Supporting Data-Driven Basketball Journalism through Interactive Visualization

Yu Fu
Georgia Institute of Technology
Atlanta, GA, USA
fuyu@gatech.edu

John Stasko
Georgia Institute of Technology
Atlanta, GA, USA
stasko@cc.gatech.edu

ABSTRACT

Basketball writers and journalists report on the sport that millions of fans follow and love. However, the recent emergence of pervasive data about the sport and the growth of new forms of sports analytics is changing writers' jobs. While these writers seek to leverage the data and analytics to create engaging, data-driven stories, they typically lack the technical background to perform analytics or efficiently explore data. We investigated and analyzed the work and context of basketball writers, interviewed nine stakeholders to understand the challenges from a holistic view. Based on what we learned, we designed and constructed two interactive visualization systems that support rapid and in-depth sports data exploration and sense-making to enhance their articles and reporting. We deployed the systems during the recent NBA playoffs to gather initial feedback. This article describes the visualization design study we conducted, the resulting visualization systems, and what we learned to potentially help basketball writers in the future.

CCS CONCEPTS

• **Human-centered computing → Visualization**; **Visualization application domains**; **Visualization systems and tools**.

KEYWORDS

data-driven journalism, sports data visualization, user interface

ACM Reference Format:
Yu Fu and John Stasko. 2022. Supporting Data-Driven Basketball Journalism through Interactive Visualization. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3491102.3502078

1 INTRODUCTION

The NBA (National Basketball Association) is the top-level professional basketball league in the United States and Canada. With millions of fans throughout the world, the NBA garners tremendous attention and is reported on by a large number of sportswriters literally daily. These writers range from professional journalists working for powerful media companies to free-lance sportswriters, both professional and amateur, who write reports and blogs about the NBA. They excel at telling vivid stories about games, teams, and players by observing games, collecting information, conducting interviews, and weaving it all together into complete stories.

The work of NBA writers has become more challenging due to the recent wide availability of data and statistics about the league and the rise of new types of *sports analytics* [37, 48, 49] about basketball. While data and statistics have been pervasive in sports throughout history, the NBA has taken two steps to massively increase available game data. First, the league outfitted each stadium with sophisticated video cameras that track the ball and all players, providing second-by-second location information [48]. Second, the league has created data output pipelines on the internet/web distributing data such as play-by-play information and data and statistics about plays. A video of each action in a game is also available.

Another challenge to basketball writers is the development of new forms of sports analytics about the game: advanced statistical models and/or the gathering and computational analysis of large data sets. The sport of baseball was the first to be impacted by these new forms of sports analytics. The book and subsequent movie *Moneyball* [31] helped show how one baseball team used analytics in team management, game planning, player development, and even player health management. Now, baseball teams routinely use sports analytics in an attempt to create a competitive edge.

While baseball has been the clear leader in adopting analytics, basketball has not been far behind. Many new types of data about the sport have emerged, as well as new analyses on those data to measure this more dynamic team sport. Such data and analyses include, for example, the efficiency of specific types of offensive plays for each player/team and the opponents' three-point shot accuracy when a subset of players is in the game.

Although we characterized the wide availability of new forms of data and the emergence of new analytics as "challenges" to basketball writers, the two can also be viewed as opportunities. The data and analytics could be an integral part of innovative new stories and narratives about why teams and players succeed and why games ended as they did. However, NBA writers are typically not trained statisticians, nor do they know how to structure and manipulate databases or how to write software to do analysis. Most NBA writers will not be writing R programs or building D3 visualizations.

How then can we make the massive data and innovative analytics available to writers? How do we help writers rapidly sift through the vast amount of data generated in games to find reliable and helpful information that supports the narratives and makes the narratives more credible? How do we help writers construct new narratives based on data and analytics in more engaging ways and present their insights to broader audiences in a more appealing fashion?

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3502078

Our approach in this work is to use interactive visualization to make these new data and analytics widely available and easily accessible to NBA writers. While most writers are not trained statisticians or programmers, they are usually comfortable working with and manipulating interactive interfaces one typically finds on the Web. Furthermore, visualization provides an opportunity to bridge from complex data and analytics to a more comprehensible understanding of what has occurred. In fact, NBA basketball already has examples of this. Visualizations like Kirk Goldsberry's pioneering shot charts [19] have, arguably, changed the course of modern-day basketball. Shot charts show, superimposed on the outline of a basketball court, where each shot was taken and whether it was successful. Goldsberry's book, *SprawlBall* [21], combines visual representations, narrative storytelling, and cartoons to illustrate the impact that analytics has brought to modern-day basketball as well as the evolution of shot location selection in the NBA.

In this research, we undertake a type of *visualization design study* [34, 36, 47] on the work and context of NBA writers and journalists. To help identify their needs and challenges, we adopted a mixed-methods approach, including ethnographic observations, a survey of the state-of-the-art, and interviews with nine stakeholders. By connecting and interpreting our findings, we identify a desire for more interactive visualization tools to aid NBA writers in exploring massive amounts of data, finding insights, and constructing narratives. We then designed and developed two interactive visualization systems, **NBA GameViz** and **NBA LineupViz**, based upon our findings.

To evaluate the systems, we deployed them in three ways to gather feedback: (1) Expert evaluation: we recruited five sportswriters/analysts to use the systems during the NBA Finals, and we gathered feedback from them. (2) Self-deployment: we used the systems to help publish over 300 visual articles/posts on different media platforms and gathered over a thousand comments from end audiences. (3) In-the-wild deployment: we deployed the systems on the web to foster real-world use and collect spontaneous feedback. Finally, we reflected on the design process and evaluation feedback and discussed the lessons learned during our study.

2 BACKGROUND AND RELATED WORK

2.1 Data-driven Sports Writing and Journalism

In the digital media era, game and interview content about sports are more accessible to broad audiences, which has shifted the landscape of sports storytelling. Audiences do not settle for just knowing "what" – they are becoming more interested in "why." This shift in audiences' taste urges journalists to add a more analytical slant to their work. Howard introduced the art and science of data-driven journalism and discussed the potential, the pitfalls, and the challenges throughout the adoption of data journalism [24]. He listed 14 recommendations and predictions, including "better tools will emerge that democratize data skills" and "data journalism will be held to higher standards for accuracy and corrections." Wiske and Horky thoroughly discussed the "new and substantial challenges on sports communicating" placed by digitalization, saying that "the quality of sports journalism is being tested." [64]. Hammond discussed the relationship and differences between "computer-assisted journalism" and data-driven journalism [23]. Diakopoulos's extensive work on computational journalism [12–14] explored multiple aspects within this field, from coining the concepts, examining the paradigm shift, to mapping out the new challenges and opportunities. This prior research informed and inspired us to understand and investigate the problem space from a holistic perspective and in its interdisciplinary context.

2.2 Sports Analytics

Data and statistics have long been used to measure and communicate sports participants' performances. Much of the data about sports is simply counting occurrences of events: wins, losses, points, and assists, for example. Over the last 10-15 years, however, a series of richer and more complex statistics and statistical models have emerged. The development of some of these so-called *analytics* has resulted in part from larger data sets becoming available and computational analysis of that data being performed. The field of sports analytics [35, 63] now is a rich and deep area driven by both academics and practitioners. It involves a comprehensive process from collecting data, processing data, analyzing data using statistical models and algorithms, to presenting data to stakeholders.

Basketball is a dynamic and fluid team sport. Each individual player in the game is involved in results and events. It is a team working in coordination that succeeds or fails. The continuous flow of a game and the team dynamic makes it very difficult to analyze statistically. Shea and Baker note, "the statistics typically captured in an NBA box score are not sufficient"[49]. Driven by the demands of sports analytics and powered by video tracking and wearable technology [8, 15, 62], an unprecedented amount of data has begun to be collected and analyzed in professional basketball. A broad spectrum of algorithms and models have been developed to quantify and analyze the complexity of the game. Dean Oliver, an eminent contributor to basketball analytics, published his book *Basketball on Paper* [37] in 2004, laying a foundation for the field. Shea and Baker introduced and discussed a variety of advanced metrics in their book *Basketball Analytics* [49] with topics ranging from player evaluations to team construction. Shea's following book *Sports Analytics-Spatial Tracking* introduced methods for analyzing the newly-emerging spatial tracking video data captured by SportVU [48]. Seth Partnow, a former director of basketball research for an NBA team, recently published his book *The Midrange Theory* [40]. Partnow gave a walkthrough of what basketball analytics is about and what it is not, as well as the conceptual foundation of basketball analytics and ongoing challenges.

2.3 Sports Data Visualization

Perin et al. conducted a thorough review of both academic and practitioner contributions to sports data visualization [42]. They pointed out two primary roles of sports data visualization: analytical (or exploratory) purposes and narrative (or communicative) purposes. These two purposes align well with sportswriters' roles: to analyze a game and to communicate insights about it to audiences. Perin et al. also categorized sports data into three overarching types: box-score data, tracking data, and meta-data. The visualizations that we develop primarily represent box-score data, but also some tracking (event) data and advanced metrics derived from these two.

Recently, the sport of soccer has been the focus of much data visualization work. The SoccerStories system [41] shows players' movements on a soccer pitch and assists analysts in understanding strategies of play that might lead to more success. The visualizations present an image of the pitch and represent players as circles and their movements as lines on the pitch. Sacha et al. similarly present soccer player movements on a pitch, showing trajectories and patterns, with a focus on abstracting and aggregating to handle the large amount of data that often leads to overplotting [46].

For the sport of basketball, our focus, Goldsberry's shot maps have become a well-known example of sports data visualization [19, 20]. Whitehead has turned "boring statistics" into appealing, stylish visualizations, such as the *Passing Lane Chart* [61], and *Hot Hand* [60]. Phillips has done similar work, such as the *25 Most Common Shot Locations by Teams* [43]. And Taylor, the author of *Thinking Basketball* [52], has added animated charts into his analytical video series, such *Deep Analysis of Larry Bird's Impact* [54].

Besides these author-driven explanatory practices, more commercial sports data websites have adopted exploratory visualizations or visualization techniques to assist data exploration and analysis. Analytic-oriented websites, like Cleaning the Glass [55], Bball Index [25], and ShotQuality [44], have leveraged color-coded spreadsheets as well as some standard visualizations like scatter plots. Pivot Analysis [1] applied more advanced interactive visualizations, such as heatmaps, beeswarm charts, and parallel coordinates plots.

On the more academic research side, Chen et al. introduced multiple narrative visualization techniques to visualize NBA games in the GameFlow [10] system. They provide court views, temporal event views, statistics, and links to videos of plays. BKViz [32], developed by Losada et al., is a visual interface dedicated to analyzing single basketball games. It contains views or representations for a game outline/overview, for showing a play sequence, for a player's shot pattern, team player rotations, and other statistically-focused analyses. Metoyer et al. proposed a text-visualization coupling approach leveraging natural language processing, quantitative narrative analysis, and a custom coupler [33]. They demonstrated this approach with a basketball article linked to corresponding visualization components and states. Zhi et al. developed GameViews [65], two visualization prototypes to support data-driven sports storytelling for both sports journalists and sports fans. The authors observed a professional journalist and reviewed a corpus of basketball game recaps to drive their work. The project also focused on the fan experience. The system's interface included a line chart-style game flow chart with textual play-by-play summaries and statistics integrated with the game video.

Our visualization systems not only provide access to a larger collection of sports data but also compose existing artifacts, standard visualizations, and custom views. In addition, we augment the visuals with further interactive controls that enable basketball writers and analysts to construct distinctive narratives by focusing on specific segments of provided data. With functionalities such as ubiquitous connection, brushing and global updating, and universal linking to game clips, our systems aim to contextualize the insights and make components more actionable from a domain perspective.

3 UNDERSTANDING THE PROBLEM

To better understand the challenges faced by basketball writers, we conducted an investigation into their current work, practices, and existing tools. Our examination followed a mixed-methods approach that included ethnographic observation, a review of the current state-of-the-art, and interviews with nine working basketball journalists/analysts. We then conducted open coding and thematic analysis to distill the emerging themes. These activities were a portion of our overall study process, as shown in Figure 1.

3.1 Methodology

Ethnographic Observation. Our ethnographic observation consisted of work setting observation, existing artifacts observation, and online work and community interaction observation.

Documenting the work settings: During the recent NBA season, we conducted a field study by visiting an NBA media room, a typical workplace for NBA journalists. One author took a standby role and observed the artifacts made available to journalists by NBA teams in the media room, their characteristics, and how journalists used them. We documented the artifacts via pictures, physical objects such as game notes (statistic details), stat sheets, and field notes.

Browsing sports data websites: Sports data websites, either official or informal, have become a primary medium for sportswriters and fans to explore massive sports data. Additionally, they reflect the state of the art of sports analytics in the public space and the norms of sports data usage. To further investigate the characteristics of data, first, we collected twelve frequently used basketball stats websites. Then we summarized the features for each one regarding the traits of the data, the usage of visual representations, and the interactions that support data exploration.

Observing online work and community interactions We started following many sports journalists, sports visualization practitioners, and sportswriters on sports media and social media platforms. We observed their posts/work, including narratives, articles, and visual content, as well as the comments about them. This part served three purposes: (1) a formative approach to acquire contextual information about the domain in broader social settings; (2) a preliminary approach for us to build up the collection of sports visualization practice, and narrow down the candidate interviewee pool; (3) inform and inspire our visual and interaction design.

Sports Visualization Practice Review. To understand the application of sports visualization in data-driven journalism, we collected 55 data-driven sports articles and 84 social media posts. The inclusion criteria were that the selected article/post must include at least one visualization. We started the collection by browsing through websites and searching for examples that include basketball visualization. The articles came from five media platforms, including three major sports media outlets (ESPN, the Athletic, Tencent) and two data-driven media outlets (FiveThirtyEight, Nylon Calculus). Finally, we examined both collections along four dimensions: the data, the chart types, the visual techniques, and the narratives/topics.

Expert Interviews. After reviewing the data-driven sports articles and posts (Twitter/Instagram), we sent recruiting emails and direct messages to 24 candidate experts. To gain a holistic view,

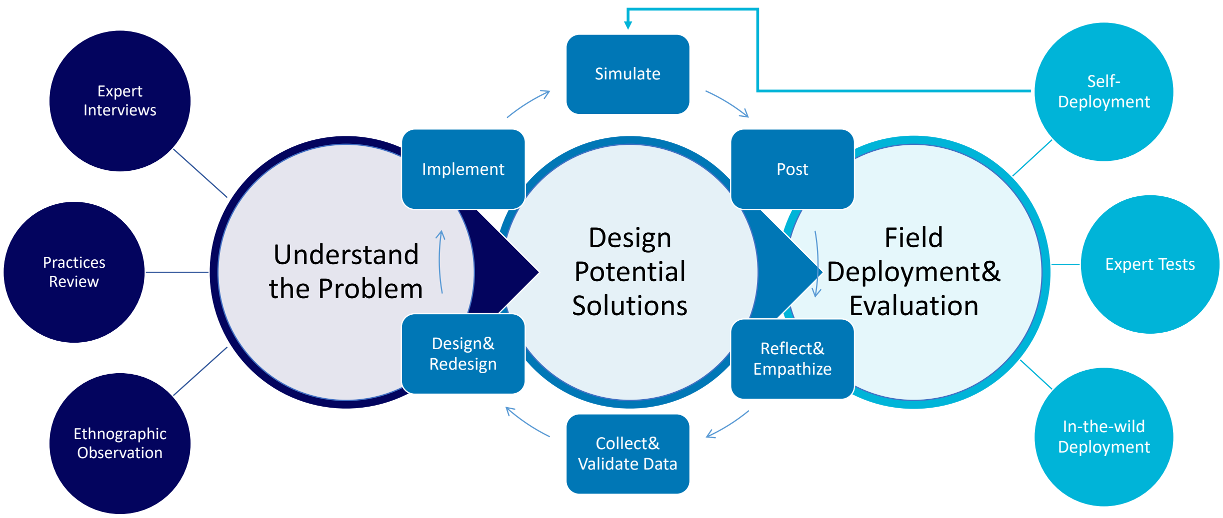

Figure 1: An overview of our study: three major parts that followed a sequential order but also were highly iterative and dynamic. For example, the design process helped us better understand the problem, and self-deployment played a crucial role in helping us evaluate the design rapidly.

we included both primary stakeholders (sportswriters) and secondary stakeholders (e.g., sports analysts, basketball visualization practitioners) in our candidate pool. We asked the nine experts who agreed to participate in the interviews to fill out a consent form and a preliminary questionnaire to gauge their background and experience. We categorized experts into three different roles: **Sports journalists (J1-J3)**: three credentialed NBA journalists who have been covering NBA games for major media outlets, including one who has a voting right in the NBA awards. **Sports visualization practitioners (P1-P3)**: three influential sports visualization practitioners who approach sports storytelling with visualizations. **Sports analysts (A1-A3)**: three sports data analysts who write articles/posts based on sports data and/or build models to generate advanced stats, including two sports analytical data website founders. The interviews were semi-structured, and the protocols were designed based on their roles, the answers from the questionnaires, and the notes from reviewing individual work. Each included: **Core section**— uniform questions that we asked every expert (e.g., their experience with visualization and sports analytics, the challenges of using them); **Group-specific section**— questions based on experts' primary roles (e.g., the impact of visualization and sports analytics on their work); **Individual section**— questions based on their unique experiences and individual work (e.g., their experience using sports visualization to write articles/ teach programming). The interviews were conducted remotely, recorded, and kept as video files. The conversations lasted approximately 30-60 minutes and were transcribed and validated for our data analysis. Each participant was compensated with a $15 Amazon gift card.

3.2 Findings

We conducted method triangulation [7] and consolidated the emerging themes to five primary findings (F1-F5):

F1. Towards faster yet credible digital journalism. Though truth-telling has always been the core of journalism, two other features, **user participation** and **a fast continuous news cycle**, have been catching up as journalism shifts towards digitalization [27]. Sports journalism, in particular, is facing the challenge of being faster. Due to the nature of sports games, immediacy exerts a significant impact on audience engagement. Both J1 and J2 spoke on the urgency of sports writing, with J1 stating that "*it is different from doing newspaper now, you have to write something immediately after the game, or even before the game ends.*" While immediacy has become an essential and valuable feature of online journalism [38], it threatens the quality of journalistic reporting [5, 27, 39] because finding insights and evidence often costs additional time, and being immediate could result in dubious and incomplete reporting [27]. In sports journalism, narratives and arguments are often subjective. Our interviewees pointed out that many hardcore sports fans expect them to support narratives with solid evidence. J2 told us that "*it is hard to convince people, but if you want to make an argument, you have to use something to back it up.*" J1 also stated that "*those hardcore fans really care about whether your analyses are convincing. You want them to be solid and accurate; you cannot just write nonsense. Because fans nowadays are knowledgeable. They can detect those nonsenses and hold you accountable.*".

F2. The increasing role of sports analytics in sports journalism. For many sportswriters and journalists, sports analytics has become an essential approach to narrative construction. Sports data, consequently, presents itself as one of the primary evidence sources, along with game video analysis, to support their narratives. A3 shared that his sports data website has many journalists and blogger subscribers, "*we partnered with the Ringer, Bleacher Reports, Forbes, and NBA.com.*" J3 added that "*if you are not using data to make arguments or at least part of your arguments, you are* *beaten. Some people are doing it, and if you are not doing it, it is going to be much harder for you to get a job...The stats sheets offered by NBA teams in the media room are very vanilla...I need to use these websites to look up advanced stats.*" In the 55 data-driven stories and 84 visualization posts we examined, advanced metrics/data (computed by algorithms/models) were most frequently used. 37 articles (67.3%) and 40 posts (47.6%) used advanced metrics/data while only 8 articles (14.5%) and 13 posts (15.5%) used traditional boxscore statistics.

On the other hand, sports analytics create usage contexts for sportswriters and journalists to use the sports data. For example, the "Four Factors of Basketball Success," analytic metrics proposed by Dean Oliver [37] have become widely accepted to evaluate teams' performance in an NBA game. However, without a sophisticated knowledge of proper usage contexts, sports data can be presented in misleading ways. Nevertheless, more sportswriters are accustomed to incorporating sports data and sports analytics into their work. A2 stated "*I feel like most serious reporters now are just at least comfortable with citing statistics in the most comments, advanced statistics in the NBA, or whatever sport they are in a lot.*".

F3. Spreadsheets are prevalent to explore and present data. The expanding sports data collected and delivered by sports statistics providers such as Sportradar and Second Spectrum spread across various sports websites. They consist of official stats websites (e.g., nba.com/stats), independent sports data services (e.g., Synergy, Basketball Reference, Cleaning the Glass, etc.), and news websites (e.g., ESPN, Yahoo, etc.). After carefully examining twelve websites and the physical objects we collected from the NBA media room, we found that spreadsheets are prevalent in presenting these data. None of the game notes (around 80 pages) and stats sheets accessible to journalists in the NBA media room utilize visualization. Out of the twelve websites we examined, all the websites had sortable spreadsheets. Two had color-coded columns. While seven had utilized at least one visualization, the use of visual representations and interactions were limited. The most used visualizations are bar charts displaying counts. Interactivity on these websites was very minimal.

F4. Static author-driven visualizations are prevalent. While visualizations have been increasingly used in sports journalism, only a handful have been used to support sportswriters conducting in-depth exploratory data analysis. As we mentioned in F3, the websites we investigated only supported interactive visualization for a tiny portion of their data, not to mention supporting in-depth analysis. Some applications from sports analytics community have increased the use of visualization. But the capability of interactive visualization has not been fully exploited yet. Advanced exploratory interactive visualization like Beshai's Buckets [3] and Scattershots [4] are rare finds. Most practices focused on leveraging the communicative power of visualization and tended to be author-driven [57]: analysts used static visualizations to communicate the findings from their statistical data analysis. None of the data-driven articles we examined provided interactivity. Only two articles (3.6%) have screenshots from an interactive visualization interface. Moreover, the 55 data-driven articles were from only ten sportswriters. Only a small number of data-savvy analysts appeared to have the capability to make appealing and effective communicative visualizations. P2, a well-known practitioner, told us that he "*has never used interactive visualization*" to explore data. P3 was the only practitioner we interviewed who had used visualization for insight finding rather than communication. However, as he added, "*it is not for exploring data... it is more of confirmatory analysis.*" P3 also addressed the observation of prevalent communicative visualization, arguing that these author-driven visualizations can be abused,"*you are only showing a very small subset, a selected subset of the data... and people cannot tell.*" One representative case was Goldsberry's shot chart. Some criticism from the analytics community has been directed at Goldsberry, citing that his "Mid-range is dead" [18] type of narrative-driven visualizations overlook the differences in players' roles and skillsets, as well as the context of the rule changes and defensive schemes [2, 53], resulting in misunderstanding from the general audiences and some media. P1 also spoke on the issue, mentioning that exploratory visualizations can "*let the users have a stake in what they are viewing*" and "*empower them to understand the data on a more thorough level.*"

F5. Towards broader sports media and fluid roles. Before the advent of the Internet era, sports writing was primarily performed by professional sports journalists. The growth of the internet, especially social media, has invited more independent writers to contribute. Many hobbyists have leveraged broader access to sports data and game videos to inform their analysis and narratives. Social media allows them to share content and transition to independent writers or even professional journalists. There are also cases of team analysts/management becoming journalists. Meanwhile, high-profile media outlets and journalists with more sources/resources have taken on most news-breaking roles and exclusive content. Social media also made the direct interaction between players and fans prevalent, leaving even less "reporting" room to other journalists. Consequently, more journalists are shifting towards "analyst" roles, and lines between these roles have become increasingly blurred.

To help these broader sportswriters through visualizations, we first acknowledge their diverse backgrounds and niche audiences to cater to, *"they have fundamental differences in their needs, motivations, and goals as well as their capabilities"* [28]. As J1 mentioned, "*many sports journalists received liberal art educations, they do not have much engineering training and knowledge*" and "*it is unrealistic for them to create a visualization with coding.*" (J2). The next step is to identify the "pain points" where computing can intervene. Through our interview with experts and the review of basketball writers' work, we summarized their storytelling/narrative construction process. They started by taking a "insight seeker" role and analyzing external data sources, such as sports data, game observation, and game videos. Then they took the "insight communicator" role and constructed and communicated their narratives based on the insights they distilled. These narratives were presented as media products, such as articles, podcasts, and videos. The last step in the process is to engage audiences. We concluded the three tasks during the process: (1) analyzing external data and finding insights, (2) communicating insights through media products, (3) engaging audiences to consume the products. These tasks echo with the two uses of sports visualizations: analytical (exploratory) and narrative (communicative) [42]. Besides exploring data, seeking insights, and

FINDINGS	TAKEAWAYS	GENERIC DESIGN IMPLICATIONS
F1. Towards faster yet credible digital journalism	**T1.** Sportswriters need to write faster while maintaining accountability	**DI1.** Automate the data pipeline, support rapid large-scale visualization **(T1, T4, T8)**
F2. The increasing role of sports analytics in sports journalism	**T2.** Sports analytics offers ways to construct narratives, and sports data is used as evidence to "back them up"	
	T3. There are a massive amount of sports data, and visualizing them requires defining specific usage contexts to decompose fuzzy tasks	**DI2.** Be resilient to diverse needs and reporting styles **(T7, T8)**
	T4. Sports data are constantly updating after each game	
F3. Spreadsheets are prevalent to explore and present data	**T5.** Spreadsheets advantages: showing exact numeric values and their rankings, etc. Disadvantages: high cognitive bandwidth, error-prone and tedious [22]	**DI3.** Accelerate information retrieval and sense-making process **(T1)**
F4. Static author-driven visualizations are prevalent	**T6.** Interactive visualization systems can intervene and help journalists explore the massive data and make sense of it	**DI4.** Support "drilling-down" exploration and analysis with interactivity **(T1, T5, T6)**
F5. Towards broader sports media and fluid roles	**T7.** Different purposes of using visualization: analytical purposes, communicative purposes, and "eye-catcher"	**DI5.** Provide charts that sportswriters can use to communicate the insights and support their narratives **(T2, T7)**
	T8. Not realistic for them to create visualizations with coding and authoring tools. But they can interact with visualization interfaces and construct unique narratives	

Table 1: Three-level interpretation process. We concluded five primary findings based on the themes; then derived eight takeaways from the findings. Ultimately, we connected these takeaways and raised five generic design implications.

communicating them, another use of visualization emphasized by interviewees is its role in "catching eyes." P1 stressed that "*it is the first step into the door.*" P2 added, "*people naturally gravitate to charts, which is a lot.*" And P3 elaborated, saying "*people just want to scroll through something quickly. If it is a chart, they want to look at it... It is harder to get eyeballs on something just texts.*"

3.3 Implications for Design

Our findings characterized different aspects of the field: the needs, the challenges, work settings and existing tools, and the opportunities for visualization to contribute. By interpreting these findings, we derived eight takeaways **(T1-T8)** as shown in Table 1. By connecting the takeaways, we further distilled five high-level design implications **(DI1-DI5)** that can be applied to a broader scope.

The vast amount of sports data, the difference in their characteristics and usage contexts (T3), along with the diversity in sportswriters' backgrounds, data and visualization literacy (T8), and fluidity in their roles and purposes (T7), creates a "*wicked problem*" [6, 34], as "there is no definitive formulation of the problem" [45]. Visualization design study emerges as a suitable way to "tackle the wicked problem" [34]. To contain the problem within manageable bounds [11], we need to decompose the high-level fuzzy tasks (e.g., exploring data more efficiently) to lower-level crisp ones [47]. Reflecting on our study, we summarized and projected sportswriters' high-level activities into three dimensions: **Task** (e.g., post-game recap, season preview), **Topic** (e.g., offense/defense/shooting performance), and **Entity** (e.g., players/teams). There are many usage contexts sitting in between these three dimensions. In this study, we selected two usage contexts: **Post-game summary** and **Lineup performance evaluation**.

The post-game summary is one of the primary sports journalistic reporting types. Sportswriters lead their audiences to "rewatch" games through their lens by blending interview quotes, descriptive narratives, game statistics, and video footage. Post-game summaries often require the utmost immediacy and can be approached from various reporting angles, from analysis-driven to narrative storytelling. Sportswriters have been leveraging more advanced metrics and videos to support their narratives. Within this context, we further examined over 100 NBA post-game articles published on various media platforms, including ESPN, the Athletic, Sports Illustrated, and other local news platforms such as the Houston Chronicle, to gain a deeper understanding of post-game writing tasks. We identified a series of lower-level tasks that visualization can enhance: (1) identifying outstanding/critical stats for players/teams; (2) comparing performance between players/teams/time intervals; (3) locating critical moments/periods; (4) identifying critical events and corresponding game video; (5) inferring trends/patterns. To support these tasks, provide more context, and make it resilient to diverse needs, we append four contextualized design implications in addition to the generic design implications aforementioned:

- Allowing sportswriters to focus on a time window or a player
- Supporting global game statistics/metrics/charts updating
- Ubiquitous connections to bind related information
- Universal linking to corresponding game footages

Lineup performance evaluation in dynamic team sports is a complicated yet crucial matter. Most team sports allow different combinations of players to be used in a game. Frequently, such selections are among the most critical coaching adjustments, making the topic a frequent appearance in sports articles. Players' talents and skillsets may complement or conflict with each other, while teams' strategy and opponents' lineups can also add complexity. Sports analysts have developed various approaches using rich sports data to evaluate the efficiency and effectiveness of different lineups. These approaches, often built by designing statistical models or machine learning models, may fall into Anscombe's Quartet-like situations [16]. Some analytical models are tweaked to mitigate the effects of lacking contexts, such as excluding the data from garbage time and projected heaves [56]. Nevertheless, it remains exceptionally challenging to reveal the complex context using the traditional statistical methods. Furthermore, as the sizes of measuring variables and lineup cases grow, distilling insights from large data tables with only limited interactions can be onerous, and some valuable outliers could easily be overlooked. For this usage context, we built our understanding by examining the articles/posts that we collected during our ethnographic observation concerning this specific topic, team lineups. We then examined existing tools that provide related analytical data, such as the lineup stats pages on NBA.com and Cleaning the Glass. Based on our understandings, we concluded four lower-level design implications:

- Supporting multi-layer drilling-down exploratory analysis
- Focusing on revealing insights that are difficult to find with existing tools
- Utilizing and enhancing the rich advanced analytical stats to answer a broader range of analytical questions
- Supporting novel narrative construction

4 INTERACTIVE VISUALIZATION SYSTEMS

To develop our visualization systems, we followed the nine-stage design study methodology framework [47], with minor tweaks to leverage our background and suit the domain characteristics. This section introduces the resulting two visualization prototypes: *NBA GameViz* and *NBA LineupViz*, as efforts to address the two usage contexts we selected. For each system, we started by brainstorming and drawing up sketches based on the design implications. We then built initial prototypes with single views and limited interactivity. Following that, we entered our iterative design cycle that centered around self-deployment and included these six stages:

- **Simulate**: We simulated the writing process by observing NBA games and writing data-driven sports articles using the visualization systems. This stage helped us discover new narratives and usage scenarios within the context. The first author has multiple years of experience as an NBA journalist and held both *domain expert* and *visualization researcher* roles. The collaboration between the different people in these two roles is a fundamental and mandatory part of the nine-stage framework, but "*the same person can hold both roles in strong problem-driven work*" [47].
- **Post**: We leveraged social media by posting data-driven articles on various sports media and social media platforms. We used the articles as probes to obtain feedback from real-world audiences. Some feedback directly impacted our design choices, while others hinted toward future research opportunities in a broader social context, which we present in later sections.
- **Reflect and empathize**: We reflected on the writing process, the new narratives/usage scenarios, and the feedback from audiences and asked these questions: (1) Does visualization help? (2) Are the interactions effective in accomplishing the tasks? How can we improve usability? (3) Is the public data sufficient and retrievable? (4) Can we generate the needed data from existing data using algorithms?
- **Collect and validate the data**: If specific data were necessary for critical analytical narrative construction, we automated its collection and integrated it into our data pipeline. If the data was generated by algorithms using the existing data, we validated it through comparisons with the data on existing websites. This process helped us fix issues in our data pipeline.
- **Design and embed**: We designed the visualizations for the incoming data and embedded the visualization and interaction into the existing interface with focuses on the synergy of different visual components and interactions. For example, when we embedded the rotation chart into NBA GameViz, we changed the layout to align it with the existing game trend. We added two vertical lines corresponding to the selected time window as visual cues.
- **Implement**: We implemented the visualizations and fixed the issues that occurred during the interaction, then we repeated the self-deployment cycle.

4.1 NBA GameViz System

The NBA GameViz focuses on surfacing the game data and analytics about a particular single game. The information provided within a larger context aims to support rapid sense-making and narrative construction.

Data. The system is fed by three categories of datasets: play-by-play data (PBP data), historical contextual data, and metadata. We employed NBA_API, an API client package to retrieve raw datasets from NBA.com. The PBP data is the primary dataset for this system. It is aggregated from multiple raw data sources that include play-by-play descriptions (textual), team rotation data (temporal), timestamps (temporal), shot locations (spatial), shot types and results (categorical), event type (categorical), and video links (textual). We retrieved all 1165 games of the 2020-21 NBA season, including 1080 regular-season games and 85 playoff games. Each game's data includes 48 columns (variables) and approximately 500 rows (plays). The second category of data, contextual, includes advanced game stats and averages for all thirty teams from the regular season and playoff games. Finally, the metadata is a supplemental dataset incorporating game schedules, team information, and player information. To support our field deployment, we built a data pipeline to collect, clean, and aggregate data automatically. New game PBP data was retrieved and piped into the visualization system within 30-60 minutes after each game finished and other data was updated from there. We also embedded analytical algorithms to compute advanced metrics dynamically. The visualization interface is generated automatically according to the selected game.

Visualization and Interactions. The interface of NBA GameViz is shown in Figure 2. It includes multiple types of visualizations and subviews. The interface focuses around the familiar scoring summary (each game contains four 12-minute quarters) at the top center. Below that is the brushable **Game trend (A)**, an area chart that shows the game score trend (differential) from left-to-right, with team colors and height/depth indicating which team is in the lead and by how much. Below that are **Event chart (B)**: a visualization with color-coded marks/glyphs representing different event types (e.g., made/missed three-point shots, turnovers). Every mark can be selected, which shows a video clip of the corresponding play in the game. To the top left and right are the **Shot location charts (C1)** for the two teams, showing the spatial location of game shots (C1), color-coded by the result of the shot (red-miss, green-make). Next to the location charts are the aggregated **Shot type chart (C2)** showing the hierarchy of different shot types and

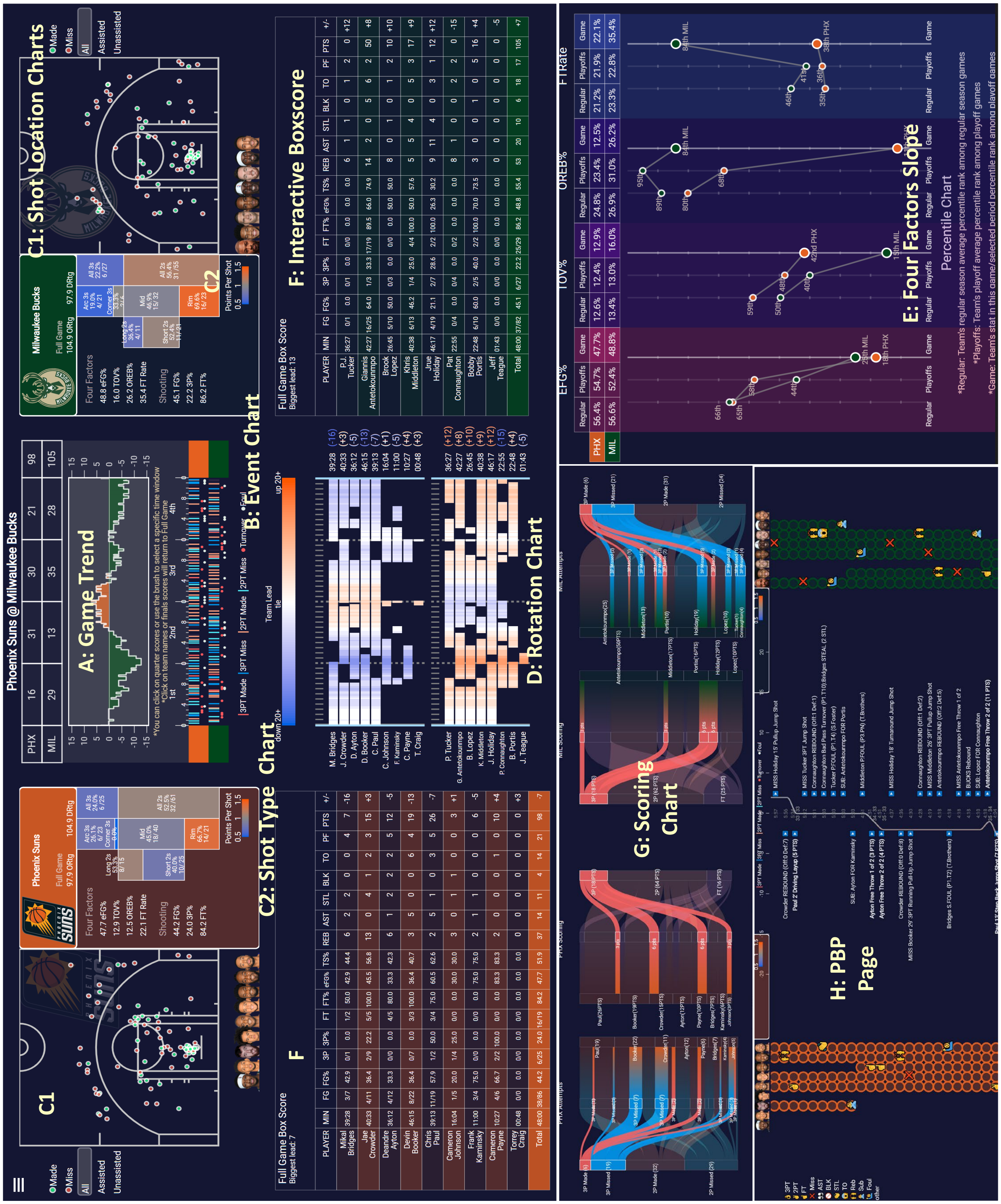

Figure 2: GameViz Interface. At the top is the main view. The Scoring Chart (G) and PBP Page (H) are the alternative bottom views.

Figure 3: Examples of using GameViz for player evaluation. The highly connected visual components could provide a quick overview of players' performance. Users can narrow focus by selecting different portions of the game and hovering on players and different views.

locations. The height of a segment indicates a number of that type of shot, and color indicates the make/miss efficiency. The lower center contains a **Rotation chart (D)**, a Gantt chart showing which five players were active for each team at different moments in the game. On the bottom right is the **Four Factors slope (E)**: that reveals the four factors percentile change from teams' regular-season average to playoff average to the current game.

The viewer can also select one of three different types of views to the left and right below the Shot Location Charts. The first option is an **Interactive boxscore (F)** listing many traditional statistics such as shots, rebounds, assists, turnovers, etc. The second option is the **Scoring chart (G)**, a customized Sankey diagram that presents the distribution and composition of both teams' points and attempts. Bar width represents the count/number, and color represents the categories. Finally, the viewer can select the **PBP page(H)** with lineups. This scrolling view places the players into columns and has game time move from the top (start) to the bottom (end). Each row is an action or event in the game, each of which can be selected to see a video clip of it. The view also contains a faint vertical gradient line chart to indicate the score difference at that point in time.

The interface is highly interactive and connected. The control panel on the sidebar allows viewers to choose a particular game and to navigate between the different lower views in the system. In order to support inquiry into different temporal segments of the game, a viewer can select (brush) any horizontal segment of the top-center Game trend (A), which then makes that temporal period of the game be the focus. When such a selection occurs, all the statistics and visuals are updated to only reflect what occurred during that specific range of time as illustrated in the supplemental video. The viewer also can select any of the four quarters on top of the game trend to quickly change the focus to that quarter. Clicking the final scores or the team names restores the selection to the complete game.

To help a journalist understand how specific players performed, a viewer can hover the cursor on a player's headshot picture, row in the boxscore (F), or node in the scoring chart (G). This action highlights all three items and filters the shot location chart (C1) and event chart (B) to only present actions by that player; any shots or events committed by other players fade, and the aggregated shot type chart is updated. Hovering the cursor on different attempt types or score types will also filter the event chart (B). Clicking the shot type buttons on the side of shot location charts enables the viewer to filter shots by miss/made or assisted/unassisted. Hovering on a shot-type icicle plot region also filters the shot location chart.

Hovering the cursor on a specific shot (small circle on floor map) or an event bar brings out a tooltip that contains the specific play description on-demand. Clicking on shot circles, events, or a play description on the PBP page (H) pops up a video window showing a clip of the play. Hovering the cursor over the cells within the boxscore (F) shows relevant data and filters the shot location (e.g., hovering on the *assist*s cell shows who the recipients are, how many points resulted from the assists, and where the shots were taken).

Usage Scenario and Insight Generation. To help explain how the system can assist sports journalists with story creation and authoring, we introduce a hypothetical usage scenario. Suppose that Leonard is an Atlanta Hawks columnist and focuses on in-depth coverage of the team. Tonight, he needs to write a single game analysis on Game 5 of the Hawks vs. 76ers playoff series. Shortly after the game finishes, Leonard opens up NBA GameViz and selects this game. He already knows it was a 26-point comeback win for Atlanta Hawks, but what jumps out on the game trend chart (A) immediately is that Hawks never had a lead until the final two minutes of the game. By looking at the aggregated shot charts (C2), he quickly sees that 76ers shot better than Hawks did, particularly from beyond the 3-point line. Both teams did not have many three-point shot attempts, and only five of them were from the corner area for both teams. This probably means that their defensive schemes were trying to force mid-range jumpshots. Leonard scrolls down to the Four Factors [37] slope (E) and immediately observes that ***the 76ers did better than the Hawks in shooting efficiency and free throw rate by a decent margin, but the Hawks were better at limiting their turnovers.*** Next, he clicks the buttons on the side of the shot locations chart (C1) and notices that most of the unassisted shots were from the mid-range area, so it seems that the defenses were largely successful in their schemes.

Leonard next decides to browse through all the players by hovering their headshot pictures. The players' stats are highlighted accordingly, and he can see the shooting details and the timing of the key events by each player. He identifies three top-scoring players for both teams: Young (39pts), Embiid (37pts), and Curry (36pts). Additionally, he sees more context and details as shown in Figure 3, which helps him to construct some narratives quickly: ***Young led scoring, but he did not shoot that well. A large portion of his points was from free throws. Seven of his ten field goals came in the second half. Embiid was dominating in the first half, hitting nine out of eleven, probably the biggest reason for the 76ers 26-point lead. But that was because he was making most of his mid-range jumpers(6/8), which is not very sustainable. He only hit 3/9 in the second half, and none of his shots were near the rim! Seth Curry was on fire from everywhere, shooting exceptionally well with a true shooting percentage of 87%!***

Leonard wants to dig deeper on the causes of the 76ers' early lead and the Hawks' comeback. He utilizes the time brush and other charts. Suddenly he has more narratives in mind: ***The Hawks started cutting into the lead late in the third quarter when their reserve lineup was on the court. As indicated in the rotation chart, the comeback was actually led by Williams, who had 13 points in these 14 minutes, and Gallinari, who hit some big shots. Gallinari also appears to have had a significant impact on crowding the paint with Capela and Collins to stop Embiid's close-in shots.*** Leonard explores further by clicking shots/events to see the corresponding game clips. He can also grab screenshots from the system and share the visual findings in his article to support his narratives.

4.2 NBA LineupViz System

The LineupViz system focuses on assisting assessments on performances of different combinations of players (lineups). It demonstrates how the lineups were used across multiple games and how the specific lineups performed.

Data. The datasets used in NBA LineupViz include all the PBP data introduced in NBA GameViz, aggregated by teams, as well as lineup performance and player performance data from NBA.com. The datasets are grouped by game types: regular-season games and playoff games. The lineup performance data consists of 38 numerical variables that describe various aspects for different numbers of lineups. Similarly, the player performance data contains 63 numerical variables to measure single-player performance. The data is updated after each game finishes.

Visualization and Interactions. The interface of the LineupViz is shown in Figure 4. The system focuses around the **Ranked Lineup Chart (A)** in the top center, an interactive text table with a search bar on top and a ranked bar chart on the right side. Each row presents a particular set of players being "on the floor" (in a game) together. The length of the bar to their right represents the minutes these players played together, and the color scale represents the net rating of how well they did (red is good, blue is bad). The left side contains the **Control Panel (D)** that allows a series of interactions, including selecting different teams, different sizes of player subsets (1-5), switching between regular season and playoffs, and filtering by playing time. To the right is the **Game Heat Map (B)**, an interactive heat map. Each row represents a single game, and the color scale shows the score difference at that point of the game. The row bars are separated horizontally into the quarters of a game. The **Lineup Performance Scatterplot (C)** at the bottom includes an interactive scatterplot that displays different statistics of the selected team's lineups, each of which is represented by a bubble. The size of the bubble represents the playing time of that subset of players. Via the control panel, the viewer also can select the two statistics to be shown on the scatterplot axes (C). The Game Heatmap (B) in Figure 4 shows games from the playoffs, but it can be swapped with a view of games from the regular season (E). Clicking on a single row of the heat map or hovering on a bubble in the scatterplot brings up a tooltip that contains more details, and that filters the game heat map, the primary function of the system, as shown in Figure 6.

Usage Scenario and Insight Generation. As done above for the GameViz interface, we include hypothetical usage scenarios to illustrate some of the kinds of insights the system can deliver. **Usage Scenario 1**: Our NBA journalist, Leonard, is reporting on the Atlanta Hawks NBA playoffs series with the New York Knicks. After the Hawks' Game 2 loss, Leonard noticed that the Hawks were not able to run their offense schemes when Young and Bogdanovic were both off the court. A narrative occurs to him that Hawks should stack these two guards' minutes and make sure one of them

Figure 4: Interface of the NBA LineupViz system. The control panel to the left supports selecting and filtering

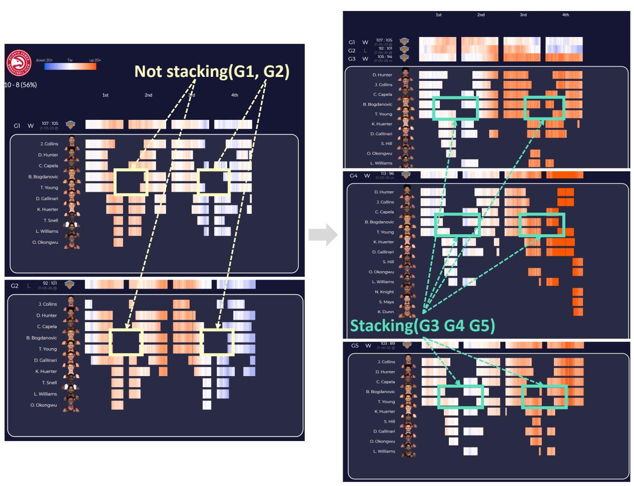

Figure 5: LineupViz Usage Scenario 1 - Left: The Hawks did not stack two players' minutes in the first two games, that is, both players left the game for the same time. Right: The Hawks stacked those two players' minutes in the rest of series.

is always on the court. To confirm and find evidence to support his narrative, he opens up the NBA LineupViz and selects Game 1 and Game 2 to see the specific rotations in these two games. Using the game heat map and the rotation chart, he confirms that Hawks' coach McMillian took both players off the court at the ends of the 1st and 3rd quarters and put them back about 3 minutes into the 2nd and 4th quarters. The changes of the colors also suggest that Hawks were losing the battle during these intervals. Having this in mind, Leonard wants to observe and see if this narrative still stands after Game 3 or if the Hawks made adjustments. Unsurprisingly, coach McMillian started to stack their minutes much more than the first two games (Figure 5). The charts in NBA LineupViz make it evident that the Hawks were able to either hold their lead or expand it during these minutes.

Ultimately, they won three games in a row after the adjustment to take the playoff series. Leonard was able to use these findings to interview Hawks' coaches about the strategy and write a subsequent analytical article explaining the Hawks' critical adjustments.

Usage Scenario 2: When the Hawks advanced to the playoffs second round against the Philadelphia 76ers, the top team in the Eastern Conference, Leonard wanted to find another angle to write about their adjustments. He observed that coach McMillian played the Hawks "Big Lineup" more often during this series, with Capela, Collins, and Gallinari on the court together. He examined LineupViz again, selected 3-man lineups in the control panel, then typed in these three players' names in the search bar. He found the relevant highlighted row in the upper-left table and selected it. All the time that these three played together jumped out immediately on the game heatmap, as shown in Figure 6. The bubble representing this lineup also was highlighted in the scatterplot. By selecting different statistical attributes, he was able to drill down and investigate why coach McMillian increased their usage and understand the move from different perspectives. Switching to the regular season

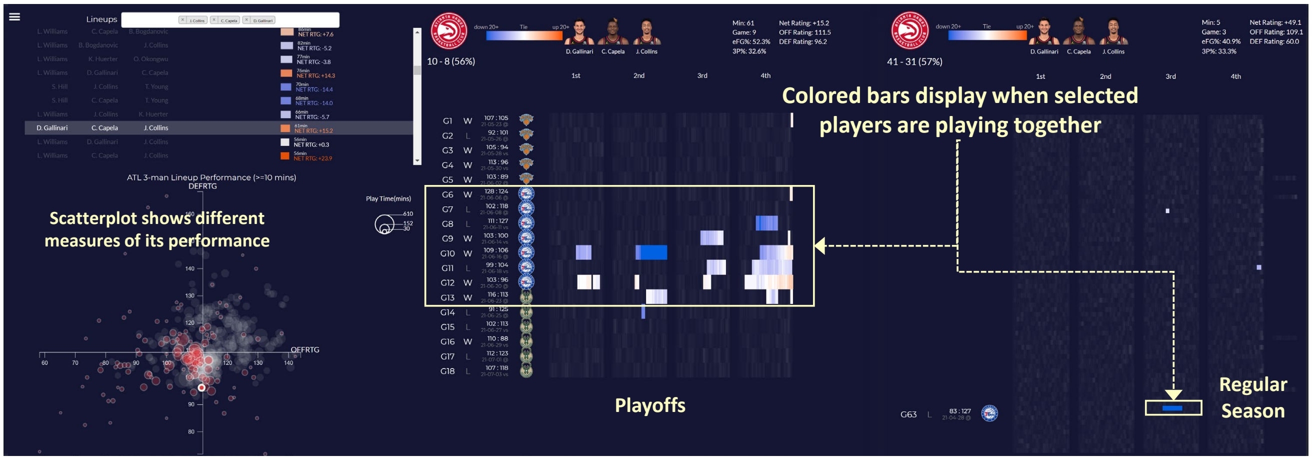

Figure 6: LineupViz Usage Scenario 2 - Left: Minutes when selected three players were on the court together in the playoffs. This lineup was mostly used against the 76ers. Right: The Hawks rarely played these three players together in the regular season.

view, Leonard also discovered that the Hawks rarely played these three together. They only logged 5 minutes together in 72 regular-season games, and it is against the 76ers as well. Combining the discoveries using NBA LineupViz and his observations, Leonard was able to construct insightful narratives: ***The Hawks did not use this lineup much in the regular season and the first round. In the second round, after trailing 1-2, coach McMillian started to increase the usage, and it became very effective on the defensive end because they could sag in the paint area and limit Joel Embiid, the 76ers' superstar center's dominance. The 76ers' offensive rating dropped to 96 when the Hawks were using this 3-man lineup.*** These narratives and insights would otherwise be difficult to discern and/or challenging to back up with evidence using existing tools. LineupViz supports a journalist to find such insights and provide visualizations to help communicate the data and support resulting narratives.

5 FIELD DEPLOYMENT AND EVALUATION

5.1 Deployment

To assess and validate our designs and prototypes, we primarily used a field deployment consisting of three parts: **Expert Feedback**, **self-deployment**, and an **In-the-Wild Deployment**. These parts correspond roughly to the **Semi-controlled Studies**, **Convenience Deployment**, and **In the Wild** components advocated by Siek, et al [50]. Ultimately, we deployed both systems online and allowed open access to the general public with tutorials and website links published on social media.

Expert Feedback. Five experts we interviewed earlier (J1, J2, P1, P2, A3) also participated in a user study of the two visualization systems. We first sent demos of and web links to the systems to them and asked them to use both visualization prototypes during the NBA Finals (two weeks) on their own. Afterward, we held remote user feedback sessions consisting of two parts: **Demonstration and think aloud**—we asked each participant to demonstrate how they would interact with the systems and think aloud while doing so; **Post-study interview**—we asked follow-up questions to further inquire about: (1) Their overall experience; (2) Whether the systems accelerated any tasks, and if so, what aspect(s) of the visualization/interactivity did so; (3) Whether the experts discovered novel insights/narratives; (4) Whether they felt that the charts would assist them in communicating insights/narratives to their audiences. Each feedback session lasted about 40-60 minutes and was recorded as a video for further analysis. We compensated participants with a $15 Amazon gift card for each session.

Self-deployment. During the 2020-2021 NBA season, we submitted over 300 data-driven articles/posts across four media platforms, and we received over a thousand of comments from sports fans. Each post/article contained at least one screenshot of a visualization generated by one of the systems. In the beginning, we rarely attached narratives or contexts to the posts. Later, we gradually added more descriptions, narratives, context, and explanations to the visuals.

In the Wild. We deployed the systems on the web and attached a survey, so anyone could use the systems and leave spontaneous feedback. The survey consisted of consent forms, validation questions, peripheral questions, and rating (0-100) questions. We adopted Insight and Time-savings as two evaluation metrics from the ICE-T value-driven evaluation framework [58], and added three further components. More specifically, we asked participants to evaluate the efficacy of the systems in terms of: (1) Insight: answering questions, helping to understand the game, discovering new insights. (2) Time: saving time in accomplishing tasks, understanding the game, and exploring the data. (3) Additional capability: allowing the user to accomplish tasks that are impossible/difficult to do using existing tools.(4) Communication: providing charts that help communicate insights/ findings. (5) Fun: making data exploration fun.

5.2 Feedback

Overall, all the expert participants spoke highly of both systems, and the systems were also well received by many domain experts after the release. Below is key feedback from our field deployment.

***Expert Evaluation*.**

NBA GameViz

Accelerates game information retrieval and analytical sense-making: All five experts agreed that GameViz could speed up sports data exploration and the sense-making process. P1 stated "*Those are stressful moments when you do not have much time to mess around and find stuff. A tool like this would put it right in front of your face so much quicker and allow you to access the memory of the game that you just watched.*" J2 added, "*This tool can help me remember what happened in the game immediately; otherwise, I have to look at the box scores for a long time and try to remember those plays.*" A3 said, "*This is something that I can see myself turning to and really quickly being able to generate many insights.*"

Supports deep-dive analysis: We observed that all five experts were able to conduct a more in-depth analysis of the games they selected. They were able to choose an interesting portion of the games and hover on different players to see details. P1 commented, "*It allows you to take one step further than just looking at box scores...You can dig into that next layer of the details and figure out what you want to focus on with your wrap-up. You can see where the game turned and what happened in that part of the game? Who was in the game? Who made the shots?*"

Time window selection is the most outstanding features: All the experts deemed GameViz's capability to allow time window selection and global stats/charts update as the most important interactive operation that separates it from existing tools. J1 said "*We journalists tend to focus on important runs and the turning points...Now I can focus on this specific period and see all the stuff that happened inside the window...This is extremely helpful.*" P1 added, "*Being able to pull out the stats for specific parts of the game using the window finder is a really neat feature...it really distinguishes it from other versions of a box score that you already see online.*"

Divergent opinions on the scoring chart: We intentionally designed the scoring chart as a probe to understand broader audiences' reactions to a more sophisticated chart (customized Sankey chart). Unsurprisingly, it received the most contrasting reviews. Four experts (J2, P1, P2, A3) expressed fondness for it. P2 stated "*Personally, I think it (scoring chart) is very pretty whenever I looked at them, and it gives me a much easier overview of where the points and shots are coming from.*". P1 added that the scoring chart is "*something he has never seen*", and it is "*really neat and innovative.*" But J1 found it less useful, citing that "*if interpreting takes more time than just describing it with words, why would I need a chart?*" and "*it might be too fancy for me to really get the message.*"

Recommended feature: hover to browse, click to focus: Three experts (J2, P1, A3) expressed that they would like to focus on a specific player, but A3 also mentioned "*it is convenient just to hover on players and their related charts updated immediately.*" Browsing allows journalists to quickly identify a noteworthy angle and then dive into it. A future revision could integrate the two actions to support "hovering to browse" and "click to focus."

NBA LineupViz

Reveals important obscure information: The most critical capability of LineupViz is that it allows users to select different combinations or subsets of players and display the periods when they played. All the experts agreed that this capability allows them to see patterns that could be very challenging to spot using other tools. J2 covered the Mavericks vs. Clippers playoff series. During the demonstration, J2 selected the team she covered and clicked on the first two 5-man lineups. J2 said "*It became obvious to me that Mavs changed their lineups in the last three games. It is probably because of Porzingis's bad performance, and they have to put Boban in.*"

Supports novel narrative construction: The rotation patterns uncovered by the Game Heatmap can be used in combination with the performance scatterplot to help journalists construct novel narratives. J2 commented "*Besides teams' starting lineup, I can easily see which other lineups the coach prefers, or when they were trying to catch up...I do not really dive into those before, but since this tool already puts it on paper, I can take a look and write about it.*" P1 said, "*It is potentially useful when you are writing a piece on a playoff series, 'why the coach needs to change the lineup?'... Load management was a big topic a couple of years ago. With a visual like this, you can trace the gaps in players' stints.*"

Provides more context: Lineup performance evaluation has been well studied with statistical models. LineupViz, as the experts suggested, provides more context to the statistical models. P1 shared a usage scenario where he used the system to refute a narrative raised by another journalist. "*It is cool to see how coaches actually used them, and it also gives more context...Some players may have amazing plus/minus numbers, but when you look at it, it is in garbage time...Some people are really good against opponents' second units. Being able to see that here is really useful to me.*"(P2).

Recommended feature: negation filtering Currently, LineupViz allows users to select any lineup or subsets of players on the floor together. Three experts (P1, P2, A3) suggested adding a filter to exclude specific players. "*I want to see the Warriors' performance when Curry and Draymond played together, but with Wiseman off the floor.*"(P2).

***Comments from Self-deployment*.**

Visualizations can foster credibility of narratives: The most frequent theme we distilled from the self-deployment comments was that visualizations lent credibility to narratives. It is quite common to see disagreements and contentious responses under narrative-only posts and articles, even when they are accompanied by statistical support. However, it was rare to see fans react with counterarguments under our visualization-accompanied posts. Instead, responders generally provided support, with feedback such as "*These are so professional, even though I don't understand the charts*" or substantiating the claims with analogous arguments.

Complex customized visualizations could be demanding to interpret, however...: Both GameViz and LineupViz provide unique custom views. Unsurprisingly, we received many comments on readability of some views, such as "*This is really aesthetically appealing, but it's hard for me to understand.*" or "*Are these charts for professionals? I couldn't understand.*" We then reposted such comments with short explanations attached. The follow-up comments were often positive. For example, one person replied that "*Now I understand... Compared to statistics, your visualizations are more impressive and informative. It became straightforward how players performed throughout the whole game.*". Another added that "*It is great if you can understand them, but there certainly is a bar.*"

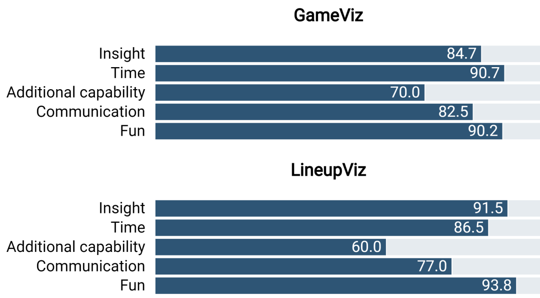

Figure 7: Average scores from in-the-wild survey. Each aspect was evaluated on a 0-100 scale

Feedback from the wild.
Although we received many comments and messages on social media, we received only nine valid spontaneous survey responses (R1-R9) in the two-week window. The quantitative results are shown in Figure 7. It is difficult to generalize from this small number of responses, but some qualitative comments were worth noting. R1 suggested a similar feature to "hover to browse, click to focus" as discussed in section 5.2. R2 recommended a "simpler version" for general fans, saying that "*The only downside of these systems is that they might be too complicated for amateurs or general fans. Maybe you can create a fan version by cutting some functions?*" This echoed a comment from a sports editor, stating "*The visuals might not be that attractive or easy to understand for casual fans...you may want to make insights more digestable for them.*"

6 REFLECTION AND DISCUSSION

Coalescing all of the feedback we received about the systems allows us to reflect on the strengths and weaknesses, as well as the challenges and experiences encountered during the design process.

6.1 On visualization systems

Leverage existing artifacts as anchors and couple them with context and interactivity. Components like boxscores, court diagrams, player images, and play-by-play descriptions have been in existence for a long time, so they are more familiar to journalists. We observed that users of the system tended to act on these artifacts to start their exploration (e.g., hovering on players' names or pictures). Moreover, attaching customized visualization and interactivity to these existing artifacts can provide more context and visual hints and make them more actionable. For example, we appended a vertical scoreline, customized on-court lineup, and activity icons to commonly-seen play descriptions on our play-by-play page. A3 commented "*A normal play-by-play is just a big chunk of information...but here I can scroll through and get a good sense of what is going on and which player it belongs to... It is better than any other play-by-play.*"

Make interactivity visible. We noticed that some interactions were overlooked by participants, especially for NBA GameViz due to its rich interactivity. One expert told us that he missed some interactivity that would otherwise have been very useful because he had not watched the tutorial video. All the experts missed at least one interactive capability (e.g., clicking the scores to fast select quarters or return to the full quarter state, hovering on players' boxscore cells to generate data facts that help narrative construction). P1 said "*Making sure that people know how everything works is big*". While we sent tutorial demos that showcase these interactions, we note that it is better to integrate on-site initial/real-time guidance [26] and more visual cues to help users get started with such multi-layered visualization interfaces, therefore flattening the "learning curve barrier" indicated by P2.

Cope with limited space. Space layout issues are common in interface design, particularly when multiple views are needed to showcase different datasets. With the rich connections that we implemented in these two systems, displaying connected views all together was a design priority. We carefully designed the layouts to mitigate the space issue while prioritizing "showing connected components together." However, some sub-views are still challenging to read and act upon, such as the game events bars. A2 noted, "*journalists usually write in the media room, so they have to use laptops with smaller screens. The font size can be challenging for them to read.*" A more dynamic layout mechanism that allows users to enlarge the focused part could be advantageous.

6.2 From Exploration to Communication

Towards Visual Data Storytelling. During our interviews, we asked practitioners "*If technical skill were not an issue, what type of visualization project would you work on?*" P1, an influential practitioner, told us that he would try "interactive data storytelling" had he known the technical aspects of doing that. Data-driven media such as FiveThirtyEight and the Pudding have attempted to bring data-driven storytelling to sports [17]. Narrative visualization techniques like scrollytelling have been applied to such visual data stories. While capable of delivering data-driven content in a more appealing fashion, these visual data sports stories are still somewhat scattered. Producing such visual data stories often involves "exploring data, making a story, and telling a story." and requires collaborative work among multiple roles, including data analyst, scripter, and presenter [29]. While major media platforms like The New York Times have adequate staff to make such endeavors, smaller media, independent journalists, or other content creators do not share comparable capability. Additionally, visualization developers/designers often take core roles in the storytelling process, whereas domain experts take lesser ownership because they lack technical skills. Commercial tools such as Tableau have lowered the barrier with products like Story Points. However, as we demonstrated with our systems, customized visualizations and interactions are capable of uncovering deeper insights.

Bridging the Gap. Even though we attempted to enhance our systems' capability to facilitate communication and we received positive feedback from experts on this front, the systems ultimately are still complex interactive dashboards. During self-deployment, we often relied on external annotation tools to highlight certain parts or add narratives to the screenshots. There were occasions when we needed to use visuals from both systems or compare different states within one system to support a narrative. While we could rearrange narratives and visuals, our ability to act upon these views was limited. This is not unexpected because exploratory visualization and explanatory visualization often aim at different targets and have different purposes [57]. To transition from exploring data to communicating insights, we advocate adding a layer on top of visual analytics systems. In this layer, journalists or analysts can act on the excerpts extracted from exploratory visualization systems and turn them into author-driven visual data stories. Chen et al. touched upon this notion with their "Story Synthesis" framework [9] and Sultanum et al. introduced a text-chart linking strategy with their prototype VizFlow [51].

6.3 On Field Deployment

Expert Feedback. While our expert participants provided us with rich feedback, one limitation was that their real-world usage of our systems was largely shielded from us. We asked them to use the systems "at their own will" due to our understanding of the high stakes during the NBA finals. We suggest future deployment studies take consideration of the real-world setting of study context, such as the NBA schedule in our case.

Self-deployment. It was beneficial for us to rapidly evaluate small design decisions, identify the pain points in situ, and draw inspiration for functionality design. By simulating the writing process and posting mock-up articles on media platforms, we provoked more input that helped us understand the problem through the lens of secondary stakeholders (e.g., journalists' audiences). Executing a self-deployment, however, can be challenging, mainly because visualization designers/developers do not always have sufficient **domain expertise** to perform "front-line" analysis, or they may lack the **time** to do so. Also, this approach is more applicable when the **application area** is public-facing, such as journalism, since other areas may involve confidentiality, or the audiences may be difficult to reach. On the flip side, self-deployment in a larger social setting might yield messy data. In our case, we posted a significant number of articles to generate spontaneous feedback without much active inquiry. However, we felt that more active and planned intervention (e.g., test reactions on annotated charts) could produce more informative data.

In-the-wild Deployment. As self-deployment gave us perspectives of the end audience, in-the-wild deployment provided a broader understanding of secondary users such as fans and coaches. For future in-the-wild deployment studies, we suggest a longer deployment window and strategies to curate more spontaneous feedback.

6.4 Broader Impacts

Our holistic approach and the use of sports writing as probes provided us with a broader vision of the value and potential of visualization in data-driven journalism. Below are some directions for future research.

Study the value of interest-driven data consumption in data literacy/education. Watching sports games has long been regarded as a form of entertainment by many. However, our study hinted that more sports consumers are being exposed to data and statistics with the evolution and prevalence of sports data and sports analytics. The "debate culture" in sports has been a double-edged sword: it can lead to cyber conflicts among fans and media, but "constantly trying to convince others"(J2) could encourage passionate fans to learn and argue with data-driven evidence. Many times have we seen visualizations presented as "evidence" by sports media and fans. We have also seen that ill-intended charts can misinform general audiences. Additionally, two of the visualization practitioners we interviewed have been leveraging sports visualization to teach programming, citing that "*data visualization makes it fun for beginners to approach data science*" (P3), and "*sports datasets are more interesting than those 'dull Titanic datasets' for people to start with.*" (P2). This echoes another finding of our study: the visual data articles/posts during our self-deployment received many comments and messages from general fans, inquiring "*What tools did you use to make these charts?*" and "*Can you share some good D3.js tutorials?*" Therefore, we wonder if ***sports data visualization could contribute to data literacy education***. With the enormous real-world data generated, the widespread presence of data analytics, and the vast audience base, we believe that interest-driven data visualization, like sports visualization, has the potential to contribute.

Study the social impact of sports visualization. From Wattenberg and Kriss's NameVoyager [59] to data practices during COVID [30], the use of visualizations in real-world information transmission contexts and their impact has attracted many visualization researchers. Relatedly, one theme caught our attention from the data we collected during self-deployment practice: narratives with visualization support usually received fewer comments. These comments tended to be more **friendly**, **benign**, and **rational**, however. Sports are replete with emotion, which sometimes translates to conflicts and cyber abuse. Many times abuse is directed at sports journalists. We acknowledge that some media may seek more engagement but also believe such topics are worth examining and can be expanded to similar areas, such as political journalism.

6.5 Implications for Other Sports

Although our visualizations were tailored for basketball and NBA journalists in particular, we believe that some of the design implications we derived can be transferred to other sports or even other areas.

Findings and Design Implications. We believe our findings and the resulting design implications listed in Table 1 are generally applicable to sports journalism. These findings and implications are driven by shifts in journalistic reporting (e.g., data-driven journalism), sports evolutions, and particularly sports analytics, technology, and social media.

Visual components. Every sport has its own unique characteristics, making it difficult to apply custom visualizations from our systems directly to other sports. However, we believe that there is still an inspiration to draw. For instance, other team sports could borrow the core feature of LineupViz to gain knowledge of lineup performance. Further, allowing users to flexibly select various time periods from a game for analysis also can be transferred to other sports.

7 CONCLUSION

This work explored an approach to assisting basketball writers in coping with rapidly growing sports data and analytics through interactive visualization interfaces. To inform our design, we collected holistic data via mixed methods to understand the opportunities and the obstacles for basketball writers. The data we gathered informed us of their needs, limitations, existing artifacts, and usage contexts. Based on our informed interpretation, we designed and constructed two interactive visualization applications **NBA GameViz** and **NBA LineupViz** to support basketball writers to conduct rapid, in-depth analysis and create novel narratives in two selected usage contexts. We deployed two robust prototypes in the wild and gathered feedback from target audiences and other potential stakeholders. Finally, we presented feedback and lessons learned from the design and evaluation process, as well as future work directions surfaced by our study.

REFERENCES

[1] Pivot Analysis. [n.d.]. Pivot Analysis | Basketball Analytics for Every League of Play. https://www.pivotanalysis.com/.

[2] BBallytics. 2020. Is the mid-range really dead? https://www.bballytics.co/en/2021/02/24/mid-range-evolution.html.

[3] Peter Beshai. [n.d.]. Buckets: NBA Shot Visualization. https://buckets.peterbeshai.com/app/#/playerView/201935_2015?playerSelector=true.

[4] Peter Beshai. [n.d.]. Scattershot. https://scattershot.peterbeshai.com/.

[5] Rena Kim Bivens. 2008. The Internet, mobile phones and blogging: How new media are transforming traditional journalism. *Journalism Practice* 2, 1 (2008), 113–129. https://doi.org/10.1080/17512780701768568

[6] Richard Buchanan. 1992. Wicked problems in design thinking. *Design issues* 8, 2 (1992), 5–21.

[7] Nancy Carter, Denise Bryant-Lukosius, Alba Dicenso, Jennifer Blythe, and Alan Neville. 2014. The Use of Triangulation in Qualitative Research. *Oncology Nursing Forum* 41 (09 2014), 545–547. https://doi.org/10.1188/14.ONF.545-547

[8] Dan Cervone, Alexander D'Amour, Luke Bornn, and Kirk Goldsberry. 2014. Pointwise: Predicting points and valuing decisions in real time with nba optical tracking data. In *Proceedings of the 8th MIT Sloan Sports Analytics Conference, Boston, MA, USA*, Vol. 28. 3.

[9] Siming Chen, Jie Li, Gennady Andrienko, Natalia Andrienko, Yun Wang, Phong H. Nguyen, and Cagatay Turkay. 2020. Supporting Story Synthesis: Bridging the Gap between Visual Analytics and Storytelling. *IEEE Transactions on Visualization and Computer Graphics* 26, 7 (2020), 2499–2516. https://doi.org/10.1109/TVCG.2018.2889054

[10] Wei Chen, Tianyi Lao, Jing Xia, Xinxin Huang, Biao Zhu, Wanqi Hu, and Huihua Guan. 2016. GameFlow: Narrative Visualization of NBA Basketball Games. *IEEE Transactions on Multimedia* 18, 11 (2016), 2247–2256. https://doi.org/10.1109/TMM.2016.2614221

[11] Nigel Cross. 1982. Designerly ways of knowing. *Design studies* 3, 4 (1982), 221–227.

[12] Nicholas Diakopoulos. 2012. Cultivating the landscape of innovation in computational journalism. *Tow-Knight Center for Entrepreneurial Journalism* (2012).

[13] Nicholas Diakopoulos. 2016. Computational journalism and the emergence of news platforms. In *The Routledge Companion to Digital Journalism Studies*. Routledge, 176–184.

[14] Nicholas A. Diakopoulos. 2011. A Functional Roadmap for Innovation in Computational Journalism.

[15] Ben Dowsett. 2018. Nylon Calculus: How Second Spectrum is redesigning the NBA. https://fansided.com/2018/06/28/second-spectrum-redesigning-nba/.

[16] Jean-Daniel Fekete, Jarke J. van Wijk, John T. Stasko, and Chris North. 2008. The Value of Information Visualization. In *Information Visualization - Human-Centered Issues and Perspectives*, Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North (Eds.). Lecture Notes in Computer Science, Vol. 4950. Springer, 1–18. https://doi.org/10.1007/978-3-540-70956-5_1

[17] Russell Goldenberg and Amber Thomas. 2019. How many high school stars make it in the NBA? https://pudding.cool/2019/03/hype/.

[18] Kirk Goldsberry. [n.d.]. (1) Most Common Shot Locations In The NBA This Season / Twitter. https://twitter.com/kirkgoldsberry/status/1374807806876868608.

[19] Kirk Goldsberry. 2012. Courtvision: New visual and spatial analytics for the nba. In *2012 MIT Sloan sports analytics conference*, Vol. 9. 12–15.

[20] Kirk Goldsberry. 2019. How Mapping Shots In The NBA Changed It Forever | FiveThirtyEight. https://fivethirtyeight.com/features/how-mapping-shots-in-the-nba-changed-it-forever/.

[21] Kirk Goldsberry. 2019. *Sprawlball: A visual tour of the new era of the NBA*. Mariner Books.

[22] Samuel Gratzl, Alexander Lex, Nils Gehlenborg, Hanspeter Pfister, and Marc Streit. 2013. LineUp: Visual Analysis of Multi-Attribute Rankings. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2277–2286. https://doi.org/10.1109/TVCG.2013.173

[23] Philip Hammond. 2017. From computer-assisted to data-driven: Journalism and Big Data. *Journalism* 18, 4 (2017), 408–424.

[24] Alexander Benjamin Howard. 2014. *The Art and Science of Data-Driven Journalism.*

[25] Bball Index. [n.d.]. Home - Basketball Index. https://www.bball-index.com/.

[26] Hyunmo Kang, Catherine Plaisant, and Ben Shneiderman. 2003. Helping Users Get Started with Visual Interfaces: Multi-Layered Interfaces, Integrated Initial Guidance and Video Demonstrations. In *Proceedings of the 2003 Annual National Conference on Digital Government Research* (Boston, MA, USA) *(dg.o '03)*. Digital Government Society of North America, 1.

[27] Michael Karlsson. 2011. The immediacy of online news, the visibility of journalistic processes and a restructuring of journalistic authority. *Journalism* 12 (04 2011), 279–295. https://doi.org/10.1177/1464884910388223

[28] Bongshin Lee, Eun Kyoung Choe, Petra Isenberg, Kim Marriott, and John Stasko. 2020. Reaching Broader Audiences With Data Visualization. *IEEE Computer Graphics and Applications* 40 (03 2020). https://doi.org/10.1109/MCG.2020.2968244

[29] Bongshin Lee, Nathalie Henry Riche, Petra Isenberg, and Sheelagh Carpendale. 2015. More Than Telling a Story: Transforming Data into Visually Shared Stories. *IEEE Computer Graphics and Applications* 35, 5 (2015), 84–90. https://doi.org/10.1109/MCG.2015.99

[30] Crystal Lee, Tanya Yang, Gabrielle D Inchoco, Graham M. Jones, and Arvind Satyanarayan. 2021. Viral Visualizations: How Coronavirus Skeptics Use Orthodox Data Practices to Promote Unorthodox Science Online. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 607, 18 pages. https://doi.org/10.1145/3411764.3445211

[31] Michael Lewis. 2004. *Moneyball: The art of winning an unfair game*. WW Norton & Company.

[32] Antonio G. Losada, Roberto Therón, and Alejandro Benito. 2016. BKViz: A Basketball Visual Analysis Tool. *IEEE Computer Graphics and Applications* 36, 6 (2016), 58–68. https://doi.org/10.1109/MCG.2016.124

[33] Ronald Metoyer, Qiyu Zhi, Bart Janczuk, and Walter Scheirer. 2018. Coupling Story to Visualization: Using Textual Analysis as a Bridge Between Data and Interpretation. In *23rd International Conference on Intelligent User Interfaces* (Tokyo, Japan) *(IUI '18)*. Association for Computing Machinery, New York, NY, USA, 503–507. https://doi.org/10.1145/3172944.3173007

[34] Miriah Meyer and Jason Dykes. 2020. Criteria for Rigor in Visualization Design Study. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2020), 87–97. https://doi.org/10.1109/TVCG.2019.2934539

[35] Tobias J. Moskowitz and L. Jon Wertheim. 2011. *Scorecasting: The Hidden Influences Behind How Sports Are Played and Games Are Won.* Crown Archetype.

[36] Tamara Munzner. 2009. A Nested Model for Visualization Design and Validation. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 921–928. https://doi.org/10.1109/TVCG.2009.111

[37] Dean Oliver. 2004. *Basketball on paper: rules and tools for performance analysis*. Potomac Books, Inc.

[38] John O'Sullivan. 2005. Delivering Ireland: Journalism's Search for a Role Online. *Gazette* 67 (02 2005), 45–68. https://doi.org/10.1177/0016549205049178

[39] John O'Sullivan and Ari Heinonen. 2008. Old Values, New Media. *Journalism Practice* 2 (10 2008), 357–371. https://doi.org/10.1080/17512780802281081

[40] Seth Partnow. 2021. *The Midrange Theory*. Triumph Books.

[41] Charles Perin, Romain Vuillemot, and Jean-Daniel Fekete. 2013. SoccerStories: A Kick-off for Visual Soccer Analysis. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2506–2515. https://doi.org/10.1109/TVCG.2013.192

[42] Charles Perin, Romain Vuillemot, Charles D. Stolper, John T. Stasko, Jo Wood, and M. Sheelagh T. Carpendale. 2018. State of the Art of Sports Data Visualization. *Computer Graphics Forum* 37 (2018). https://doi.org/10.1111/cgf.13447

[43] Owen Phillips. 2020. 25 Most Common Shot Locations by Teams. https://twitter.com/owenlhjphillips/status/1317111245741359106?s=20.

[44] Shot Quality. [n.d.]. ShotQuality | Homepage. https://shotquality.com/.

[45] Horst Rittel and Melvin Webber. 2018. Dilemmas in a General Theory of Planning. (02 2018), 52–63. https://doi.org/10.4324/9781351179522-6

[46] Dominik Sacha, Feeras Al-Masoudi, Manuel Stein, Tobias Schreck, Daniel Keim, Gennady Andrienko, and Halldor Janetzko. 2017. Dynamic Visual Abstraction of Soccer Movement. *Computer Graphics Forum* 36 (06 2017), 305–315. https://doi.org/10.1111/cgf.13189

[47] Michael Sedlmair, Miriah Meyer, and Tamara Munzner. 2012. Design Study Methodology: Reflections from the Trenches and the Stacks. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2431–2440. https://doi.org/10.1109/TVCG.2012.213

[48] Stephen M Shea. 2014. *Basketball analytics: Spatial tracking*. CreateSpace Independent Publishing Platform.

[49] Stephen M Shea and Christopher E Baker. 2013. *Basketball analytics: Objective and efficient strategies for understanding how teams win*. Advanced Metrics.

[50] Katie A. Siek, Gillian R. Hayes, Mark W. Newman, and John C. Tang. 2014. Field Deployments: Knowing from Using in Context. In *Ways of Knowing in HCI*, Judith S. Olson and Wendy A. Kellogg (Eds.). Springer New York, New York, NY, 119–142. https://doi.org/10.1007/978-1-4939-0378-8_6

[51] Nicole Sultanum, Fanny Chevalier, Zoya Bylinskii, and Zhicheng Liu. 2021. Leveraging Text-Chart Links to Support Authoring of Data-Driven Articles with VizFlow. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 16, 17 pages. https://doi.org/10.1145/3411764.3445354

[52] Ben Taylor. 2016. *Thinking Basketball*. CreateSpace Independent Publishing Platform. https://books.google.com/books?id=GgWgDAEACAAJ

[53] Ben Taylor. 2019. Is the midrange shot dead? The surprising math behind basketball's least efficient attempt. https://www.youtube.com/watch?v=lZZry6Aed3k.

[54] Ben Taylor. 2020. The unique skills that made Larry Bird a GOAT candidate | Greatest Peaks Ep. 4 - YouTube. https://www.youtube.com/watch?v=S8n5tPbWB50&t=8s.

[55] Cleaning the Glass. [n.d.]. Cleaning the Glass – Toward a Clearer View of Basketball Decisions. https://cleaningtheglass.com/.

[56] Cleaning the Glass. [n.d.]. Guide - Garbage Time. https://cleaningtheglass.com/stats/guide/garbage_time.

[57] Alice Thudt, Jagoda Walny, Theresia Gschwandtner, Jason Dykes, and John Stasko. 2018. Exploration and explanation in data-driven storytelling. In *Data-Driven Storytelling*, Nathalie Riche, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale (Eds.). AK Peters/CRC Press, 59–83. https://doi.org/10.1201/9781315281575

[58] Emily Wall, Meeshu Agnihotri, Laura Matzen, Kristin Divis, Michael Haass, Alex Endert, and John Stasko. 2019. A Heuristic Approach to Value-Driven Evaluation of Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 491–500. https://doi.org/10.1109/TVCG.2018.2865146

[59] M. Wattenberg and J. Kriss. 2006. Designing for social data analysis. *IEEE Transactions on Visualization and Computer Graphics* 12, 4 (2006), 549–557. https://doi.org/10.1109/TVCG.2006.65

[60] Todd Whitehead. 2017. Hot hand. https://crumpledpaperjumper.com/portfolio/shot-distances-after-misses.

[61] Todd Whitehead. 2019. Nylon Calculus: Charting player passing lanes. https://fansided.com/2019/02/17/nylon-calculus-charting-player-passing-lanes/.

[62] Mark Wilson. 2012. Moneyball 2.0: How Missile Tracking Cameras Are Remaking The NBA. https://www.fastcompany.com/1670059/moneyball-20-how-missile-tracking-cameras-are-remaking-the-nba.

[63] Wayne L. Winston. 2012. *Mathletics: How Gamblers, Managers, and Sports Enthusiasts Use Mathematics in Baseball, Basketball, and Football.* Princeton University Press.

[64] Jana Wiske and Thomas Horky. 2021. Digital and data-driven sports journalism. *Insights on Reporting Sports in the Digital Age: Ethical and Practical Considerations in a Changing Media Landscape* (2021), 2. https://doi.org/10.4324/9781003010944-4

[65] Qiyu Zhi, Suwen Lin, Poorna Talkad Sukumar, and Ronald Metoyer. 2019. GameViews: Understanding and Supporting Data-Driven Sports Storytelling. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300499