling teammates, for example, dutifully filled out their weight-lifting progress on 'lift cards' assigned to them by strength trainers each week. This progress (or the lack thereof) was visible on a 'big board' which Seth described as a low-tech white board in the wrestling room that was updated weekly to showcase each team member's training and match-day statistics (e.g., wins, losses, technical fails, color-coded weight range targets). Their goals were visible during every moment of training. Meeting those goals required around-the-clock effort. Seth recalled uniting with fellow athletes across social media to make fun of an NCAA 'Day in the Life of a Student-Athlete' promotional video that they thought edited out all the stress, pain, and sweat of their college careers.

Beyond the training room or playing field, every athlete we interviewed also mentioned tracking their sleep and, sometimes, their steps. Most also carefully tracked their diets, though this was most intense—sometimes to a dangerous degree—in wrestling and women's track. In women's track, 'pace' had been internalized over years of practice. Lei said pace became "habitual." This meant that runners would sometimes disagree with the pace set by a coach in a particular training session. If it was too slow, a group of runners might collectively decide to beat it and enjoy the feeling of exceeding expectations. If it was too fast, they might perform something of a work slowdown and collectively slow their pace to avoid injury and send a message to their coach that the expectations were too high relative to what their bodies could take at the moment.

Wrestling and women's track are heavily subsidized by the university. They do not draw the crowds—let alone the TV deals and corporate sponsorships—of basketball or football. Football player Ben, for example, did not fill out his own weight card—strength coaches dedicated to football specifically did that for him. Nor was his weight measurement as simple as climbing on a scale. Instead, BodPod full-body scanners were used to assess not just total body weight but fat versus water weight and bone density. The interpretation of the scans was affected by one's position within the organization. At the end of the season, Ben's coach called every player into his office and showed them their body-fat percentages. He said Ben's was too high and it would affect his performance. The team nutritionist informed Ben the coach hadn't correctly interpreted the BodPod data. There was an important social effect beyond the concrete findings of the data: The coach's interpretation was established as authoritative, and Ben understood his data was to be interpreted for him, not by him. And while, as we noted, Ben could make some choices about which film to review at the team level (e.g., watching receivers play or not), his own individual film assignments were very clearly demarcated by staff using the Thundercloud video storage and replay software.

Football is perhaps on the extreme end of this spectrum: Flush with resources, closely managed by staff, with fewer choices for players. Women's track might be on the other end. Other sports sit in the middle. Softball, for example, left a lot of training and film review to students. But Nicole, a softball player, still reminded us that a degree of surveillance came with the territory, when, for example, she paid for meals with a special debit card loaded and monitored by State U athletics staff. "You don't have to worry about paying because it goes through athletics," she said. "But the nutritionist can see all of your receipts, so they see like what you buy, so if you go and say, hey, she'll be like 'Lay off the french fries' or something." Golf, both men's and women's, provided an interesting example of high levels of data collection with low levels of institutional investment. Both golfers' families provided them with personal coaches from a young age. This meant that David and Amanda had access to radar-trackers and film review technology worth tens of thousands of dollars, far more advanced than State U's equipment. But because the equipment was paid for by a personal coach, who was hired by their families, how they used it, and how it changed their game was up to them—not their State U coach.

Tension 3: The Personal Context Motivates and Drives Data Practices but Athletes Need Help

Student-athletes were highly motivated to analyze data—particularly as it related to their own individual play. However, we also observed that participants often needed help with the complexities of the required analyses and so their coaches would provide this scaffolding. Our participants had a well-honed sense for the context of data collection. They were analyzing data to understand and inform their practice, play, and even daily habits (e.g., sleep, screen time). This depth of engagement was a natural result of how deeply personal their data was to them. Raquel, a thrower on the track team, complained of how keenly she felt an error in a public tournament database because, "Now it's on my name." Indeed, data literacy practices are deeply embedded into players' game play. Athletes in ball sports (e.g., softball, golf) talked about data on the trajectories of their hits and throws. Speed and power athletes (e.g., runners, wrestlers) closely strategized around their weight and nutrition. Athletes in sports that start and restart (e.g., American football) analyze video to assess and adjust their form. All participants paid close attention to trends in their performance and fitness data. Two particular analytical strategies stood out.

Analysis of Different Levels of Data. First, athletes described their analysis of different scales of data as their data is measured from single performances or moves (e.g., gathering projectile data from a single golf hit), whole games, entire seasons, and over the course of their athletic careers. Athletes described how they would review these data over these different time-scales in order to decide whether a meaningful trend in their performance existed.

Comparisons to Other Players. Second, student-athletes placed a lot of attention on comparisons to other players, particularly professional players of their sport and their direct competitors whose data they obtained from public websites. If they were not measuring up as well as they thought they would, they would begin to figure out which data were most important to consider in their comparisons. For example, Amanda, a golfer, explained that she used "the percentage of makes at several distances on the green" to compare herself to professional players. This helped her realize that even professional players missed certain types of shots, especially at further distances and helped her to feel better if she missed one. Similarly, Raquel recognized that there were numerous throwers in track and field she could compare herself to with various statistics. But she realized it was most useful to focus on the data of her nearest competitor at the next meet because that comparison was the most actionable. This helped set an achievable benchmark for each meet. Raquel would know, for example, when she did not need to risk over-extending herself to reach a personal best, because the most meaningful metric was just the other woman's average throw.

Athletes also described the complexity of these comparisons, and the risks of overlooking relevant criteria (e.g., years of experience, different types of events or performance data). Raquel recognized the importance of considering experience alongside performance data: "For me, sometimes you just have to step back and look at the bigger picture, 'Oh, I'm a freshman throwing 14 meters, she's a freshman throwing 16 meters. What's the difference?' I'm actually 18 years old, a true freshman, she's been throwing since she was six and is 20. So, she has all these additional years on me, maybe I'll be at her place when I'm 20, too."

Coaches' Scaffolding. While student-athletes looked at their data over different time scales and thought reflectively about comparing themselves to other players, athletes often needed support to engage in these data analytic practices. Coaches helped athletes determine which data were important to consider, when, and why. Coaches' scaffolding also involved helping athletes face performance data they would prefer to avoid—plays that went wrong or embarrassing statistics. Caitlin, a thrower on the track team said, "I had a throw that got away in my third round, and I just sat down with my coach, because I didn't want to watch it for a couple of days because I was still salty about it, but I sat down and watched it with him. We scrolled through it and we did like a technical analysis of what went wrong at the end."

Often, coaches' guidance helped athletes to see their athletics performances in a new, more encouraging light. Amanda's golf coach, for example, helped her reinterpret a recent slide in performance: "He came up to me and said, 'Last year you were averaging at 54 and now you're averaging at 59. That's four meters difference.'" This helped her realize that it wasn't her personal record that mattered for her long-term success, but her year-over-year improvements.

Tension 4: The Need for Data Versus Stepping Back

Our participants saw data as an integral part of their sports play, particularly in the intensely competitive environment of Division I sports. Lei described what this meant for a runner, "Basically every girl on our team was the number one girl in their state or was the number one girl in their county, or district [in high school]. And then you come here and all of a sudden you're the 12th runner at best." When the competition is so stiff, Lei reflected, "You have to start counting the little things." But it is precisely this demanding level of competition that led athletes to step back from in-depth analyses at moments of high stress, for fear of inaction or, worse, undertaking dangerous behaviors to meet certain metrics.

Athletes' data was key to making meaningful enhancements to their sports play. Yet, while players saw the value of studying their performance data, across sports, student-athletes talked about times when they specifically needed to step away from their data so that they would not, in Seth's words, "get in their head" about it or overthink it. The flood of available data could be a distraction from their sports play. In fact, even though golf coaches expected them to be intimately familiar with their numbers at practice, when it came to game time, David said, "They just want you to play. Just play. Don't get bogged down in all the numbers and details. … That's why we practice so we don't have to think anymore."

Players either chose to ignore the data themselves or had coaches tell them to ignore it during a game or tournament. Student-athletes talked about stepping away from stats, film, and their bodyweight measurements during high stress periods: right before or during tournaments or as they were having particularly good or bad streaks. Seth actively stepped away from tracking his weight during or a couple of days before a wrestling competition, when he needed to make a certain weight to compete in his weight class. During her three-day golf tournaments, Amanda's coach did not allow her to visit the public website that showcased player performance data. Jade did not watch video of her softball gameplay, particularly as her patterns of play during the season were starting to emerge (e.g., if she was playing well or playing badly). Instead, she opted to record herself during practice for assessments of her technique.

However, an athlete's ability to take this step away from data was often mitigated by their coaches. In training, Raquel and her position coach ignored her throwing distances in favor of improving her technique. They were chastised by the head coach who only wanted to see progress: "[My head coach is] like 'why aren't you throwing farther out of high school? We brought you in with the expectation you would throw further than you did.' So it's mainly my individual coach who's telling me not to pay attention to the data because he sees how I am growing outside the number value, outside of my actual throw marks."

Sometimes participants stepped away from their data for their own health and well-being. We specifically observed this when it came to counting calories. Runners and wrestlers in particular knew that obsession with weight and nutrition data frequently had dangerous results. One of Lei's roommate's, also a runner, struggled with orthorexia, a condition that includes obsessive behavior around eating healthy and is associated with symptoms that frequently occur with anorexia and other eating disorders [15]. Lei observed that such eating disorders are common for female runners and within other individual sports (e.g., gymnastics) in which bodyweight is assumed to be a performance indicator, e.g., lighter runners are assumed to run faster. She stopped counting calories herself for that reason.

Our findings suggest that while keeping track of player data may be important to the team, student-athletes also need to balance analyses of their performance with the need to focus on their play in the moment, and to focus on their health in the long term. Their coaching staff played significant roles in their ability to do this, sometimes helping them focus by forbidding certain data, sometimes hurting by insisting on the importance of other data.

Tension 5: Family Atmosphere Can Be Disrupted by Data Sharing

Student-athletes across sports talked about their close-knit relationships with other athletes on their teams. Several described the social environment as a 'family' atmosphere, deeply appreciating that aspect of their team experience. Caitlin said of the other State U throwers, "Even though you're competing against each other, you're still kind of like a family. Everyone's got each other's back no matter what."

Lei explained that girls on her track team live together, eat dinner together and do homework together. Golfer Amanda, recognized that teammates' friendly communication is important for bringing out the best in one another. Similarly, Alex, the strength and conditioning coach, emphasized that building this camaraderie generated "buy in" for his training regime: "If you don't really act interested in what's going on in their life, they tend not to buy into what they're doing with you in the weight room."

While the family setting is an integral part of their team experience, athletes across sports talked about how it could be disrupted by the bare fact of competition. Athletes were always either directly competing with each other in the same event or indirectly for a starting spot. Because of this, they were often uncomfortable sharing personal data with one another. As Amanda said, "It's like a bunch of friends, honestly, we're all pretty close. But I think it's like an individual and team sport at the same time. It's pretty hard to talk about really specific stuff, like numbers, since we're all competing with each other to make the lineup…"

Instead, players were much more comfortable sharing data with people with whom they were not as close. Seth does not discuss the wrestling team's 'big board' with his teammates, even though he and his wrestling teammates use this analog tool to keep track of one another's accomplishments, struggles, and progress. Even his coaches did not call too much attention to it, he observed, only occasionally bringing it to the team's attention to highlight players who were doing well. And while golfers, runners, and throwers were hesitant to compare their numbers with one another, it was a normal part of their lives for fans and analysts to pore over their public data.

DISCUSSION

Supporting Athletes' Data Literacy

Our findings suggest student-athletes are engaging quite authentically in the first three components of data literacy [39, 50, 60]. Mapping back to those components, they have a heightened awareness of data and its role in sports because of its prevalence in their day to day sport lives. They routinely identify relevant data based on their needs for self-assessments (e.g., understanding when to use specific statistics and how to include other relevant data of interest in their analyses). They know exactly where to go to find these data and are heavily engaged in evaluating and analyzing their own personal data, interpreting it to make adjustments in their technique and performance. With respect to their understanding of the context of data, they leverage their deep understanding of their sport context and manage the complexity of data by narrowing their focus from the wide range of data to specific metrics most relevant for their comparisons and self-assessments.

The Division I context scaffolds and facilitates these practices. For example, we found that coaches help athletes figure out the important metrics to consider, know, and face up to, and know when to look at data and how to compare them. However, there are also ways in which their data literacy practices are limited by the organizational infrastructure of Division I athletics. First, there is an entire body of staff members (i.e., strength and training coaches, sport coaches, sports medicine trainers) that are responsible for athletes' day-to-day sports training and game play. Often there are data-driven decision-making practices that are not readily apparent to athletes. For example, though the athletic trainers work hard to develop data-informed training activities workouts are handed to athletes on individualized lift cards with the work already done for them and masked from view.

Second, while some athletes could communicate to their coaches when a workout was too tough or easy, even in those cases athletes did not report open conversations in which they could discuss workout plans and negotiate with coaches. This is in contrast with Rapp & Tirabeni's [52] findings with elite athletes outside of universities. They reported close consultation between coaches and athletes to determine workout plans. This limited discussion in the collegiate context likely inhibits some of the learning benefits Rapp & Tirabeni [52] report elite athletes experiencing regarding their sport performance (i.e., why certain decisions were made, how training impacts fitness and performance, etc.). However, the limited discussion also inhibits student-athletes' access to other relevant aspects of data literacy, including the process of making data-driven decisions, understanding data management (e.g., data storage, cleaning) and ways of communicating data (e.g., how trainers create presentations for coaching staff) [50, 60].

While concealing of data literacy practices is likely practical for very busy individuals and teams, there is room to consider how to create opportunities for life-relevant learning for athletes (and potentially even non-athletes interested in sports), particularly given the call for athletics programs to consider learning and development opportunities for players through their sports play [68, 69]. Our analysis reveals some opportunities for such learning in the context of day to day sports experiences.

First, we might find ways to present some of the data driven training and playing decisions to players. For example, designers could create systems that include data guiding workout decisions on lift cards and practice plans given to student-athletes. Lift cards could show athletes how their practice workouts have changed over time and how that has corresponded to their overall performance (e.g., "You squatted 10 pounds more this week and your average speed during soccer games has increased by 0.5 mph"). Such information could spark more discussion between players and coaches about workout and game play decisions.

Second, players' deep engagement with personal data suggests that data visualizations that make visible links between players' individual data and the performance of the team may help players engage more deeply with team level analytics. Additionally, such interactions may help them better understand their coaches' team-level decisions and emphases—or at least be better able to communicate data-driven reasons for their objections to them.

However, such designs for data literacy must be balanced with concerns athletes expressed about getting lost in their data. Student-athletes need to step away from data from time to time and their needs for managing data obsessions must be taken into account. This finding is similar to findings in prior personal informatics and quantified-self research that amateur athletes avoided data so that they could focus on the joy of the sport, nature, and the physical experience of activity. Our findings within the collegiate context suggest specifically that designs for student-athletes' data literacy must help them face and not be wary of "bad plays" (e.g., balancing presentation of data from good and bad plays, helping athletes see their data in hopeful and encouraging ways) and help them to recognize and balance between too much focus on data (e.g., through monitoring and visualizing their focus on data in ways similar to the iPhone's Screentime features). In designing such data systems for sports organizations, it is imperative to understand and account for the diversity of roles within the organization (e.g., differentiating access to personal data by role).

Taking Into Account the Social Dynamics on the Team

Division I athletics approaches a 'total institution' [24] where the campus guides every aspect of a player's life. Any design interventions that attempt to enhance performance or help students learn from their data needs to take into account the dynamics of this highly-structured social setting. The value of a certain kind of data, and what players and staff can do with it, will differ in different parts of the institution. These organization-level findings thus distinguish our study from other social analyses of player data that either focus on recreational sports or elite individual athletes.

There are many similarities between the athletes we study and those who play for pure recreation, particularly in their day-to-day experience of the sport and their data. Runners in our study, for example, have a very physical sense of their data similar to elite runners in [64]. But institutional policies and the high level of performance expected for a higher level of competition leads to higher-level differences.

Recreational athletes often use personal technology to organize themselves and their peers [64, 71]. Obviously, those peers are found for our athletes by their institution and the bodies that govern their sport dictate the equipment, spaces, and training conditions in which they work. Rapp & Tirabeni [52] do not analyze the organizational context of their elite athletes, but do note that they used their performance data to create personal brands and build social networks of elite athletes. For our Division I student-athletes, however, social media was not a branding exercise but an outlet for adolescent socialization; indeed Seth and others deliberately resisted the NCAA's efforts to sanitize and brand their athletic identities. Nor was performance data shared in search of community. Rather, it travelled along the channels of existing relationships maintained by the institution. Sometimes this meant not sharing data between friends. Sometimes it meant sharing data with a position coach to work on technique. Sometimes it meant discipline, even when the interpretation was 'off.' Raquel's reprimand from her head coach is one example. Ben's experience with his new head football coach was another. Each sport had a different social dynamic with respect to data sharing, but State U structured the channels in which those data flowed. Similarly, in recreational contexts, data-sharing and data analysis is a way for athletes to motivate each other to keep training [64, 71]. Our athletes do not need this motivation.

Through raw talent and years of training, they've self-selected into an elite level of competition. They are self-motivated and they are surrounded by support staff that maintain that motivation if it flags. Athletes support each other in recovery and in down times, but their competitiveness does not waver. Rather, it is the athletics staff who seem to collaborate most on data analysis. Sometimes this is because of competing interpretations that need to be worked out, like the football team's nutritionist versus their coach. Sometimes this is the result of a division of labor within a large organization. Coaches rely on trainers to provide up-to-date analyses of players' strength, speed, and endurance in order to structure practices and choose roles. For athletes to engage in these higher-level analyses of their personal data, coaches and staff may have to give up some of their control of the team routine.

In recreational settings, athletes often report an appreciation of the 'pain' or 'burn' that comes with hard training. Their data acts as a receipt for this endurance [64]. At the other end of the spectrum, Rapp & Tirabeni's [52] elite individual athletes hire personal coaches to find the right performance level across training tasks and meet their specific individual goals. And while that dynamic is reminiscent of our golfers, the majority of our athletes said that they did not have the agency to either choose a level of pain satisfying to them or analyze their data in cooperation with a single coach in order to design the perfect training regime. Instead, our athletes expressed their lack of agency within State U through wishes for greater boundaries between athletics and academics, and increased sensitivity on the part of staff to problems of burnout, exhaustion, and scheduling constraints. Their ability to learn from or act on their data was determined not just by the opportunities offered to them within the organization, but the opportunities that were foreclosed. Their weight or bone density were meaningful not just because of speed or performance but because of injuries and eating disorders. They measured their sleep not just to maximize their energy, but because their sleep was threatened by overstuffed schedules.

Not that this organizational context acted only as a restriction. Data became meaningful in the context of intensive social relationships, the team as family. Where elite athletes in [52] worked out alone or with coaches, and amateurs work out with friends and relatives, our athletes see their teammates every day—often several times a day, or even in their apartments. They were best friends, but also competitors. This complex social dynamic drove interesting data analytic practices, like the greater willingness to share data with strangers (i.e., the public) than roommates.

In recreational contexts, relationships drive social experiences—they are voluntary and initiated by athletes [64, 71]. Our participants all accepted that their game-day data were public. Recreational athletes certainly share and publicize their race times or game scores [64, 71]. And elite individual athletes may publicize their data as part of their personal brand [52]. But both of these practices are voluntary. Our athletes were often televised to millions of people, and they had little control over what data were collected or what was said about it. Recall, for example, Raquel's regret over a misplaced data point that she felt damaged her reputation. They still made choices about how to analyze that data, however. Most often this meant choosing to ignore or avoid it during particularly sensitive periods, sometimes at staff's encouragement, in order to focus on the task at hand. Performance data may be revealing, but those insights may not be equally useful to everyone—including the person being measured.

In designing new systems for student-athletes' data literacy our findings suggest two important considerations. First, effective data systems will change the information flow within the organization, thus adjusting inter-role relations. Second, designers must be cognizant of the limits of such interventions: Information asymmetries persist in part because of power differentials built into campus athletics. Coaches are adult authority figures employed by the university for the long term. Players are younger, uncompensated, and there for a short period.

CONCLUSION

Our findings have three major implications for research on HCI and Sports. First, personal informatics practices in sports settings could have an impact beyond the sport itself to include data literacy practices more broadly. Frameworks for data literacy can help us identify these practices and push them forward. Second, our findings suggest that the level of competition in Division I sports creates social and emotional dynamics that differentiate athletes' experiences with data even from elite athletes in other contexts. Finally, our findings indicate that HCI and sports researchers need more study and consideration of the influence of organizational contexts on data practices. Further work is needed to more deeply explore the impacts of organizational infrastructure within specific sports, across multiple institutions, and beyond collegiate contexts. Additionally, more ethnographic approaches are needed to further explain how organizational infrastructures impact data literacy practices on a moment-to-moment basis. This pilot work guides our next steps geared toward understanding such nuances to develop learning tools and experiences for data literacy in sports play. Beyond HCI and sports, student-athletes' experiences of data monitoring, particularly of health data (or even their academic data [63]), are similar to other workers' experiences of, e.g., biometrics. However, Division 1 athletics' high investment in athletes means they are likely the leading edge of trends that will filter out to the broader population as technologies cheapen and such managerial practices become more socially acceptable.

ACKNOWLEDGEMENTS

We thank our participants and reviewers, as well as Priya Kumar and Pramad Chundury for feedback on this paper.

REFERENCES

[1] Ackoff, R.L. 1989. From data to wisdom. *Journal of applied systems analysis*. 16, 1 (1989), 3–9.

[2] Baca, A. and Kornfeind, P. 2006. Rapid feedback systems for elite sports training. *IEEE Pervasive Computing*. 5, 4 (2006), 70–76. DOI:https://doi.org/10.1109/MPRV.2006.82.

[3] Bachynskyi, M., Oulasvirta, A., Palmas, G. and Weinkauf, T. 2014. Is motion-capture-based biomechanical simulation valid for HCI studies? Study and implications. *Conference on Human Factors in Computing Systems - Proceedings*. (2014), 3215–3224. DOI:https://doi.org/10.1145/2556288.2557027.

[4] Bal, B. and Dureja, G. 2012. Hawk Eye: A Logical Innovative Technology Use in Sports for Effective Decision Making. *Sport Science Review*. 21, 1–2 (2012), 107–119. DOI:https://doi.org/10.2478/v10237-012-0006-6.

[5] Battle, L., Chang, R. and Stonebraker, M. 2016. Dynamic prefetching of data tiles for interactive visualization. *Proceedings of the 2016 International Conference on Management of Data* (2016), 1363–1375.

[6] Blumer, H. 1954. What is wrong with social theory? *American sociological review*. 19, 1 (1954), 3–10.

[7] Bourdon, P.C., Cardinale, M., Murray, A., Gastin, P., Kellmann, M., Varley, M.C., Gabbett, T.J., Coutts, A.J., Burgess, D.J., Gregson, W. and Cable, N.T. 2017. Monitoring athlete training loads: Consensus statement. *International Journal of Sports Physiology and Performance*. 12, (2017), 161–170. DOI:https://doi.org/10.1123/IJSPP.2017-0208.

[8] Caine, K. 2016. Local standards for sample size at CHI. *Proceedings of the 2016 CHI conference on human factors in computing systems* (2016), 981–992.

[9] Cardinale, M. and Varley, M.C. 2017. Wearable Training-Monitoring Technology : Applications , Challenges , and Opportunities. (2017), 55–62.

[10] Charmaz, K. 2011. Grounded theory methods in social justice research. *The Sage handbook of qualitative research*. 4, 1 (2011), 359–380.

[11] Chiapello, E. and Fairclough, N. 2002. Understanding the new management ideology: a transdisciplinary contribution from critical discourse analysis and new sociology of capitalism. *Discourse & society*. 13, 2 (2002), 185–208.

[12] Choe, E.K., Lee, N.B., Lee, B., Pratt, W. and Kientz, J.A. 2014. Understanding quantified-selfers' practices in collecting and exploring personal data. *Conference on Human Factors in Computing Systems - Proceedings*. (2014), 1143–1152. DOI:https://doi.org/10.1145/2556288.2557372.

[13] Clegg, T. and Kolodner, J. 2014. Scientizing and Cooking: Helping Middle-School Learners Develop Scientific Dispositions. *Science Education*. 98, 1 (Jan. 2014), 36–63. DOI:https://doi.org/10.1002/sce.21083.

[14] Corbin, J. and Strauss, A. 2008. *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage.

[15] Donini, L.M., Marsili, D., Graziani, M.P., Imbriale, M. and Cannella, C. 2005. Orthorexia nervosa: validation of a diagnosis questionnaire. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*. 10, 2 (2005), e28--e32.

[16] Esposti, S.D. 2014. When big data meets dataveillance: The hidden side of analytics. *Surveillance and Society*. 12, 2 (2014), 209–225.

[17] Fairclough, N. 2013. *Critical discourse analysis: The critical study of language*. Routledge.

[18] Fountain, J. and Finley, P. 2011. Academic Clustering: A Longitudinal Analysis of a Division I Football Program. *Journal of Issues in Intercollegiate Athletics*. 4, (2011), 24–41.

[19] Fountain, J.J. and Finley, P.S. 2009. Academic majors of upperclassmen football players in the Atlantic Coast Conference: An analysis of academic clustering comparing white and minority players. *Journal of Issues in Intercollegiate Athletics*. (2009).

[20] Frey, J. 2011. Institutional control of athletics: An analysis of the role played by presidents, faculty, trustees, alumni, and the NCAA. *Sports & Athletics in Higher Education*. J.W. Scatterfield, R.L. Huges, and K. Kearney, eds. Pearson Learning Solutions.

[21] Fritz, T., Huang, E.M., Murphy, G.C. and Zimmermann, T. 2014. Persuasive technology in the real world. (2014), 487–496. DOI:https://doi.org/10.1145/2556288.2557383.

[22] Galbraith, J.R. 2014. Organizational Design Challenges Resulting From Big Data. *Journal of Organization Design*. 3, 1 (2014), 2. DOI:https://doi.org/10.7146/jod.8856.

[23] Gee, J.P. 2004. Discourse analysis: What makes it critical? *An introduction to critical discourse analysis in education*. Routledge. 49–80.

[24] Goffman, E. and others 1961. On the characteristics of total institutions. *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates*. Anchor Books. 3– 124.

[25] Greene, D. and Shilton, K. 2018. Platform privacies: Governance, collaboration, and the different meanings of "privacy" in iOS and Android development. *new media & society*. 20, 4 (2018), 1640–1657.

[26] Guest, G., Bunce, A. and Johnson, L. 2006. How many interviews are enough? An experiment with data saturation and variability. *Field methods*. 18, 1 (2006),

59–82.

[27] Häkkilä, J., Helander, V., Jamoido, D. and Colley, A. 2018. Designing an Interactive Ice Skating Dress for Young Athletes. *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (2018), 734–737.

[28] Harper, S.R. 2018. Black male student-athletes and racial inequities in NCAA Division I revenue-generating college sports: 2018 edition. (2018), 24.

[29] Haugen, T. and Buchheit, M. 2016. Sprint Running Performance Monitoring: Methodological and Practical Considerations. *Sports Medicine*. 46, 5 (2016), 641–656. DOI:https://doi.org/10.1007/s40279-015-0446-0.

[30] Hemerly, J. 2013. Public policy considerations for data-driven innovation. *Computer*. 46, 6 (2013), 25–31. DOI:https://doi.org/10.1109/MC.2013.186.

[31] Herdal, T., Pedersen, J.G. and Knudsen, S. 2015. Designing Personal Visualizations for Different People: Lessons from a Study with Elite Soccer Teens. *2015 IEEE VIS Workshop on Personal Visualization: Exploring Data in Everyday Life* (2015).

[32] Lazar, A., Koehler, C., Tanenbaum, J. and Nguyen, D.H. 2015. Why we use and abandon smart devices. *UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. (2015), 635–646. DOI:https://doi.org/10.1145/2750858.2804288.

[33] Lee, V.R. 2013. The Quantified Self (QS) movement and some emerging opportunities for the educational technology field. *Educational Technology*. 53, 6 (2013), 39–42.

[34] Lee, V.R. and Drake, J. 2013. Digital physical activity data collection and use by endurance runners and distance cyclists. *Technology, Knowledge and Learning*. 18, 1–2 (2013), 39–63.

[35] Lee, V.R., Drake, J.R., Cain, R. and Thayne, J. 2015. Opportunistic uses of the traditional school day through student examination of Fitbit activity tracker data. *Proceedings of the 14th International Conference on Interaction Design and Children* (2015), 209–218.

[36] Li, I., Dey, A. and Forlizzi, J. 2010. A stage-based model of personal informatics systems. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10* (New York, New York, USA, Apr. 2010), 557.

[37] Li, I., Dey, A.K. and Forlizzi, J. 2011. Understanding my data, myself: supporting self-reflection with ubicomp technologies. *Proceedings of the 13th international conference on Ubiquitous computing* (2011), 405–414.

[38] Losada, A.G., Theron, R. and Benito, A. 2016. BKViz: A basketball visual analysis tool. *IEEE Computer Graphics and Applications*. 36, 6 (2016), 58–68. DOI:https://doi.org/10.1109/MCG.2016.124.

[39] Maybee, C. and Zilinski, L. 2015. Data informed learning: A next phase data literacy framework for higher education. *Proceedings of the Association for Information Science and Technology*. 52, 1 (2015), 1–4. DOI:https://doi.org/10.1002/pra2.2015.1450520100108.

[40] NCAA: *ncaa.org*.

[41] NCAA 2019. *NCAA Demographics Database*.

[42] Ness, K. 2012. Constructing masculinity in the building trades:'Most jobs in the construction industry can be done by women.' *Gender, Work & Organization*. 19, 6 (2012), 654–676.

[43] Nguyen, L.N.N., Rodríguez-Martín, D., Català, A., Pérez-López, C., Samà, A. and Cavallaro, A. 2015. Basketball activity recognition using wearable inertial measurement units. *ACM International Conference Proceeding Series*. 07-09-Sept, (2015). DOI:https://doi.org/10.1145/2829875.2829930.

[44] Nylander, S. and Tholander, J. 2014. Designing for movement - The case of sports. *ACM International Conference Proceeding Series*. (2014), 130–135. DOI:https://doi.org/10.1145/2617995.2618018.

[45] Ono, J.P., Gjoka, A., Salamon, J., Dietrich, C. and Silva, C.T. 2019. HistoryTracker: Minimizing human interactions in baseball game annotation. *Conference on Human Factors in Computing Systems - Proceedings*. (2019), 1–12. DOI:https://doi.org/10.1145/3290605.3300293.

[46] Owens, S.G. and Jankun-Kelly, T.J. 2013. Visualizations for Exploration of American Football Season and Play Data. *1st Workshop on Sports Data Visualization, IEEE VIS*. (2013).

[47] Paule, A. 2010. Gaining equity in all the wrong areas: An analysis of academic clustering in women's Division I basketball. *Scholarly Conference on College Sport, Chapel Hill, NC* (2010).

[48] Perin, C., Vuillemot, R., Fekete, J., Perin, C., Vuillemot, R., Kick-off, J.F.S.A., Perin, C., Vuillemot, R., Fekete, J. and Member, S. 2013. SoccerStories : A Kick-off for Visual Soccer Analysis To cite this version : SoccerStories : A Kick-off for Visual Soccer Analysis. *IEEE Transactions on Visualization and Computer Graphics*. 19, 12 (2013), 2506–2515.

[49] Pileggi, H., Stolper, C.D., Boyle, J.M. and Stasko, J.T. 2012. Snapshot: Visualization to propel ice hockey analytics. *IEEE Transactions on Visualization and Computer Graphics*. 18, 12 (2012), 2819–2828.

[50] Prado, J.C. and Marzal, M.Á. 2013. Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*. 63, 2 (2013), 123–

134. DOI:https://doi.org/10.1515/libri-2013-0010.

[51] Provost, F. and Fawcett, T. 2013. Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*. 1, 1 (2013), 51–59. DOI:https://doi.org/10.1089/big.2013.1508.

[52] Rapp, A. and Tirabeni, L. 2018. Personal informatics for sport: Meaning, body, and social relations in amateur and elite athletes. *ACM Transactions on Computer-Human Interaction*. 25, 3 (2018). DOI:https://doi.org/10.1145/3196829.

[53] Read, A. 2016. Monitoring Recovery in Collegiate Strength and Conditioning. *Theses, Dissertations, Professional Papers*. 10678, (2016).

[54] Rhoden, W.C. 2010. *Forty million dollar slaves: The rise, fall, and redemption of the Black athlete*. Broadway Books.

[55] Rooksby, J., Rost, M., Morrison, A. and Chalmers, M. 2014. Personal tracking as lived informatics. *Proceedings of the SIGCHI conference on human factors in computing systems* (2014), 1163–1172.

[56] Sands, W.A., Kavanaugh, A.A., Murray, S.R., McNeal, J.R. and Jemni, M. 2017. Modern techniques and technologies applied to training and performance monitoring. *International Journal of Sports Physiology and Performance*. 12, (2017), 63–72. DOI:https://doi.org/10.1123/ijspp.2016-0405.

[57] Sapp, L. and Vaughan, K.T.L. 2017. Connecting the Libraries and Athletics through Instruction and Outreach. *Medical Reference Services Quarterly*. 36, 2 (2017), 187–195. DOI:https://doi.org/10.1080/02763869.2017.1293999.

[58] Schaefer, S.E., Ching, C.C., Breen, H. and German, J.B. 2016. Wearing, thinking, and moving: testing the feasibility of fitness tracking with urban youth. *American Journal of Health Education*. 47, 1 (2016), 8–16.

[59] Sellers, R.M., Kuperminc, G.P. and Damas, A. 1997. The college life experiences of African American women athletes. *American Journal of Community Psychology*. 25, 5 (1997), 699–720.

[60] Shields, M. 2005. Information Literacy, Statistical Literacy, Data Literacy. *IASSIST Quarterly*. 28, 2 (2005), 6. DOI:https://doi.org/10.29173/iq790.

[61] Sisneros, R. and Moer, M. Van 2014. Expanding Plus-Minus for Visual and Statistical Analysis of NBA Box-Score Data. *Workshop.Sportvis.Com*. November (2014), 7.

[62] Steen, L.A. 2000. Numeracy: The new literacy for a data-drenched society. *Educational Leadership*. 58, 3 SUPPL. (2000), 10–11.

[63] Sun, K., Mhaidli, A.H., Watel, S., Brooks, C.A. and Schaub, F. 2019. It's My Data! Tensions Among Stakeholders of a Learning Analytics Dashboard. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (2019), 594.

[64] Tholander, J. and Nylander, S. 2015. Snot, sweat, pain, mud, and snow: Performance and experience in the use of sports watches. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (2015), 2913–2922.

[65] Verhagen, E. and Bolling, C. 2015. Protecting the health of the @hlete: How online technology may aid our common goal to prevent injury and illness in sport. *British Journal of Sports Medicine*. 49, 18 (2015), 1174–1178. DOI:https://doi.org/10.1136/bjsports-2014-094322.

[66] Wang, J. 2016. New political and communication agenda for political discourse analysis: Critical reflections on critical discourse analysis and political discourse analysis. *International Journal of Communication*. 10, (2016), 19.

[67] Watson, H.J. 2014. Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems*. 34, 1 (2014), 1247–1268.

[68] Weight, E.A., Cooper, C. and Popp, N.K. 2015. DIOrgStructure. (2015), 510–522.

[69] Weight, E.A., Navarro, K.M., Smith-Ryan, A. and Huffman, L.T. 2016. Holistic Education through Athletics: Health and Health-Literacy of Intercollegiate Athletes and Active Undergraduate Students. *Journal of Higher Education Athletics & Innovation*. 1, 1 (2016), 38–60. DOI:https://doi.org/10.15763/issn.2376-5267.2016.1.1.38-60.

[70] Wilson, B.D. 2008. Development in video technology for coachingy. *Sports Technology*. 1, 1 (2008), 34–40. DOI:https://doi.org/10.1080/19346182.2008.9648449.

[71] Woźniak, P.W., Fedosov, A., Mencarini, E. and Knaving, K. 2017. Soil, rock and snow: On designing for information sharing in outdoor sports. *DIS 2017 - Proceedings of the 2017 ACM Conference on Designing Interactive Systems*. (2017), 611–623. DOI:https://doi.org/10.1145/3064663.3064741.

[72] Zimmermann-Niefield, A., Turner, M., Murphy, B., Kane, S.K. and Shapiro, R.B. 2019. Youth learning machine learning through building models of athletic moves. *Proceedings of the 18th ACM International Conference on Interaction Design and Children, IDC 2019*. (2019), 121–132. DOI:https://doi.org/10.1145/3311927.3323139.