Received: 1 December 2021 | Revised: 1 December 2022 | Accepted: 2 December 2022

DOI: 10.1002/tea.21842

RESEARCH ARTICLE

JRST | WILEY

Data everyday as community-driven science: Athletes' critical data literacy practices in collegiate sports contexts

Tamara L. Clegg[1] | **Keaunna Cleveland**[1] | **Erianne Weight**[2] | **Daniel Greene**[1] | **Niklas Elmqvist**[1]

[1]College of Information Studies, University of Maryland, College Park, Maryland, USA

[2]Department of Exercise and Sports Science, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Correspondence
Tamara L. Clegg, College of Information Studies, University of Maryland, College Park, MD, USA.
Email: tclegg@umd.edu

Abstract

In this article, we investigate the community-driven science happening organically in elite athletics as a means of engaging a community of learners—collegiate athletes, many of whom come from underrepresented groups—in STEM. We aim to recognize the data literacy practices inherent in sports play and to explore the potential of *critical data literacy practices* for enabling athletes to leverage data science as a means of addressing systemic racial, equity, and justice issues inherent in sports institutions. We leverage research on critical data literacies as a lens to present case studies of three athletes at an NCAA Division 1 university spanning three different sports. We focus on athletes' experiences as they engage in critical data literacy practices and the ways they welcome, adapt, resist, and critique such engagements. Our findings indicate ways in which athletes (1) readily accept data practices espoused by their coaches and sport, (2) critique and intentionally disengage from such practices, and (3) develop their own new data productions. In order to support community-driven science, our findings point to the critical role of athletics' organizations in

promoting athletes' access to, as well as engagement and agency with data practices on their teams.

KEYWORDS
critical data literacy, data justice, data science, sports and learning

1 | INTRODUCTION

From investigating, diagnosing, and treating injuries to developing winning practices and playbooks that maximize talent on a team's roster, data science plays a key role in guiding athlete training and competition. The science gets more advanced at collegiate and professional levels where athletics is a high-revenue industry and teams seek any resources that facilitate a competitive edge. Athletes analyze film[1] of themselves and competitors, track their performance metrics, and outwit their opponents with game-time decisions deriving from hours of analysis synthesized with in-the-moment observations. Hence, data cleaning, collection, analysis, and interpretation, as well as data-driven decision-making—all important data literacy skills (Maybee & Zilinski, 2015; Prado & Marzal, 2013)—fuel science in sports. Although analytics are a critical and ingrained aspect of this industry, the data literacy skills athletes develop and use are often unrecognized and underdeveloped (Weight et al., 2016; Weight et al., 2020; Weight & Huml, 2016).

Within this context, we situate our stance on community-driven science (CDS). First, we aim to understand the CDS happening in places we might not recognize, especially in organizations and institutions where learners from minoritized and resource-constrained communities are deeply engaged. In order to broaden our conceptualization of science in socially just ways, we take an asset-based perspective (Kretzmann & McKnight, 1996; Mathie & Cunningham, 2003), acknowledging that it is imperative to identify and characterize CDS in these settings so the science engagement of all learners is acknowledged and built upon. In this article, we investigate the CDS happening authentically in elite athletics as a means of engaging a community of learners—collegiate athletes, many of whom come from underrepresented groups in STEM (Comeaux, 2018). Our first research question in this context is therefore, *what critical data literacy practices are athletes engaging in through their participation in athletics?*

Second, we acknowledge that communities come with social, political, and historical contexts that community members must navigate. These contexts both empower and constrain learners' science engagement, particularly as it relates to promoting learners' well-being and social good (Paris & Alim, 2014). In promoting CDS, we must therefore understand how learners engage in CDS within the sociopolitical historical contexts of their communities. Our second research context within the context of collegiate athletics is therefore *how do athletes' roles within the socio-historical political context of intercollegiate athletics shape their access to and engagement with critical data literacy practices on their teams*? *How do these practices inform the possibilities for CDS?*

The community in which this study of CDS is situated is the most competitive and commercialized level of collegiate athletics within the United States National Collegiate Athletics Association, Division 1, Power 5, where extensive budgets fund robust sports analytics.[2] This intense context fuels and hinders data practices for athletes, with disproportionate effects on young black men (Comeaux, 2018). This context and the dynamic struggle athletes face between their athletic and academic commitments (Weight et al., 2015) has led to mounting pressure for

examination of the athlete experience and consideration of educational development through their athletic experience (Weight et al., 2015; Weight et al., 2016).

These issues present an *opportunity* for athletes to explore technical, social, and political aspects of data science in a context in which they are deeply embedded. We aim to reframe CDS within elite collegiate athletics by leveraging it as an educational tool that can empower the athlete well beyond their sport play. Our goal is to build upon the community-driven data practices within athletics to help (1) athletes recognize the science expertise they develop through day-to-day experiences on their teams so that they can extend their expertise in other contexts (e.g., academics) and (2) them to critically reflect on these data practices so that they can use and/or fight against them for their own liberatory effects within athletics. This analysis is an initial step toward this goal.

This analysis addresses calls from science education, data science education and critical data literacies research for approaches, particularly at the collegiate level, that focus on the technical aspects of data science *and* the social, historical, and political contexts in which such efforts are situated (Philip et al., 2013; Wilkerson & Polman, 2020). Research in sports science and legal perspectives of data practices in high-revenue sports (Greenbaum, 2018; Jessop & Baker III, 2019) especially call for helping athletes to be suitably informed of data practices in athletics that drive intense collection, analysis, sharing, and decision-making with athletes' biometric and performance data.

We join a growing body of research focused on sports as rich contexts for STEM[3] learning, where most approaches focus on K-12 learners and athletes (Jones et al., 2020; Zimmermann-Niefield et al., 2019). We leverage critical data literacies as a lens to present case studies of three athletes' critical data literacy practices at a Division 1, elite university athletics program spanning three different sports. Our findings indicate ways in which athletes (1) readily accept data practices espoused by their coaches and sport and (2) critique and intentionally disengage from such practices, and (3) develop their own new data productions. These findings point to the importance of recognizing and building on the critical data literacies athletes develop through their sports play and point to the critical role of athletics' organizations in influencing athletes' access to, engagement, and agency with team data practices. Our findings point to the importance of CDS experiences that help learners understand and navigate power dynamics in their own community contexts.

2 | BACKGROUND

We draw on data science education and critical data literacy research to provide a framing to understand the critical data literacies athletes in our study were developing. Data science and critical data literacies are central concerns for science education because current societal science and engineering problems rely extensively on data collection, analysis, and modeling and the ubiquity of data science tools in society has led to an increased reliance on data practices for decision-making in society (O. Lee & Campbell, 2020). We situate our framing of critical data literacies in the context of sports sciences research and legal perspectives of sports science that reveal the broader context of data use in elite competitive athletics.

2.1 | Data literacy and data science defined

We define *data literacy* as the ability to collect, manage, analyze, and visualize data sets to understand its underlying phenomena (Maybee & Zilinski, 2015; Prado & Marzal, 2013).

While this definition emphasizes the technical processes involved in data science, researchers have called for data science educators to foreground the situated nature of data, grounding data science experiences in contexts relevant to learners and placing experiences within settings in which learners can take action and have impact (V. R. Lee & Dubovi, 2020; Philip et al., 2013; Van Wart et al., 2020; Wilkerson & Polman, 2020).

Thus, data literacy involves multiple components (Table 1). First, learners need *content proficiency and discursive fluency* to engage in the collection, management, analysis and visualization of data, including flexible skills for using novel tools and approaches in the process (Maybee & Zilinski, 2015; Philip et al., 2013; Prado & Marzal, 2013; Provost & Fawcett, 2013; Shields, 2005). Our definition of data literacy includes the development of practices and skills foundational for *data science*. Data science emphasizes large data sets often containable only in specialized database systems, and data-driven decision-making—practically applying data insights to real-world decisions (e.g., Baca & Kornfeind, 2006). Hence, data science researchers emphasize the importance of carefully considering the context in which data are situated, looking deeper into the metrics to see when something is beneficial, what problems might arise, what is being counted, and what is not. Data science also includes having an eye for when and how data analytics, data mining, and big data sets would be helpful for new problems and domains.

TABLE 1 Critical data literacies synthesis

Mechanical/technical data literacies	**Socio-technical data literacies**
Content proficiency and discursive fluency	Contextual, cultural, and social understandings
Data literacy	*Socio-technical data understanding*
Data collection, management, analysis, and visualization	*Questioning where data comes from and who works with it*
Data safety and management	Data identification, understanding, reflexivity
Data hacking	Recognizing the use of data to tell stories
	Questioning what stories are being told and who is telling them
	Critiquing the validity of those stories
Data science	*Socio-technical data productions*
Principled extraction of information and knowledge from data	Choices and actions regarding when to engage and when to disengage with data (e.g., managing settings)
Big data fluency—making decisions and developing insights from data	
Reading, analyzing, and comprehending large data sets	Producing one's own novel or clarifying stories with data—leveraging data in new ways or forms
Individual capacity to work with big open data sets	
Statistics, systematic study of the organization properties and analysis of data	

2.2 | From data literacy to critical data literacies

Our focus on critical data literacies is motivated by trends of data use in society that necessitate efforts toward a more just, equitable society. Specifically, our approach integrates technical data literacy with *contextual*, *cultural*, *and social understandings of data* and data use. Learners need contexts of study relevant to their lives and interests (V. R. Lee & Dubovi, 2020; Philip et al., 2013; Van Wart et al., 2020). Additionally, they need to understand the politics of knowledge. Data are never neutral and is always shaped by the context of its collection, analysis, and circulation (Johnson et al., 2021; Philip et al., 2013). Learners must also reflect on the limitations of normative language and traditional disciplines, leveraging opportunities to analyze data through new disciplines and perspectives, including their own everyday sociocultural ways of knowing (Philip et al., 2013).

Advances in data generating tools, devices, and storage mechanisms have led to the proliferation of big data analytics across all sectors of life (Galbraith, 2014; Watson, 2014). Additional risks are therefore posed to the subjects from whom these big data insights are extracted (D'Ignazio & Bhargava, 2015; Taylor, 2017; Tygel & Kirsch, 2016). D'Ignazio & Bhargava, 2015 cite several of these risks: data collected without approval or awareness of subjects, data collection by third parties not meant for observation or consumption, denying agency to the subjects of data collection, use of data to make impactful decisions on subjects' lives, and denying their participation in decisions that affect them.

Additionally, technical aspects of data science are so complex that they privilege and welcome certain actors (i.e., those from dominant groups by whom data science language and culture were developed and oriented toward) (D'Ignazio and Bhargava (2015); Philip et al., 2013). Thus, such complexities of data science disadvantage and push away others from minoritized backgrounds, with different linguistic and cultural norms, who often have less access to data science opportunities (Philip et al., 2013). Yet, since data science is situated in all aspects of everyday life with profound impacts on minoritized, resource-constrained communities (D'Ignazio & Bhargava, 2015; Taylor, 2017), it is imperative that learners in these communities understand and are aware of data practices and how they are impacting their lives. Furthermore, these communities must then have voices and agency in how these practices are enacted and how they are impacted by them (Johnson et al., 2021; Taylor, 2017). Indeed, social justice in the context of data science should allow *all* citizens to have a voice. Additionally, special attention is needed to break down systemic barriers that hold minoritized and resource-constrained communities from having equitable voices in how data are used to make decisions.

Researchers have lauded pre-college data literacy efforts for emphasizing socially motivated data experiences that expose learners to a range of technical, social, and political data considerations (Wilkerson & Polman, 2020). However, approaches at collegiate levels are often situated within disciplinary courses and programs (e.g., biology, computer science), where data literacy is framed in the narrow terms of particular industry tools and data science techniques (Wilkerson & Polman, 2020). We expand data science education at the collegiate level by considering collegiate athletes' informal learning in their everyday sports experiences as an avenue for socially motivated, deeply situated data science learning.

2.2.1 | Critical data literacy practices

In light of the sociohistorical and political nature of data science practices *and* the complexity of the technical aspects of data science, what has been called for is empowering "subjects" of

data collection through critical data literacy *practices* (D'Ignazio & Bhargava, 2015; Špiranec et al., 2019; Tygel & Kirsch, 2016). Critical data literacy practices situate data science skills deeply within sociohistorical and political aspects of data in society and in peoples' lives—recognizing the use of data to tell stories, questioning what stories are being told (and not), who is telling them (and who is not) and critiquing their validity (Calabrese Barton et al., 2021; D'Ignazio & Bhargava, 2015; Johnson et al., 2021; Tygel & Kirsch, 2016). Critical data literacy practices empower learners to find novel ways to tell their own stories with data—especially those in marginalized communities (D'Ignazio & Bhargava, 2015; Tygel & Kirsch, 2016). Data justice perspectives similarly call attention to the ways people are made (in)visible, (mis)represented, and (mis)treated in the production of data (Taylor, 2017). Critical data literacy practices then help learners to recognize and address disempowering framings and orient toward empowering ones through producing their own novel or clarifying data stories (Calabrese Barton et al., 2021; D'Ignazio & Bhargava, 2015; Johnson et al., 2021; Tygel & Kirsch, 2016).

We therefore draw on the following definition of critical data literacies: practices with hoped for liberatory effects, oriented around data, including its production, consumption, and sharing (Calabrese Barton et al., 2021; Milan, 2019). Critical data literacies hence empower learners through helping them connect sociohistorical political aspects of data science to technical data science practices for their own empowerment and that of their communities.

We leverage research on critical data literacy practices as a means to promote empowerment for learners to act upon data for their own freedom, well-being, and opportunities (Calabrese Barton et al., 2021; D'Ignazio & Bhargava, 2015; Taylor, 2017; Tygel & Kirsch, 2016). This means data practices must be situated in their broader sociohistorical political contexts (Calabrese Barton et al., 2021). Critical data literacy in action within novel contexts may look completely different from the big data of "experts" that have historically dominated industry and perspective, yet they are both valid forms of engagement.

More work is needed in novel contexts, particularly where minoritized learners are *already* deeply engaged with and directly impacted by data. Our CDS perspective considers ways we can leverage *local* community data practices as a context for engaging learners with sociohistorical political aspects of data in ways that might help them more broadly in society.

2.3 | Intercollegiate athletics as a context for critical data literacies

In this analysis, we extend the current work on CDS to sports as a means of reaching and engaging learners, athletes in particular, who have traditionally disengaged from science learning. While athletes at every level (from youth to professional sports) are deeply engaged in their sport, they often do not see the connections between their sport expertise/talent and their academic ability. Nor do they see the relevance of science-based processes and topics for their sport performance (Lubker & Etzel, 2007). In presenting opportunities for critical data literacies in athletics, we distinguish between technical and sociotechnical components because currently, they are each motivated and guided by distinct goals and disciplines. Mechanical data practices are largely driven by sport competition—that is, winning, optimizing training, nutrition, and revenue. Meanwhile sociotechnical concerns with data are being studied and put forward by researchers concerned with athletes' rights, policies, and education (e.g. Weight et al., 2020). In our CDS approach, our goal is to understand athletes' experiences navigating practices that integrate technical and sociotechnical components of critical data literacies.

2.3.1 | Mechanical data literacies in NCAA athletics

Advances in miniature technology, sensor modalities, apps, and wearable analytics tools over the last decade have prompted exponential growth in sport-related data (Halson et al., 2016; Malone et al., 2017; Piwek et al., 2016). Integrated devices can chart all aspects of physiological, mechanical, and psychological loads in real time. Heterogeneous aspects of team analytics are being utilized to understand group movement, insights into improving performance, effects of strategies and tactics, and event-related data. With increasing access to sensor-based technologies, film, and large-scale statistical data sets generated from sports training and performance, most major sports teams use extensive analytics to guide and inform training, game play, personnel decisions, fan engagement, and business decisions (Patel et al., 2020). Elite collegiate sport contexts thus provide a rich opportunity to understand athletes' access, engagement, and reaction to the technical aspects of data science.

2.3.2 | Sociotechnical data literacies in NCAA Division 1 athletics

Collegiate athletes live and work on college campuses and their lives are actively managed by the athletics institution. Complex sociohistorical political dynamics present the *need* for athlete empowerment. Research in the legal field has brought to light concerning risks inherent in NCAA Division 1 sports regarding big data trends (Greenbaum, 2018; Jessop & Baker III, 2019). The risks mirror those put forward by data science education researchers (D'Ignazio & Bhargava, 2015; Pangrazio & Selwyn, 2019; Taylor, 2017) further illuminating the opportunity and need for situating critical data science learning in this context.

This research highlights the growing trend of multimillion dollar deals between Division 1 university athletics programs and large corporations that grant big data companies like Nike and Under Armor exclusive access to athletes' performance, training, physiological, and even lifestyle (e.g., sleep habits) data. While these deals have significant benefits for corporations and universities, college athletes have little say in the negotiations. Designated as students, not employees, athletes at these elite sports institutions are not allowed to unionize or develop player associations to advocate for their rights (Jessop & Baker III, 2019). Jessop and Baker III (2019) hence put forward the need for policies and practices within collegiate athletics for informed consent, policies, and education (for athletes and staff) around big data use in NCAA sports. They also point to the need to consider the complex power dynamics on sports teams for helping athletes and staff to formally and informally advocate for the well-being of athletes. We build on these calls to contribute an analysis from athletes' perspectives on their own orientations to data practices at their universities and ways they take action—or not.

Like D'Ignazio and Bhargava (2015), Tygel and Kirsch (2016), Calabrese Barton et al. (2021) and Johnson et al. (2021), we take a learning oriented approach to critical data literacy, focusing on ways athletes empower themselves and their communities through their own understandings of data. We draw on Nasir and Hand's (2008) work which contrasted the contexts of high school basketball and mathematics classrooms to understand how these environments were promoting and limiting black males' identities. Their findings indicate the ways sports teams provided access to the domain of sports, opportunities to take on integral roles and for self-expression in ways unavailable in math classrooms. Their work inspires our aim to see ways athletes' experiences within intercollegiate sports gives them access to data science practices

and enables (and limits) their opportunities to take on integral roles in data science (Nasir & Hand, 2008).

3 | METHODS

This analysis is a part of a larger interdisciplinary research project focused on understanding and enhancing athletes' experiences with critical data science through sport. We also aim to: understand organizational practices with data in athletics, promote athletes' health and well-being with data, and develop data analysis and visualization platforms for athletes. With a focus on CDS within collegiate athletics, our methods started with understanding the data perspectives and experiences of athletes and staff from diverse sports and positions within the community before narrowing our analytic focus to focal athlete participants. We take a case-based approach to this analysis so that we can understand: (1) athletes' data practices within the socio-historical political context of athletics and (2) athletes' broader orientations to data practices in their sport. Both of these goals require analysis of individual athletes' experiences across excerpts of their interviews to observe patterns or themes evident across multiple experiences.

3.1 | Context/setting

We conducted 48 semi-structured interviews with athletes (22) and athletic staff (26) from a variety of sports within two Division 1, elite athletic universities to understand athletes' experiences with data. In particular, we were interested in exploring how athletes collect, analyze, and apply data and how these practices fit into social, athletic, academic, and health contexts. To recruit participants, we used chain-referral (snowball) sampling, a technique widely used in qualitative research (Valerio et al., 2016). Athletes were recruited from a variety of sports (e.g., wrestling, American football, and cheerleading), while staff were recruited from various athletic offices and departments related to a specific sport or sports (e.g., strength and conditioning coach: football). We adopted this approach in order to compare data practices within different sport contexts.

Interviews were conducted on the university campuses between May 2019 and December of 2020 (interviews conducted in 2020 were completed virtually due to the COVID-19 pandemic). All interviews ranged from 25 to 60 min and were conducted by members of the research team and transcribed afterward. Questions were open-ended and centered on two key topics: ***data for sports play*** ("What kinds of things do you count or measure for your sports play?"), and ***organizational practices*** ("Who's in control of your data?").

3.2 | Data analysis

All 48 athlete and staff interviews were coded for the larger team project with codes focused on (1) *What*: Data Practices, (2) *Where*: Organizational Context, and (3) *Why*: Identity/Discourses athletes were indicating orientations toward in their excerpts. Two authors did an initial round of coding of 11 athlete interviews and 1 staff interview to iteratively develop subcodes in each category. Two additional authors did a second coding pass of these codes, discussing any coding disagreements. To take our case-based approach, we selected 12 focal participant interviews to

get a range of positions within the athletics organizations. To present in-depth cases of athletes' critical data practices, we focused on athlete interviews, with the exception of one coach who emphasized a particular set of data practices with athletes on his team. We then leveraged remaining staff interviews when needed to further understand the context of focal athletes' interview reflections. We selected athletes from a range of sports (10 sports) and where there were >1 sport, we selected athletes who indicated a range of orientations toward data on their sports teams. To organize our data highlighting the most relevant parts of their interviews to this analysis, we structurally coded our data (Saldaña, 2015), starting with the most relevant codes from the original data set/analysis and we also coded interviews for athletes' personal orientations toward data.

For each case, we used the excerpts generated from our structural coding and mapped that participants' themes and excerpts according to: (1) data practices (social and technical), (2) orientations toward data, and (3) mapping to well-being, as this showed how athletes took up agency (or not) within the athletics organizational context for issues that mattered to them (i.e., sociopolitical historical aspects of their (dis)engagement with data). In an axial round of analysis, these mappings generated broader themes of athletes' orientations toward data practices (i.e., deference, fighting back, reflection, agency) that highlighted patterns and distinctions between different athletes' data practices on their teams and the sociopolitical historical influences on those practices. The first two authors reviewed, discussed, and iterated upon these themes.

There were two dimensions within each orientation type that characterized athletes' orientations toward data practices. First, there was a continuum with respect to the social focus of athletes' orientation—from focusing on their own individual data to focusing on the data and data practices of their team, broader sport or Division 1 athletics more generally. Second, there were ranging foci with respect to the type or topic of data athletes were oriented toward. Some focused more narrowly on data for their sport performance and others' focus was broader, also including data and practices to advance their whole selves beyond their sport (e.g., long-term health, academics).

Three focal participants were selected to present a range of social foci and type of data focus and to highlight the experiences of athletes within different parts of the athletics institution (e.g., high-revenue and low-revenue teams). In selecting these three cases, we focused on athletes at the two furthest extremes with respect to their reactions to data practices on their teams—those who were able to enact practices on their teams (fighting back) and one who was primarily deferential to athletics in his perspectives and use of data on his team. We did not include cases from the two categories in between this continuum (agency and reflection) as these categories drew out a different set of themes and implications that were not as tightly connected to the research questions posed in this manuscript.

Based on our approach to CDS, we leverage a critical data literacy framework to analyze participants' descriptions of everyday athletics practices. Our interview protocol was designed to help athletes identify and talk about data practices in their own lay and sports specific terms (e.g., we asked athletes to walk us through a typical day in their sports practice and play, and recount everything that was counted or measured about them). We used a critical data literacy lens to identify relevant critical data literacy practices even if athletes did not label these practices as such. Additionally, our codes focusing on "where" and "mapping to well being" helped to reveal key aspects of the sociohistorical political contexts that drove and hindered athletes' data practices. Axial coding was especially attuned to generating overarching themes related to the influence of the sociohistorical context in athletes' data practices (e.g., data practices were grouped by where they took place and who managed them).

3.3 | Positionality

Our diverse and interdisciplinary research team brings a complex positionality to the study. Our team includes researchers from Education, Human-Computer Interaction, Sociology, and Exercise and Sport Science. These fields drive our approach to this project and the analytic lens we use to understand critical data literacy practices and CDS. Our interpretivist approach is largely guided by the fourth authors' background in Sociology, orienting us toward a more critical and organizational perspective. Our approach to understanding and representing learners' critical data literacy and CDS is guided by the first two authors' backgrounds in the Learning Sciences and HCI. Specifically, we orient toward a more emic approach to this research, valuing and prioritizing the inclusion and welcoming of participants' unique voices and ways of engaging, especially those underrepresented in STEM (i.e., those from minoritized communities, women, athletes). Yet, we leverage frameworks from the Learning Sciences that also bring an external lens to illustrating the complex data science practices that these unique voices and engagements represent.

Some members of our team bring an insider perspective with respect to their lived experiences—that is, the third author is a former Division 1 college athlete who works in and studies the college sport industry. Other members bring lived familiarity (Holmes, 2020) to the minoritized athlete experience as the first two authors are black. Additionally, the first author is an avid runner and fitness enthusiast, especially valuing and recognizing the STEM practices that go into achieving their health and fitness goals. The first and third authors are on our universities' athletics councils, though the third author has more regular interactions with athletics via their research, consulting, informal mentorship, and physical proximity to athletics on campus. These diverse vantage points bring a range of perspectives to this study that both help us understand the complexity of our community context and that may also mask or bias our views of interactions within collegiate athletics and their relation to CDS.

4 | FINDINGS

We present cases in terms of sociohistorical political motivations for athletes' critical data literacy and ways those motivations promoted and limited their critical data practices.

4.1 | Byron: Fighting back to enhance my performance

A sophomore shot put thrower on the track and field team at State U, our analysis suggests Byron used critical data practices to fight back against athletics in support of his physical training and performance in his sport. He shared reflections in his interview of ways in which he had to fight for his well-being with respect to his sport performance. While Byron had power to use data to enhance his sport performance in some aspects, there were also ways in which he did not have power to fight back against athletics authorities—even if backed by data.

4.1.1 | Sociohistorical political motivations mediating data practices

Byron's data practices were guided and motivated by his broader socio-technical concerns. He reflected on how athletics was a business driven by winning and his coaches' motivation was

focused on winning, even at the expense of his long-term health. In such a system, Byron recognized his and his athlete friends' roles as investments—and for Byron, an international athlete, a pretty costly investment—athletics (and his coach) had made and the large dividends expected in return. He observed how this business focus could sometimes go against an athlete's health,

> *Because at the end of it, this is all just a business. Even if we think of throwing, or running, or basketball as something that we love... For the coaches and athletic directors, it's all just a business. They just want their return on investment. They're not spending 100% scholarship on you so that you don't go to nationals because of an injury. They want their return on investment.*

Byron critically reflected on the role and power of his event coach in making decisions regarding his training, competition, and ultimately his economic mobility. He reflected on ways coaches misuse their power to make athletes' lives especially hard, acknowledging their power over their athletes in stating, "*if these coaches have a bad day, they can turn your world upside down...*" Because of these dynamics, "*if the coach and athlete are fighting, the athlete has to lose 90% of the time, even if the athlete is right.*" With these reflections, Byron carefully navigated which data practices he needed to submit to and which he could take his own agency with—sometimes quietly resisting and other times explicitly fighting back.

4.1.2 | Developing his own suite of data practices

One way Byron navigated his motivation for athletic success, reflections on the "business" of athletics, and power dynamics on his team was through taking agency to develop his own suite of data practices. Informal support from his coaches back home and athletics support staff not directly evaluated on wins and losses (e.g., the nutritionist) influenced and guided these practices. Byron extensively *reviewed film* on his own, recording, analyzing, and perfecting performance based on his analysis. He determined what he would capture, set up and stored data on his own equipment, cleaned and reviewed the data, and ultimately adjusted his technique. Additionally, Byron and his nutritionist *managed his nutrition* together. He kept track of how he was feeling relative to what he had been eating and sought the nutritionist when his felt experience indicated he needed to switch up his diet. Together they set nutrition goals (e.g., to lose weight), tracked progress, and made adjustments to his regimen based on concerns for his well-being (e.g., loss of muscle mass). Byron also kept track of his own "*felt data*" regarding what his body implicitly told him about his performance. For example, he tracked load management—or how often and hard he had worked certain muscles so that he could determine what to work more or less and when to take breaks.

Byron's coaches asked the team to track data about a range of metrics (e.g., how they felt, what they had been eating, etc.) in a book that he carried to log all of this information. Byron gave a detailed description of how his coaches used this data to determine their strength and training regimens (e.g., when to cut weights or reps). However, our analysis suggests, this coach-driven data mechanism was one that Byron actively chose *not* to engage deeply with. While this data aligned with some aspects of Byron's own individual practices (e.g., tracking his nutrition and how he felt), Byron did not talk about how *he* used this data and in fact, he mentioned that he used an even more stringent plan for his nutrition than what the coaches set.

4.1.3 | Fighting back for sport performance

Our analysis suggests that Byron's recognition of his elite status as an international athlete at "*the best throwing camp in the hemisphere*" also gave him the confidence to resist athletics-driven data practices. Byron described a prior year in which his team had a new strength coach. He was frustrated that the new coach came in and tried to correct Byron's strength training practices he had developed over the years. He was especially frustrated that this relatively new coach tried to overlook Byron's own expertise and background stating, "*I'm coming from the best throwing camp in the hemisphere. And then to come here and to have this guy say that I'm benching wrong and I'm squatting wrong pissed me off*."

Byron leveraged his *felt data* to argue with the strength coach about which of his muscles needed targeting in training. For example, when he realized that his team had not worked their hamstrings in a year during their strength training, Byron reported that he "*felt my hamstrings firing*" letting him know that he needed to do some strength training targeting his hamstrings. Keeping track of this involved leveraging his knowledge of muscles and how they work together (e.g., recognizing that he needed to focus on protagonist and antagonistic muscles to prevent injury). Fighting back, Byron asked the coach to give him a different workout. While the coach obliged, Byron still was not satisfied with the workout the coach gave him, recognizing that he could *feel* that the workouts he was given were still not as effective at targeting his hamstrings as other workouts Byron knew of. He felt his training could have been much more effective with different exercises. To continue fighting back against his dissatisfaction with the new strength coach, Byron reported his dissatisfaction on surveys from the athletics department. With similar results from other athletes, Byron reported that the strength coach was subsequently fired.

Our data suggest that Byron's ability to fight back in these situations was supported by the confidence he developed from his elite training back home. Byron had been working with a coach he absolutely trusted and saw as one of the *best*. As an international athlete so he was pretty confident about his skill level and ability to decipher discrepancies in his training program. This all seemed to play a role in his having the confidence to speak up to and against the coach. Byron himself described his belief that coaches, new ones especially, should listen to their athletes, especially international athletes:

> *I feel like if you're that young, it wouldn't hurt to take advice from your athletes, especially ... I would focus on international athletes, because a lot of money's being put into them, and ... because they actually study the sport and they have a vast amount of knowledge into doing things right.*

4.1.4 | NOT fighting back with the head coach over medical clearance/performance

Even with his agency and confidence, our data point to an area in which Byron did *not* fight back. During the prior year, he sustained an injury to his pectoral tendon and the surgeon did not clear him to throw until 6 months after his surgery. However, within 3 months his coach had him throwing, privately, away from the team (and physical trainers, who likely were aware of Byron's status and medical needs), "*We would throw there by ourselves... because with the general team, that's where the physical trainers are...*" While Byron fought back against his strength and conditioning coach, he simply submitted to his head coaches' orders to throw before he was medically cleared. In this case,

Byron's reflections on the dominant role and power of his coach and the drive in athletics to win may have overshadowed his sense of agency to fight back. Byron's decision not to fight back may have also been mediated by broader socioeconomic concerns. He was motivated to train because of his desire for upward economic mobility for his family. Sports offered a way to pay for college and the potential of a career afterward:

> *I come from a poor family back in [my home country]. I'm first in my family to come to university and actually even live on campus ... My mom cries when she goes home, and I've seen that. And that's my motivation. Sometimes when I'm here and I'm like, "I don't want to do this," my only motivation is just that I can get my mom in a better position. That's literally my drive sometimes, because sometimes I lose heart... But just my mom ... I cannot let my mom live like this.*

He reflected that many athletes did not think they had options beyond sports because they did not see themselves as academically inclined and hence it was all the more important that they remained viable (i.e., healthy and elite) in their sport, "*There are a lot of athletes who are not academically inclined, but they are really athletically inclined. And they really want to go pro, and then they're always worried about being burned out.*" These reflections may have limited his ability or willingness to risk fighting back.

4.1.5 | Case analysis

Our analysis suggests Byron had a deep understanding of who was working with his data and that he often reflected on what decisions were being made with his data, questioning the validity of those decisions. His fighting back was empowered by his confidence as an accomplished international athlete and close informal relationships with athletics staff (i.e., the team nutritionist) and his external elite coach from high school. Byron's case shows that his critical data practices mediate both the decision to resist and the decision to submit to practices not necessarily beneficial to his well-being.

4.2 | Lei: Fighting back for life balance

Lei was a senior middle distance and cross-country runner on State U's track team whose major was Public Health Science (in the Pre-physical Therapy track). Lei fought back not only for her sport performance but for general balance in life as she navigated the intense demands of being an athlete at State U along with her other personal and professional goals. With such broad goals, unlike Byron, Lei's fighting back sometimes (and often) included practices *not* geared toward enhancing her sport performance. In fact, her enacted practices and those she refrained from, could be seen as detracting from her sport performance in support of her health more broadly. Her case centers around the theme of trying to find life balance among competing goals and cultures.

4.2.1 | Motivated by the healthy aspects of running and Division 1 sports

Lei's critical data literacy practices were motivated by broader socio-technical considerations as she reflected on the healthy and unhealthy aspects of running and sports. She acknowledged

that the stiff competition in Division I sports made her a better person, "*Running is probably the best environment I've ever put myself in because it's taught me so much discipline... It's taught me how to take care of my body and it's taught me to be a student athlete*." Motivated by this disciplined and competitive culture, Lei engaged in several types of mechanical data practices focused on understanding and enhancing her sport performance and health.

Lei explained that runners used GPS watches to collect and calculate their running speeds, time, mileage, and heart rates. Additionally, athletes kept track, intrinsically of their own *felt* data, knowing through their senses the pace they were supposed to run. Lei tracked her exertion levels during runs and mapped that sensed data to reports of her actual speeds and heart rates during runs. If she could hit the same speeds with less "felt" exertion (as measured by heart rate and sensed exertion) she could see that she was improving:

> *Let's say I'm running four miles at a 7:30 pace. I'm feeling good, and my heart rate is 150... By the end of the season let's say I'm running that same four miles, but the 7:30 pace feels like I'm jogging and my heart rate is 130 average. That feels good because you feel like you're improving.*

Lei reflected that she and her teammates did not need to share their data because they all ran together and used their GPS watches, so they had a felt sense of how each person was running. Lei also measured her sleep regularly and worked with a nutritionist to get feedback on what she should be eating.

4.2.2 | Fighting back against the toxic culture of sports

While Lei appreciated the discipline and competitive culture in Division 1 athletics, she also recognized that the culture induced by such extreme competition could quickly become toxic. Athletes (and coaches) could become so fixated on competing that they engage in unhealthy practices or get to an unhealthy state. Some unhealthy aspects she discussed were burnout, injury, unhealthy mental/emotional outlooks on data and performance, and eating disorders. While the pressure she experienced on the team could be quite motivating, it could also cause a roller-coaster of unhealthy emotions:

> *There's burnout all the time. I know so many runners that have quit. I know so many runners that don't enjoy the sport anymore. Running's very up and down. You can feel like you're on cloud nine and you're killing the workouts, you're having great races. And the next week you literally tank and have the worst race of your life. And those are the times where you're like, I don't want to do this anymore.*

Data practices Lei avoided

In order to avoid the negative effects of the toxic culture in her sports program, Lei created her own data production, managing which data practices she would and not participate in, which she would and could negotiate with her coaches, and how she would participate in the ones she chose to. There were a series of data practices Lei specifically chose to avoid, limit, and stay away from.

Limiting data practices

Lei decided *not* to share her data with the team while she was actively recovering from injury. Her team had recently begun using an analysis tool called Training Peaks that would get data from the running watches most women on her team used and produce analytics of their mileage, speeds, heart rates, etc. Lei saw the benefit of using Training Peaks for maximizing her training and performance. She observed that these tools could help runners understand where they were improving and identify issues related to burnout and other medical conditions. However, while she was injured, Lei's team began using the analysis tool collaboratively so that each runner's data were shared with the coaches. Lei recognized that it may not have been in her best interest to share her data with the coaches while she was recovering, in fear that she would "*become OCD about*" her performance and risk reinjury:

> *So the reason why I never wanted to get on [Training Peaks] was they started doing it while I was hurt. And basically your coaches can see everything. So when I was coming back from injury, I wanted to come back on my own pace because I knew that if my coaches could see everything I was doing, I would probably take my runs way too fast.*

Lei knew that she could not do as much mileage as the other women so while she was making modifications to her training (e.g., supplementing runs with time on the elliptical), she did not want to be tempted to risk her long-term health to advance her immediate performance. She observed other women on her team also actively made the same decision not to share their data as they were recovering from injury for the same reason.

Additionally, seeing the risks of eating disorders personally as some of her teammates became too fixated on their weight made Lei determined to steer clear of any practices that could lead to eating disorders, especially counting calories. Lei pointed out that the line between eating healthfully to enhance performance and developing eating disorders could be quite blurry. "*There are a lot of eating issues. I feel like eating issues, performance-based eating, and eating disorders have so many lines that cross. But then they're blurred.*" Lei navigated those lines in effort to stay balanced, choosing not to count calories but to maintain an awareness of the nutritional content of what she was eating:

> *So I tried to stay as far away as possible with counting calories, but I would be lying to you if I said I don't know the nutritional content of food. I'm very aware just because I've always been interested in nutrition.... So I'm very aware of what's in food. I know what's good and bad and I understand calories. But I try not to count them or anything just because I've seen it negatively affect other people.*

In addition to managing which practices she would avoid, Lei also thought carefully about which practices she could negotiate on her team and she fought to maintain healthy perspectives of her data. Lei described ways she and her teammates had collective speed ups and slowdowns when they did not agree with the speeds set by their coaches because they could *feel* that they were too fast or too slow. In such events, they would all either collectively run faster or go slower at the end. This was one way that Lei and her teammates fought back for healthy practices in data-driven and sensory-oriented ways.

Additionally, Lei talked about the importance of having healthy and realistic expectations of her body. This really hit home for her when she was injured and feeling the pressure to continue to compare her injury recovery workouts with those of her noninjured teammates:

> *Imagine the environment I just talked about. And then imagine getting pushed back 10 steps when you get hurt and not being able to do anything about it. It's like you're already in this cutthroat environment where you're already just trying to get better than the girl next to you, and then all of a sudden, it's like, well actually you're going to sit on your butt for six weeks because you have a stress fracture in your shin. And then you have to do all the cross training alone, and then you have to try to come back. And when you come back, your first run you're literally like, why do I run?*

Lei realized she could not compare her recovery training data to her noninjured teammates. Her reflections on healthy perspectives of her data extended to understanding her body more generally to set realistic expectations. She discussed how it was important (and often difficult) to distinguish between times when she was experiencing weight gain as an indication that she was slowing down versus when weight gain as merely an indication of a typically developing body:

> *It's really hard with girls because a lot of us come from high school where we were all itty bitty, our metabolisms were furnaces. And then you come into college and I'm 21 now. I look nothing like I came in when I was 17. And I shouldn't. A 21 year old shouldn't look like a... 17 year old. But it's hard when you're running to kind of get that knowledge grasped, that my body's changing because I'm now 21. Or versus oh no I'm gaining weight and I'm slowing down.*

4.2.3 | Actively managing her identity

Like Byron, Lei recognized that athletics was a business and the focus of the organization prioritized winning and revenue over her long-term health. Hence, she noticed that interventions were much more likely when an athletes' health was hindering their performance and *less* likely when athletes' unhealthy practices or conditions (e.g., eating disorders) were resulting in enhanced performance:

> *If you come in from the summer and you've dropped 30 pounds and you look... in a normal human's eyes [like] a little skeleton, but you're running amazing, it's usually not addressed. You're performing, it's a business. The coaches are like, okay well she's our new number one girl type of thing.*

A broader way that Lei fought back for her personal health was through actively managing her identity. When asked about metrics outside of her sport that she counted or measured, Lei brought up measuring her church attendance (spending 2 hours per week) and study time "*Because, in the end, I'm trying to graduate and have a good life.*" This balance of embracing her identity as an athlete while actively maintaining other aspects of herself was a primary strategy

Lei discussed for combatting the toxic culture of sports, "*It's been one of the most toxic environments I've ever been in. Because I've seen the eating disorders. I've seen what happens if you put your identity and your sport... you don't care about school... You know? So, it's all about balance.*" In fact, at the time of our interview, Lei had run her last race as a State U Orchid. She reflected on her love for State U athletics, while at the same time, she was ready to be "free" from the intense pressure.

> *So, it's exhausting. I'd be lying if I told you that I ran my last race and I was not devastated. I was very nostalgic, I was like, I'm going to miss this, I love my team, I love everything running's given me, I love [State U] athletics, I will always be [an Orchid] and I love everything about this school. But I'm very ready to leisurely jog and to enjoy running and not feel so stressed about getting on the track every Tuesday, Thursday. Or not have that race pressure and not care so much about what I eat, and not have such a rigid schedule. I am nostalgic that I'm leaving, but I'm... I feel like a weight had been lifted off my shoulders. I literally was like, okay I'm free, I'm good, I can do what I want to do now. Love the team, love the school, love the coaches, but I'm good.*

Lei's reflections imply that she is ready to be free from the intense focus on data—capturing her run data, maximizing performance, feeling all of the pressure inherent in such a competitive environment. She was looking forward to foregrounding other aspects of her identity.

4.2.4 | Case analysis

Lei had extensive access to socio-technical aspects of data practices on her team as she reflected deeply on what data she and others were collecting about her and what decisions were being made with that data, especially those that impacted her well-being. She then leveraged these reflections to create her own data productions in the form of her active management of which data practices she would participate in and how she would go about them. Our analysis suggests that Lei's intentionality about actively fighting to *not* tie her identity to her sport alone and to prioritize activities and identities beyond her sport gave her confidence to fight unhealthy practices. Implicit in Lei's case is the agency she had with data practices because she was on a low-revenue team with less resources, scrutiny, and institutionalized practices.

4.3 | Omar: Deference to athletics for sports performance

Omar was a junior Data Science major at State U and a linebacker on the American football team's defense. Linebackers play pivotal roles in reading offensive plays, calling their team's defensive strategies on the field and preventing the other team from scoring. Omar engaged in numerous data science practices on his team, mostly driven by his coaches and the data emphasis in the National Football League (the professional league Omar aspired to). His deference was largely driven by the organizational structure of athletics which offered the football team increased access to data practices. However, the organizational structure on the team constrained Omar's access to the technical data practices and significantly limited his access to the socio-technical data practices.

4.3.1 | Extensive football staff obscures Omar's access to data practices on the team

With increased coaches and staff, the mentality within the organization of the football team emphasized assigning sole responsibility of data practices to athletics staff. For example, while athletes on other teams described completing lift cards detailing their weight lifting metrics during strength training sessions, Omar explained that due to the size of their staff, football players do not need to complete lift cards, "*They just take you through it because we have four strength coaches, and two interns*." With such extensive staffing and resources on the football team, most data practices Omar described were driven and guided by coaches and managed and enacted by the athletics staff, especially the strength and conditioning coaches, who Omar said were in charge of his data, "*He's like a liaison to tech companies if we want to use their technologies to track stuff. He gets the data, then he manipulates it however he sees fit*."

Coaches and athletics staff also drove how the data that was collected was analyzed, how they were used to make decisions about training, and how they were communicated (or not) to football players. Omar described a chart the assistant coaches maintained and displayed that showed the athletes who had the highest weight training achievements (e.g., the heaviest bench press and squat weights, the highest inclines and top speeds run). He mentioned that they also have technology that measures the speed at which they lift each weight rep. Omar only had access, however, to those top weights, speeds, and inclines that were generated—he had to ask his coach to get access to his own individual numbers.

Additionally, Omar talked about how their speeds throughout practices and games were tracked through sensors placed in their shoulder pads. This data and the decisions made with it were communicated among coaches and athletics staff, but Omar only saw the resulting decisions regarding training and practice made from the data. He was aware of these practices because of conversations with his coaches regarding how these decisions were made, "*During the games too, my coach told me they have a weekly meeting on that so they use that to determine how guys are changing speed during practice, if they need to lay off or if guys are just not putting in their full efforts*." Additionally, athletics tracked extensive data regarding Omar's weight and body composition. Omar described the complex Dexa scan measurements that were taken for every player on the team as something that was done *to* him, "*They measure our body fat percentages, our muscle tissues and stuff. They put us in this van and they scan us, and just a full body report, body fat, muscle tissue...*"

4.3.2 | Engagement in mechanical data literacies

Even with such extensive staff in charge of their data, there were key ways that Omar was engaging in mechanical data literacy practices quite extensively. First, Omar was leveraging the mounds of data collected from him, his teammates, and his opponents to compare himself—his stats and physical data—to NFL scouted players. He kept a running count of his tackles and stats that he knew by heart. Omar also appreciated the more extensive data his coaches tracked for him, especially the sensor data generated from the Catapult sensors in his shoulder pads. This data allowed him to compare his speeds to those of athletes in the NFL to get a sense of where he stood, "*... the catapult's the one that measures our speed, because they do it in the NFL all the time so it's nice to see your fastest run compared to how fast they run. It gives you a glimpse of where you stand*."

At the same time, Omar would often get discouraged when his data did not measure up to NFL standards. He particularly recognized his height and weight were low compared to NFL-scouted linebackers, thus limiting his chances of making a professional team

> *... I used to measure my height and weight because for my position group, I play outside linebacker and you normally tend to be at least 6′2″ 6′3″, 240 plus. I would go everyday and look at past combine measurements, those NFL combine measurements because those are usually the most accurate. I would see just about every outside linebacker they have there is over 6′2″, 240 plus. I'm not 240 so that would discourage me and make me think like, "I don't have the right size and dimensions to really be successful.*"

Even practices Omar managed himself such as nutrition plans were driven by coaches and NFL standards. Omar's process of managing his food and nutrition was guided by the nutritionist as he worked extensively with the her to plan meals, "*Trying to eat at least every three hours, just have meals that are dense in carbs, drink a protein shake before I go to sleep so I'm not hungry throughout the night, then I wake up in the middle of the night to drink a protein shake, so I constantly fuel my body*." The nutritionist told Omar when to incorporate more foods into this regimen and which types to incorporate. He could also text her at any time to ask questions about his diet. When determining his nutrition plan, he had to choose whose data emphasis he would focus on, the nutritionists' or his coaches'. Though the football nutritionist advocated not focusing on body fat percentage, instead focusing on "*lean muscle mass and fat mass [to] see changes in those two categories*" it was the coach's focus on body fat percentage (which the nutritionist explained was driven by the NFL) that Omar echoed as he tracked his body fat percentages to note changes and did extensive meal planning and scheduling to bulk up. While his nutritionist advocated goals related to his overall health, Omar's more motivating factor may have been NFL and coach-driven emphases because his career trajectory was at stake.

Omar further deferred to his coaches' advice (and admonition) about food and nutrition as he described how his and his teammates' eating was tracked when they used their athletics identification cards to purchase food off campus (on days where they did not eat at the "training table" where the nutritionist selected the menu), "*They track what we order on those cards to make sure the guys who are overweight aren't eating what they're not supposed to eat*." Omar recounted a time where his strength coach brought another athlete's meal order receipt and told the team they needed to do better with their eating, "*... our strength coach ... came up to us and he brought out someone's order and it had blue cheese, ... it was not a healthy order and he was telling us we got to eat better than that*." When asked how he felt about that exchange, Omar said he thought it was funny. Although the coach was bringing an athletes' meal data to the attention of the team to admonish their eating habits, Omar's response went more in line with what his coach was trying to get across.

Perhaps, most extensively, however, the data-driven practice Omar engaged in and described in the most detail was film (digital video) analysis. His deep value for film data analysis was revealed as he exclaimed several times that "*Film is everything about football*." Omar thought film was so vital in football because it's "*where people watch what you do and scheme from it*." To Omar, this meant that film revealed how an opponent's offense operated, so that they could be prepared for a game. More closely, Omar described how coaches and players

watched film of their own previous games to *catch their mistakes and scheme for success* in an upcoming game. For example, Omar described how film of a previous game revealed ways their special teams unit unintentionally gave away which play they were about to run before the play began. Their coach then, from this analysis, showed them a way to conceal their strategy to make their opponents think they were about to do something else, which led to a State U touchdown at their next game against their opponents:

> *Oh, against [a division rival]. On kickoff, when they kick the ball off, they would read the people that were blocking them, they would read our hips. If I'm turning this way, it's obvious that the turn is coming this way so I don't want you to come that way so they would come across to get inside the return. Our coach told the kick-off return unit to fake turning this way, and let them turn that way, at the last second turn the other way and then we'll run it back, and then we scored [a] touchdown.*

Additionally, Omar described how watching film of their opponents helped form an understanding of *their opponents' tendencies*, "*It gives us tendencies, so helps us play faster. Having a better idea of what they're about to do*." Although film review helped Omar understand his team's mistakes and opponents' tendencies he explained that it was another type of data analysis, reading his key, that was even more important, "*...the most important thing they emphasize is, 'Read our key,'...*" Omar explained, "*My key is the shoulder of the linemen, if he's turning down, I know it's a run block and then someone else is coming at me, so I got to play differently. Or if he's coming at me, I know it's a run block and I got to set the edge, if he's kicking back, then it's a pass rush*." Although film review helped him understand his opponents' tendencies and his mistakes, Omar reflected on how he had learned that if these data points conflicted with his key, he had to prioritize what his key was telling him:

> *It's good to have tendencies but you can't rely on that, you always have to trust and rely on your key. Each position has a key, the shoulder pad is my key so I'm supposed to lock in and look at the shoulder pad, and react off of that immediately. What messes up my data is when I lose focus on my key, when I'm trying to guess, or when I'm not reacting like I'm supposed to with my key.*

Indeed, Omar's film review analysis and key reading were critical data points Omar had to analyze throughout the week, then recall and analyze during games in split seconds at the line of scrimmage in front of thousands of fans.

Omar's description of film review on the football team suggests that even though these processes were so near and dear to him, coaches and athletics staff still drive and guide them. When asked how he schemed, Omar explained, "*I don't really scheme, it's the coaches' scheme*." However, because of this division of labor—coaches scheme but athletes make game time decisions on the field, film review practices were much more transparent to athletes and there was a scaffolded process for helping athletes understand and engage with film data. Omar explained that prior to communicating around film data, athletics staff cleaned data by breaking it up play by play and then organized clips by team formation in the play and by which personnel (i.e., players) were involved in the play. While coaches likely analyzed this data extensively beforehand, they then watched relevant film clips with the players, taught the players how they were watching it, and then helped the team see how to interpret and apply the data.

4.3.3 | Limited access to socio-technical data literacy considerations

Omar's mechanical data practices illustrate the complex data science Omar was tasked with doing with his coaches and then on his own during games. In fact, Omar's description of the extensive levels of film review, analysis, and application he did on the football team mirror the types of big data analytics decisions often relegated to large computer systems (Provost & Fawcett, 2013). Yet, Omar was asked to recall, analyze, and make decisions from extensive qualitative film review in split seconds at game time. However, our analysis suggests, Omar may not be immediately recognizing his extensive engagement in data literacy practices inherent in his football engagement. While he had not mentioned film previously, at the close of the interview, Omar asked, "*Does film count as data?*" After his question was affirmed, Omar went on to describe the extensive use of film data on the football team to enhance their game play.

In his interview, Omar did not explicitly identify any socio-technical aspects of data practices in football as guiding or at least connected to his data practices. His analysis suggests he deferred to athletics to make data-driven decisions (and to collect, manage, clean, store, and analyze his data) to advance his well-being. However, one socio-technical consideration Omar was beginning to observe was long-term physical effects on him and his teammates, "*I feel like a lot of people are just [starting] to feel the repercussions of football. A lot of people's like, especially linebackers, our fingers are really messed up, I wear this because my finger was like this one day. It was really bent.*" Omar was also noticing more generally across the team, "*I've seen a lot of people's fingers just not the same anymore, and people are feeling sore.*" While Omar had begun wearing a brace to straighten his fingers, it was his coaches he relied upon to mitigate these effects by managing their training loads, "*The head coach he would meet with the strength staff to see where a player changes speed, to see if they need to lay off. Lately, we been not doing as many reps. We still practice hard but not as long, not as many reps.*" Yet, he and his teammates were still feeling the effects of their training and game play on their bodies.

4.3.4 | Case analysis

Our analysis suggests that Omar was deeply engaged in mechanical aspects of critical data literacy. The extensive resources on the football team gave him access to more big data tools and practices than any other participant in this study (not on the football team). However, our analysis also suggests his access to mechanical and especially socio-technical aspects of critical data literacies were limited by the division of labor among the extensive athletics staff on the team. Furthermore, Omar's deference to his coaches and the coaches' extensive reach into athletes' data suggest that football coaches may have had more authority with athletics data even than coaches had on other teams. Similar to Johnson et al. (2021), we surmise the social context may have played a role in the absence of his discussion of power dynamics on his team or other socio-technical aspects of data use on his team. Even with such power dynamics, Omar's need to make complex game time decisions on the field with limited communication from his coaches afforded him extensive access to film data analysis that mapped to data science techniques often carried out by extensive socio-technical systems.

5 | DISCUSSION

Returning to our research questions, our in-depth analysis of three focal participants reveals insights about critical data literacies athletes develop through their sports participation and their implications for expanding the field's understanding of CDS. Analysis of their differing reactions to data practices on their teams and in athletics (i.e., deferring to athletics and fighting back) and ways athletes oriented toward their own and their community's well-being reveals the critical role of athletics organizations in shaping athletes' access, engagement, and agency with data.

5.1 | Recognizing and building on athletes' critical data literacies

Returning to research question 1 (*What critical data literacy practices are athletes engaging in through their participation in athletics?*) our analysis shows ways mechanical data practices were deeply intertwined with socio-technical aspects, yet this intertwining looked different for different athletes and sports.

5.2 | Orientations to data practices

We found athletes had different orientations to data practices on their teams (e.g., fighting back for sport performance, fighting back for life balance, deferring to set standards) that were largely influenced by the sociohistorical political context of athletics. The orientations in turn shaped their access to and engagement with both the mechanical aspects and socio-technical data practices on their teams. Table 2 shows ways athletes' sports engagement afforded access to mechanical data science practices. However, athletes' orientations to data practices influenced how they navigated taking agency (or not) with socio-technical data understandings and productions (also shown in Table 2).

Athletes in our study who fought back against data practices discussed the underlying motivation and driver of athletics (i.e., making money, winning) and ways those motivations were often contradictory to their well-being. From that understanding, both Byron and Lei found ways to resist data practices and to create their own data productions. In order to access the sociohistorical political aspects of data practices on their teams, Byron and Lei had to individually and proactively navigate and manage certain data practices (e.g., creating their own film data, tracking their heart rates, speeds, practice runs, and throwing distances). On the other hand, taking a deferential orientation to data practices on his team, Omar's engagement with data were largely driven and guided by coaches who determined what data Omar should review, how he should review it, and how he should make decisions based on these data. But, this deference significantly limited Omar's engagement with sociohistorical political aspects of team data practices. With much less individual responsibility for data practices on his team, Omar had less exposure to socio-technical data understandings. He did not have to make any decisions about which data to collect, tools to use, decisions regarding training and strategy to make based on those data and was consequentially more deferential to those who managed these data practices. This is particularly ironic given Omar was majoring in data science at State U. Alternatively, Byron's close examination of his training, conditioning and practices led him to identify a gap in his training and address it with coaches. Hence, our findings suggest the importance of giving learners access to and responsibility with their data in community contexts, to promote engagement with both technical and socio-technical aspects of data practices.

TABLE 2 Data literacy practices in NCAA Division I athletics

Mechanical/technical data literacies	Socio-technical data literacies

Data literacy
Felt Data
• Sensory counts, measurements and estimations
• Tracking subjective feelings (e.g., exertion, fatigue, overall sense of well-being)
• Mapping subjective metrics to sports evaluations
  ○ Mapping to sports metrics/goals for evaluation of their athleticism
  ○ Mapping to physiological knowledge to evaluate health, nutrition, and training
• Coordinating felt data with sensor-based and statistical data
Tracking
• Collection and analysis of data related to sleep, nutrition, and performance
Qualitative Analysis
• Film data
• Observation data
	Socio-technical data understanding
	Mapping Social and Organizational Understandings to data practices
	• Identifying who collects, analyzes, and tells stories about their data and what decisions are being made as a result
	• Reflecting on impact of data practices on immediate and long-term well-being
	Understanding the role of the organization of athletics
	• Benefits and challenges of elite sports contexts
	• Recognizing the business side of athletics
	• Identifying power dynamics on their teams, in athletics and across sports
	Limits to and drivers of athlete agency
	• Recognizing when they do and do not have access to their data
	• Understanding what decisions athletes can and cannot make
	• Understanding which decisions of coaches and staff can and cannot be disputed and navigating power dynamics to dispute
	• Understanding the role of organizational factors (e.g., the type of team, status on the team, resources they have access to) in promoting their agency with data

Data science *External data* Tracking individual progress relative to big data sets across teams to evaluate their progress with respect to league and data standards *Narrowing down which data to focus on* • Data validity: figuring out which sources are the most reliable and accurate • Extracting actionable relevant information and knowledge from large sports data sets *Analysis and application* • Identifying trends in performance and training • Mapping trends to sports, fitness, causes • Developing actionable insights from analysis	**Socio-technical data productions** Mapping the language of sports to the language of data literacy/data science Developing their own suite of data practices they manage apart from coaches Deciding which data practices they will and will not participate in and which they have the leeway to decide Developing new forms of data collection, management and analysis to tell their own new stories

5.3 | Felt expertise

In each case, we observed a "felt" sense of data each of our athletes brought to bear as an expertise they exhibited, developed through their extensive sports play and the ways they physically experienced data. Athletes used their sensory experiences in sports to count, measure, and estimate metrics from their sports play. Their tracked felt sensations (e.g., fatigue, exertion, hits) became data that they mechanically, mentally and physically tracked. They then connected

these felt experiences with their extensive knowledge of their sport and physiology to evaluate themselves and their peers with respect to their performance, training, health and nutrition. Hence, athletes made extensive field-based observations (Eberbach & Crowley, 2009) linking science content (e.g., physiology, kinesiology) to their sport performance and evaluation. They also coordinated their *felt* experiences with sensor data and other quantitative metrics to validate and sometimes repudiate one form of data over the other—facilitating a more critical perspective of sensor-based and statistical data. Rapp and Tirabeni (2018) label this drawing upon sensation data as a form of elite expertise with personal informatics. Our analysis shows ways these felt data points were also connected to other aspects of data literacy (e.g., validating data, comparison, connecting to science-related concepts).

Our analysis suggests this *felt* sense of data and mapping to other forms of data and science concepts (e.g., muscle anatomy and physiology) helped focal athletes to be more critical of the data practices on their teams (e.g., Byron's felt sense of the deficiency of muscle strength in his body prompting him to fight back for a different training regimen). As athletes developed this critical perspective of their *felt* data, there were diverse ways they responded to team practices, which ranged from negotiating practices (e.g., Byron with his strength and conditioning coach) and quietly resisting (e.g., Lei and her teammates' collective speed ups and slow downs) to conforming (e.g., Byron's submission to his head coach's mandate to practice before medically permitted) and being shut out (e.g., Omar recognizing that though he loved the strategizing in football, it was his coaches who strategized).

While athletes themselves did not raise this point, these connections athletes were making between data and their *felt* experiences were likely also obscured from coaches and athletics staff. With the exception of Byron's explicit negotiations with his strength and conditioning coach, most other acts of resistance based on critical reflections of data were obscured from the purview of athletics staff. The individual work Lei and Byron were doing with their own data as well as Omar's individual explorations and considerations of film data and game strategy were all expressed as private endeavors not observed by their coaches or other athletics staff. The exception to this was the work athletes did with support staff (not directly evaluated by athlete performance) to manage aspects of their lives less directly related to performance. All three athletes reported ways they were actively collaborating with nutritionists to determine, implement, and track nutrition plans.

A key consideration in the role of the sociohistorical context and athletes' use of felt data on their critical data literacy practices is that athletes are both the source and producers of the data collected in athletics. This fact creates both opportunities and challenges for critical data practices in athletics. Being the producers and sources of data makes data practices more salient and relevant for learners (Clegg et al., 2017). In fact, this can lead to enhanced critical data practices (e.g., Lei and her teammates' collective speed ups and slowdowns could be viewed as an example of data hacking). However, being the producers and source of data also makes critical data practices *riskier* for athletes, particularly amidst strong power dynamics on their teams. For example, Byron's decision to acquiesce to his coaches demands he practice well before his medical clearance was the result of a data-driven decision being overridden by his coach. Even though Byron reflected critically on this decision, he did not act upon those reflections. Yet this was his body and the decision affected him both immediately and in the long term, more so even than his coaches. Yet, Byron's reflections suggest the delicate tension he needed to navigate between advocating for and with his personal data and the need to submit to coaching authority. Each of the athletes' reflections suggested they were navigating these delicate tensions with their long-term health and/or career outcomes at stake.

5.4 | Navigating "the business" of athletics

Returning to **research question 2** (*How do athletes' roles within the socio-historical political context of intercollegiate athletics shape their access to and engagement with critical data literacy practices on their teams? How do these practices inform the possibilities for CDS?*), our analysis suggests that athletes' orientations to data practices were largely driven by the sociohistorical political context of athletics and athletes' roles within this complex institution. There were two aspects we identified as shaping athletes' orientations to data practices and hence their access to and engagement with data practices on their teams: (1) local power dynamics on their teams and within the athletics institution and (2) broader systemic issues.

5.4.1 | The role of power dynamics on teams

Our findings reveal several ways in which the organizational hierarchy and distribution of power in intercollegiate athletics, mediate athletes' engagement with critical data literacy practices on their teams. First, the types of data collected as well as analysis and use of those data differed dramatically across teams. These data practices were governed by different units of athletics. For example, athletics departments governed nutrition data collected, the NCAA governed various aspects of data practices via compliance rules and regulations (National Collegiate Athletic Association, 2022). Furthermore, data are housed across these units and athletes have different access to and engagement with data in these different units. As mentioned previously, athletes could access and process nutrition and physiology data with nutritionists, but in other contexts were only shown filtered versions of data. Hence, where in the organization data were accessed mattered—it affected ways athletes could and could not engage with that data.

Athletes, coaches, and staff members' differing levels of power across positions and teams also mediated athletes' access to and engagement with data. These power dynamics were influenced by athletes' differing roles and power on their teams. For example, as a star athlete on his team with elite training prior to his arrival at State U, Byron may have had more authority to speak up against data practices and data-driven decisions than others in our study. In addition, with less professional aspirations athletes like Lei may be more able to risk the consequences of speaking up for their well-being. These aspects of their roles affected the power they had to leverage data and data practices in new ways on their teams.

Additionally, Byron, Lei, and Omar's quieter more conforming responses to position and head coach-directed practices suggest that these coaches had the most power authority, which likely afforded athletes less access to and engagement with data practices and data decisions those coaches mediate. Additionally, athletes' access to and engagement with data were influenced by differing power dynamics between teams. Our analysis suggests the revenue generation of the team impacted the amount of staffing and data collection and analysis technologies on the team as well as the level of scrutiny of team data and standardization of practices. The amount of staffing then affected the division of labor on the team and which responsibilities fell to athletes. Data such as nutrition consumption were extensively more surveilled and regulated on high revenue teams as the nutritionist not only met with and advised football players, she also selected food options at the training table, traveled with the team, and put together hotel menus during away games. While these responsibilities may have supported healthier eating habits and better performance, they limited the ways athletes could make their own choices about foods and access foods most central to their own cultures.

Our analysis suggests there may be a connection between the relational dynamics between athletes, coaches, and staff and athletes' data practices. For example, our findings suggest Byron's critical data literacy practices were deepened by his high school coach who attuned him to specific data practices and whom he trusted. When he got to State U and did not see the extent of these practices, he may have trusted the coaches' decisions less. Similarly, Omar's close relationships with coaches may have contributed to his deference to coaches' management of his data, hence limiting his data practices. Our analysis begins to reveal the complexity of these relational dynamics, but does not study them, and hence points to the need for additional research in this area.

5.4.2 | Broader sociohistorical political issues shaping athletes' data practices

Byron, Lei, and Omar's orientations toward data practices (as well as those of other focal athletes in these categories in our data set) reveal key systemic sociohistorical political issues in athletics related to justice and equity that shape data practices in athletics. First, is the commercial and competitive orientation in athletics to prioritize winning and revenue gains over athletes' long-term health and well-being. This tension is well documented in the sports and exercise science literature (e.g., Byers & Hammer, 1997; Harrison, 2000; Sack & Staurowsky, 1998). Our work points to the ways this tension and athletes' need to balance between these competing priorities drive data practices athletes engage in and those they refrain from. The drive to play competitively, avoid injury, maintain long-term health and nutrition, avoid burnout and other known diseases and unhealthy states, and maintain other aspects of their identity both drove their engagement in data practices and influenced athletes' decisions to actively disengage from data practices. Returning to the work of Jessop and Baker III (2019), our work suggests athletes' responsibilities with these data practices can help to promote their awareness and understanding of their rights and agency with data.

Additionally, even broader sociohistorical political aspects of athletics and society were touched upon in athletes' reflections on their orientations to data practices and in their day-to-day decisions about data practices. Three that are especially salient from the cases presented are the influence of socioeconomic status, gender, and race. We observe in Byron's case ways athletes' constrained resources (i.e., SES) could limit their agency with respect to data practices. This finding resonates with Duffy's (2017) notion of aspirational labor, where people labor for free in hopes of economic success in the future. Hatton (2020) similarly discussed the prevalence of coercion in college sports where power dynamics in such an aspirational context can lead athletes to be coerced into decisions like Byron's that go against their long-term health.

Similarly, Lei's avoidance of any data practices she viewed could lead to eating disorders was motivated by local observations of her teammates' experiences. However, the prevalence of eating disorders is a known issue more broadly across collegiate and professional female sports like track and field (e.g., Homan et al., 2019). Additionally, researchers point to the prevalence of athletics practices that objectify and take advantage of athletes in racially minoritized communities (Runstedtler, 2018). Limited access to data practices and hence sociohistorical reflections within sports that are predominantly made up of black athletes could be exacerbating these issues. Our work shows ways critical data literacy practices in communities can help learners engage with these broader societal issues in life-relevant ways that advance their well-being and justice. However, learners need help identifying and navigating these complex aspects of their communities.

5.5 | From critical data literacies to community-driven science

Our findings with respect to critical data literacies in intercollegiate athletics point to important considerations and implications for CDS more broadly. First, our findings suggest that *felt experiences are an inherent component of community life that should be leveraged in CDS*. Our findings related to learners' felt experiences with data build on research in data science and sports (Clegg et al., 2020; Jessop & Baker III, 2019) which illustrates the extensive data collected in athletics that are a *felt* aspect of learners' lives (e.g., performance data, sleep, alcohol consumption, heart rate, etc.). Rapp and Tirabeni's (2018) findings similarly illustrate ways athletes incorporate sensory experiences in their analysis of their quantitative sports data. Our work names and situates this concept of *felt data* within science education. Specifically, our findings show ways *felt* expertise can be used in conjunction with other quantitative and qualitative data forms (e.g., speeds, heart rate, videos of practice and games/meets) to deepen learners' engagement with science.

Linking back to CDS more broadly, we suggest that such felt experiences are inherent in community life. Other science and data science education researchers have shown ways learners leverage their physically, socially, emotionally, and economically experienced senses with other forms of qualitative and quantitative data in their communities. For example, researchers have shown ways youth connected interviews with community stakeholders, quantitative temperature data, and their own felt experiences with heat in their neighborhood to document and study the concept of urban heat islands (Barton & Tan, 2010). Others have shown ways youth leverage felt experiences in their homes, families, and neighborhoods to investigate and reason with data on topics like air quality and asthma (Auguste & Miller, 2020; Tate, 2009). Our work suggests that we recognize the *felt* expertise learners bring to bear in these contexts and ways that expertise can serve to deepen science learning in communities. We must help learners leverage these physical, social, emotional, and economic life experiences to reason deeply and guide critical insights about data that, if recognized, could broaden the fields of STEM.

Second, our findings point to the importance of *recognizing the complexity of the organizations and institutions of communities* and ways their organization shapes science learning experiences. Science education research on youth agency situates the need for helping youth take civic action with science and points to ways to give youth platforms (e.g., blogs, videos, community gatherings) to inform and interact with community members (e.g., Barton & Tan, 2010; Polman, 2010; Polman & Miller, 2010). Similarly, Vakil (2018) argues for the importance of situating computational fluency experiences within its sociopolitical context and helping youth think critically about the effects of computation in their communities.

Similar to these bodies of work, our study points to ways joint community endeavors facilitate critical perspectives and taking action in ways that leverage and promote science learning. Yet, we also point out the role of the complex organization of communities and larger institutions that influence them on data, and hence science, practices in communities. Our findings point to the ways community organizations and governing structures are multi-layered. We show that the influences of learning in one class, session, or program go beyond that immediate environment. In CDS, we therefore must also study how these complex organizational structures both facilitate and inhibit engagement with and access to science. Additionally, our findings point to the need to recognize the role of power—and how power can be different between learners, stakeholders, sessions, programs, and organizations—on

learners' science engagement and access. Specifically, our findings suggest communities with extensive divisions of labor and strong power dynamics and hierarchies may limit learners' access to and engagement with science, especially the more critical aspects of science.

These findings point to the need to help learners navigate the complex power dynamics and organizational contexts of the communities in which they are engaged. We especially draw attention to the importance of identifying community members whose roles and power lend themselves to support roles (e.g., nutritionists and athletics staff not evaluated on athletes' sport performance in our context). Additionally, we point to the importance of not only helping learners leverage science to inform their communities, but also of helping them access community decision-makers and effectively leveraging science to advocate for themselves and their community constituents. In the context of collegiate athletes, this could involve helping athletics staff to see themselves as educators and to find ways to make the science practices underlying their work with athletes transparent and visible to the athlete, in a way that gives the athlete agency with their data and the surrounding practices. Furthermore, incorporating critical aspects of science throughout K-12 and higher education could help athletes to develop the language, expectations, and data prowess to leverage data toward their own long-term goals and to question the powers that be when their agency with data are limited.

6 | CONCLUSION

In conclusion, our data showcasing the rich critical data literacies athletes bring to bear—even amid limitations posed on their agency, access, and engagement with data—provide critical counter stories to the argument, fueled by racist stereotypes, assumptions, and biases that athletics is not academic. The rich mechanical and socio-technical engagement of athletes in critical data literacies provides a qualitative account of athletes' engagement with data literacy. Such literacies have the potential to position athletes as having expertise to bring to bear to data literacy practices and the many sciences that draw heavily on such practices.

Yet, our counter stories are in tension with quotes like Byron's assertion that many athletes are not athletically inclined. His reflection shows that science education has a long road ahead in helping learners in communities not typically associated with dominant narratives of science to recognize and be acknowledged for the science literacies they bring to bear in their everyday lives. Perhaps positioning athletes as data experts in initiatives to engage all learners in critical data literacies could help athletes to be recognized for their academic prowess and help "the crowd" (i.e., coaches, administrators, and fans) to develop the critical data literacies that enable and inspire them to critique and fight back against data practices in athletics that lead to negative outcomes for athletes and ward off (at least to some extent) toxic cultures in athletics.

ORCID

Tamara L. Clegg https://orcid.org/0000-0001-6558-6944

ENDNOTES

[1] Film is a term of art in sports referring to video data collection and analysis. Although the term connotates analog machinery, film review tools are primarily now digital, powered by complex software systems.

[2] While sports analytics practices are common at all levels of sports, they are much more pronounced, complex, and state of the art in collegiate programs, especially Division I sports.

[3] We use the term STEM to refer to the integration of engineering into science learning goals and experiences as outlined by the Next Generation Science Standards. We do not refer to STEM as Science, Technology, Engineering and Mathematics.

REFERENCES

Auguste, D., & Miller, S. L. (2020). Volatile organic compound emissions from heated synthetic hair: A pilot study. *Environmental Health Insights*, *14*, 1178630219890876.

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

Barton, A. C., & Tan, E. (2010). We be burnin'! Agency, identity, and science learning. *The Journal of the Learning Sciences*, *19*(2), 187–229.

Byers, W., & Hammer, C. H. (1997). *Unsportsmanlike conduct: Exploiting college athletes*. University of Michigan Press.

Calabrese Barton, A., Greenberg, D., Turner, C., Riter, D., Perez, M., Tasker, T., Jones, D., Herrenkohl, L. R., & Davis, E. A. (2021). Youth critical data practices in the COVID-19 multipandemic. *AERA Open*, *7*, 23328584211041630.

Clegg, T., Norooz, L., Kang, S., Byrne, V., Katzen, M., Velez, R., ... Froehlich, J. (2017). Live physiological sensing and visualization ecosystems: An activity theory analysis. *Conference on Human Factors in Computing Systems—Proceedings*, *2017-May*, 2029–2041. https://doi.org/10.1145/3025453.3025987

Clegg, T. L., Greene, D., Beard, N., & Brunson, J. (2020). Data everyday: Data analytics practices in a division I sports context. In *Proceedings of SIGCHI human factors in computing systems (CHI 2020)*. ACM.

Comeaux, E. (2018). Stereotypes, control, hyper-surveillance, and disposability of NCAA division I black male athletes. *New Directions for Student Services*, *2018*(163), 33–42.

D'Ignazio, C., & Bhargava, R. (2015). Approaches to building big data literacy. *Proceedings of the Bloomberg Data for Good Exchange Conference*.

Duffy, B. E. (2017). *(Not) getting paid to do what you love: Gender, social media, and aspirational work*. Yale University Press.

Eberbach, C., & Crowley, K. (2009). From everyday to scientific observation: How children learn to observe the biologist's world. *Review of Educational Research*, *79*(1), 39–68.

Galbraith, J. R. (2014). Organizational design challenges resulting from big data. *Journal of Organization Design*, *3*(1), 2. https://doi.org/10.7146/jod.8856

Greenbaum, D. (2018). Wuz you robbed? Concerns with using big data analytics in sports. *American Journal of Bioethics*, *18*(6), 32–33.

Halson, S. L., Peake, J. M., & Sullivan, J. P. (2016). Wearable technology for athletes: Information overload and pseudoscience? *International Journal of Sports Physiology and Performance*, *11*(6), 705–706.

Harrison, C. K. (2000). Black athletes at the millennium. *Society*, *37*(3), 35–39.

Hatton, E. (2020). *Coerced: Work under threat of punishment*. Univ of California Press.

Holmes, A. G. D. (2020). Researcher positionality—A consideration of its influence and place in qualitative research—A new researcher guide. *Shanlax International Journal of Education*, *8*(4), 1–10.

Homan, K. J., Crowley, S. L., & Sim, L. A. (2019). Motivation for sport participation and eating disorder risk among female collegiate athletes. *Eating Disorders*, *27*(4), 369–383.

Jessop, A., & Baker, T. A., III. (2019). Big data bust: Evaluating the risks of tracking NCAA athletes' biometric data. *Texas Review of Entertainment & Sports Law*, *20*, 81.

Johnson, B., Rydal Shapiro, B., DiSalvo, B., Rothschild, A., & DiSalvo, C. (2021). Exploring approaches to data literacy through a critical race theory perspective. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15.

Jones, S. T., Thompson, J., & Worsley, M. (2020). Data in motion: Sports as a site for expansive learning. *Computer Science Education*, *30*(3), 279–312.

Kretzmann, J., & McKnight, J. P. (1996). Assets-based community development. *National Civic Review*, *85*(4), 23–29.

Lee, O., & Campbell, T. (2020). What science and STEM teachers can learn from COVID-19: Harnessing data science and computer science through the convergence of multiple STEM subjects. *Journal of Science Teacher Education*, *31*(8), 932–944.

Lee, V. R., & Dubovi, I. (2020). At home with data: Family engagements with data involved in type 1 diabetes management. *Journal of the Learning Sciences*, *29*(1), 11–31.

Lubker, J. R., & Etzel, E. F. (2007). College adjustment experiences of first-year students: Disengaged athletes, nonathletes, and current varsity athletes. *Journal of Student Affairs Research and Practice*, *44*(3), 855–878.

Malone, J. J., Lovell, R., Varley, M. C., & Coutts, A. J. (2017). Unpacking the black box: Applications and considerations for using GPS devices in sport. *International Journal of Sports Physiology and Performance*, *12*(s2), S2–S18.

Mathie, A., & Cunningham, G. (2003). From clients to citizens: Asset-based community development as a strategy for community-driven development. *Development in Practice*, *13*(5), 474–486.

Maybee, C., & 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), 1–4. https://doi.org/10.1002/pra2.2015.1450520100108

Milan, S. (2019). Acting on data (fication). In *Citizen media and practice* (pp. 212–226). Routledge.

Nasir, N. S., & Hand, V. (2008). From the court to the classroom: Opportunities for engagement, learning, and identity in basketball and classroom mathematics. *Journal of the Learning Sciences*, *17*(2), 143–179. https://doi.org/10.1080/10508400801986108

National Collegiate Athletic Association. (2022). *NCAA division I manual*. The National Collegiate Athletic Association (NCAA).

Pangrazio, L., & Selwyn, N. (2019). 'Personal data literacies': A critical literacies approach to enhancing understandings of personal digital data. *New Media & Society*, *21*(2), 419–437.

Paris, D., & Alim, H. S. (2014). What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harvard Educational Review*, *84*(1), 85–100.

Patel, D., Shah, D., & Shah, M. (2020). The intertwine of brain and body: A quantitative analysis on how big data influences the system of sports. *Annals of Data Science*, *7*(1), 1–16.

Philip, T. M., Schuler-Brown, S., & Way, W. (2013). A framework for learning about big data with mobile technologies for democratic participation: Possibilities, limitations, and unanticipated obstacles. *Technology, Knowledge and Learning*, *18*(3), 103–120.

Piwek, L., Ellis, D. A., Andrews, S., & Joinson, A. (2016). The rise of consumer health wearables: Promises and barriers. *PLoS Medicine*, *13*(2), e1001953.

Polman, J. L. (2010). The zone of proximal identity development in apprenticeship1 learning La zona de desarrollo próximo de la identidad en entornos de aprendizaje de oficios. *Revista de Educación*, *353*, 129–155.

Polman, J. L., & Miller, D. (2010). Changing stories trajectories of identification among African American youth in a science outreach apprenticeship. *American Educational Research Journal*, *47*(4), 879–918.

Prado, J. C., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, *63*(2), 123–134. https://doi.org/10.1515/libri-2013-0010

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, *1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Rapp, A., & 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), 1–30. https://doi.org/10.1145/3196829

Runstedtler, T. (2018). More than just play: Unmasking black child labor in the athletic industrial complex. *Journal of Sport & Social Issues*, *42*(3), 152–169.

Sack, A. L., & Staurowsky, E. J. (1998). *College athletes for hire: The evolution and legacy of the NCAA's amateur myth*. Praeger Publishers.

Saldaña, J. (2015). *The coding manual for qualitative researchers*. Sage.

Shields, M. (2005). Information literacy, statistical literacy, data literacy. *IASSIST Quarterly*, *28*(2), 6. https://doi.org/10.29173/iq790

Špiranec, S., Kos, D., & George, M. (2019). *Searching for critical dimensions in data literacy*. Proceedings of the Tenth International Conference on Conceptions of Library and Information Science, Ljubljana, Slovenia, The university of borås: Sweden, vol. 24.

Tate, E. D. (2009). *Asthma in the community: Designing instruction to help students explore scientific dilemmas that impact their lives*. University of California.

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, *4*(2), 2053951717736335.

Tygel, A. F., & Kirsch, R. (2016). Contributions of Paulo Freire for a critical data literacy: A popular education approach. *The Journal of Community Informatics*, *12*(3), 108–121.

Vakil, S. (2018). Ethics, identity, and political vision: Toward a justice-centered approach to equity in computer science education. *Harvard Educational Review*, *88*(1), 26–52.

Valerio, M. A., Rodriguez, N., Winkler, P., Lopez, J., Dennison, M., Liang, Y., & Turner, B. J. (2016). Comparing two sampling methods to engage hard-to-reach communities in research priority setting. *BMC Medical Research Methodology*, *16*(1), 1–11.

Van Wart, S., Lanouette, K., & Parikh, T. S. (2020). Scripts and counterscripts in community-based data science: Participatory digital mapping and the pursuit of a third space. *Journal of the Learning Sciences*, *29*(1), 127–153.

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

Weight, E. A., Cooper, C., & Popp, N. K. (2015). The coach-educator: NCAA Division I coach perspectives about an integrated university organizational structure. *Journal of Sport Management*, *29*(5), 510–522.

Weight, E. A., Harry, M., & Navarro, K. (2020). Integrating athletics within the academy: Educational experiences of athletes, musicians, and traditional students. *Journal of Issues in Intercollegiate Athletics*. *13*, 143–169.

Weight, E. A., & Huml, M. R. (2016). Education through athletics. *Journal of Intercollegiate Sport*, *9*(2), 352–378.

Weight, E. A., Navarro, K. M., Smith-Ryan, A., & 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), 38–60. https://doi.org/10.15763/issn.2376-5267.2016.1.1.38-60

Wilkerson, M. H., & Polman, J. L. (2020). Situating data science: Exploring how relationships to data shape learning. *Journal of the Learning Sciences*, *29*(1), 1–10.

Zimmermann-Niefield, A., Turner, M., Murphy, B., Kane, S. K., & 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, 121–132. https://doi.org/10.1145/3311927.3323139

How to cite this article: Clegg, T. L., Cleveland, K., Weight, E., Greene, D., & Elmqvist, N. (2023). Data everyday as community-driven science: Athletes' critical data literacy practices in collegiate sports contexts. *Journal of Research in Science Teaching*, *60*(8), 1786–1816. https://doi.org/10.1002/tea.21842