This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3326910

HoopInSight: Analyzing and Comparing Basketball Shooting Performance Through Visualization

Yu Fu and John Stasko

Abstract— Data visualization has the power to revolutionize sports. For example, the rise of shot maps has changed basketball strategy by visually illustrating where "good/bad" shots are taken from. As a result, professional basketball teams today take shots from very different positions on the court than they did 20 years ago. Although the shot map has transformed many facets of the game, there is still much room for improvement to support richer and more complex analytical tasks. More specifically, we believe that the lack of sufficient interactivity to support various analytical queries and the inability to visually compare differences across situations are significant limitations of current shot maps. To address these limitations and showcase new possibilities, we designed and developed HoopInSight, an interactive visualization system that centers around a novel spatial comparison visual technique, enhancing the capabilities of shot maps in basketball analytics. This article presents the system, with a focus on our proposed visual technique and its accompanying interactions, all designed to promote comparison of two different scenarios. Furthermore, we provide reflections on and a discussion of relevant issues, including considerations for designing spatial comparison techniques, the scalability and transferability of this approach, and the benefits and pitfalls of designing as domain experts.

Index Terms—sports data visualization, sports analytics, visual comparison, basketball

1 Introduction

With the increasing prevalence of sports data collection and analysis, sports visualization has become a powerful tool for accessing, analyzing, and communicating sports data and analytics [11, 38]. Only a handful of existing sports visualizations can be considered as having fundamentally altered the trajectory of a sport, however. In the realm of basketball, shot maps (or shot charts) [18, 19] are such an example – the locations where teams take shots have changed dramatically, driven in part by the insights available through shot maps [43]. Basketball teams have modified their game plans, aiming to create more opportunities for players to take shots from their "hot spots." Such shifts in professional leagues (e.g., the NBA - National Basketball Association in the U.S.) have rippled out to millions of fans and players worldwide, affecting their appreciation of and engagement in the sport of basketball.

Although the use of data, visualization, and analytics in basketball has grown significantly over the past decade, one form of analysis that has not seen an accompanying focus is in-depth *comparison* of data and information. Comparison is one of the most important analytical activities in sports [11]. For example, how does one player compare to another, or how has a particular player's performance improved or fallen over time? Media outlets rely on comparisons to assemble a wide range of compelling stories. Fans engage in casual or fierce arguments with friends and strangers about which player is better on different occasions. Analysts and coaches make comparisons to evaluate player and team performance, analyze various circumstances, and make appropriate adjustments in team make-up and strategy.

Our particular focus in this work is the comparison of shots taken in basketball. When a player "shoots" the ball toward the basket hoping that it goes through the hoop, this is called a "shot." Obviously, whether the shot is made or missed is the most important metric, but many other accompanying variables exist for each shot. The second most important variable is likely the position on the basketball court where the shot was taken from. Additional variables include the type of shot (dunk, jumper, hook, etc.), the time of game/possession it was taken, and which other players were in the lineup at that time, among others.

• *Yu Fu and John Stasko are with Georgia Institute of Technology. E-mail: fuyu@gatech.edu, john.stasko@cc.gatech.edu*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

In general, this basketball scenario can be abstracted as a spatio-temporal data analysis problem. The basketball court is our spatial grounding, which we divide up into a series of contiguous smaller regions for aggregation and convenience. Each region then has multiple events (shots) occur within it, each event having multiple quantitative and nominal variables. The primary variables that we focus on are ***frequency*** and ***efficiency***. Frequency denotes how often shots are taken from a certain region and can be measured as the number of shot attempts or as a proportion of total attempts. Efficiency denotes how accurate/effective a player/team is shooting the ball, i.e., the proportion of shots taken that are made. Our goal is to support our primary user group, sports analysts or other analytic-savvy stakeholders (e.g., fans, journalists), examining the shot behaviors of two references, such as two different players or one player in two different time periods, and to compare the performance and patterns of the two selected references, hopefully gaining insights from the comparison.

We have developed an interactive visualization technique and an accompanying system called HoopInSight to compare and analyze shooting patterns and performance in a way existing shot maps fail to provide. Our approach revolves around a novel visual technique with flexible interactions and a real-world data pipeline. More specifically, the data pipeline and interactivity enable users to swiftly construct a spectrum of comparative conditions/scenarios, e.g., Denver Nuggets' shooting behaviors with or without their star player Nikola Jokic in the game. Meanwhile, our proposed visual technique allows users to rapidly compare multivariate spatio-temporal data across two chosen scenarios and derive insights, such as determining the locations where shots were taken more/less frequently and how much more/less efficient those shots were, and ultimately infer the contrasting patterns.

2 Related Work

2.1 Sports Visualization

Perin et al. [38] conducted a comprehensive state-of-the-art review and identified two primary roles of visualization in sports: analytical and narrative. Visualizations used to analyze sports data focus on understanding the ever-increasing amount of data and extracting insights from it. These insights can lead to novel analytical storylines, help scouts understand opponents' tendencies and behaviors, and even assist in strategy-making or team assembling. Narrative uses of sports visualization tend to be more prevalent, targeting a broader audience with more straightforward messages. Du and Yuan [11] focused on the use of visual analytics in competitive sports and proposed a more nuanced categorization of competitive sports data. They also listed *feature* *comparison* as one of the three main tasks for competitive sports data visualizations. Our work falls into that category and seeks to advance the analytical use of visualization by incorporating multiple data types revolving around spatio-temporal information about shots.

While many sports visualization practitioners have primarily focused on using visualization for communication purposes, researchers also have sought to address challenging analytical problems leveraging both novel interactions and visual techniques (e.g., glyphs) [1, 10, 25, 37, 40, 44, 49]. In the domain of basketball, projects such as GameFlow [7] and BKViz [29] aim to demonstrate complex statistics about NBA games, such as game trends, play sequences, and team rotations. Another focus is the emerging notion of supporting sports journalists to cope with increasingly advanced basketball data and analytics, aiming to assist them in exploring such analytics and identifying novel insights to build narratives [13, 52]. Other researchers have endeavored to leverage tracking data and its derivatives to further advance the use of visualization in sports analytics [42, 48]. Recently, visualization researchers have sought to integrate sports visualization with natural language and VR/AR technology [8, 9, 27, 28], thus broadening the roles of sports visualization and benefiting more stakeholders. Reflecting on their projects and collaborations with different user groups, Lin et al. provide valuable insights into working with different stakeholder experts and future research opportunities [26].

2.2 Visualization for Basketball Spatial Analytics

Our work aims to push the boundaries of basketball spatial analytics through interactive visualization. The quest to understand basketball shooting behaviors through spatial analysis dates back to the 1940s. Research on this topic [21, 39], however, was scattered, until Goldsberry's CourtVision technique paper [18] popularized the technology in both academia and industry, with his signature shot map paving the way for basketball spatial analytics. Ensuing research has emphasized developing and improving models [6, 22, 23, 32] to either reduce dimension or provide estimations/evaluations.

Although CourtVision was originally intended for spatial visual analytics, its primary use has been in sports journalism, partly due to Goldsberry's long tenure as a journalist. To communicate sophisticated insights to broader audiences, he recast his invention into a set of variants, ranging from classic hexagonal bins to more recent contour lines to aesthetically appealing shot charts featuring 3D animation and embellished visual effects. Other practitioners have likewise explored different forms to engage audiences and deliver compelling stories with shot charts. For instance, Whitehead created a physical representation of Stephen Curry's career three-pointers using toy trays [33, 45].

Despite the success in attracting audiences and driving narratives (e.g., "the midrange is dead" [3, 20, 50]), the analytic value of shot charts has seemingly been sacrificed in favor of storytelling and engagement. During an invited talk at the 2018 IEEE VIS conference, Goldsberry enumerated several reasons why NBA analytics departments have been slow to adopt visualization. Yet we have observed a trend that NBA teams and sports analysts have recently realized the importance of data visualization, although their employment of visualization has primarily focused on a mere showcase of their advanced metrics, with only a few exceptions (e.g., Buckets [4]) that employ interactivity for analysis. Based on our experience in the sports industry and visualization research, we argue that the insufficient realization of Goldsberry's original vision is largely attributable to a lack of organic integration of analytics, visual representation, and interactivity. This motivated us to design and develop HoopInSight.

2.3 Visual Comparison for Information Visualization

Gleicher et al. [17] surveyed visualizations designed to support comparative tasks and proposed a general taxonomy of visual comparison. They classified visual comparison designs into three general categories: *juxtaposition*, which displays different objects separately; *superposition*, which overlays objects within the same space; and explicit encoding of the differences/relationships [16, 17]. Based on his framework [16], the targets we try to compare are the shooting performances of two entities (i.e., players/teams). The challenge lies in the complexity of individual items (multivariate spatial data) and their relationships. The strategy we adopt falls under the *Summarize Somehow* category — we compute the difference of each variable for each spatial unit. Our main visual technique for spatial comparison falls under the third category — explicit encoding of the relationships between two objects (i.e., shot maps). However, to combat decontextualization, we also include two reference objects in the system, making our system a combination of explicit encoding and juxtaposition.

3 Understanding the Problem

3.1 Design As Domain Experts

As characterized by Sedlmair et al. [41], design studies in visualization research are projects that analyze "*a specific real-world problem faced by domain experts*" and "*support solving this problem*" with a visualization system. Collaboration between visualization researchers and domain experts is considered a *"fundamental and mandatory part"* of the framework. However, there are occasions when visualization designers/researchers are also well-versed in the domain for which they are designing visualization solutions. Our case fell into this territory where "the same person holds both roles" (visualization expert and domain expert), which Seldmair et al. opted not to address in their design methodology article [41]. We characterize this distinctive design study method as *Designing as Domain Experts* (DaDE). In this project, after careful consideration, we ultimately decided to embrace and capitalize on DaDE for the following reasons:

(i) The first author is a domain expert in basketball journalism/analytics and a visualization researcher, and he has a comprehensive understanding of domain tasks and real-world data. The second author also has considerable experience in visualization research and is an enthusiastic and knowledgeable sports fan.
(ii) The data we use are available through public APIs, thus, no necessity to collaborate with "data providers" exists.
(iii) Potential stakeholders come from a variety of backgrounds, ranging from professional team management to sports analysts/journalists, as well as avid or casual fans who may or may not possess domain/data/visual literacy.

3.2 Basketball Shot Data and Shot Maps

The fundamental unit of data we are exploring is a shot taken by a player. Each shot is either made or missed, and we can aggregate these shots temporally (e.g., minute, game, season) and spatially (coordinates on the court). Within a particular spatial region and temporal window, two primary metrics will drive the analysis: frequency and efficiency. Our objective is to visually compare these and other metrics between two different situations or scenarios. For example, we may want to compare a player's performance on shots from the "elbow" area [15] in the previous season to those same type of shots taken during the current season. More specifically, suppose he took 65 shots from that area in 2022 and made 27 (42%), but in 2023 he took 73 shots and made 28 (38%). The two important comparative metrics that we seek to communicate are that he took 8 more shots in the current year but his efficiency declined by 4%. Furthermore, we seek to do this for any players, teams, regions of the court, and time periods.

To understand current approaches to presenting this type of data, we collected and analyzed current shot map techniques and examples. Our collection comes from various sources, including research literature, renowned practitioners, and sports data websites.

Multiple methods for shot mapping exist, which differ in levels of aggregation and granularity, visual encoding (e.g., color scale, shapes), and filtering. Here we categorize these methods and explain how each aligns with a specific style of matching conventional geographic maps.

Shot (Dot distribution maps): The *Shot* technique does not aggregate data; rather, it shows the actual location where each shot was taken, represented by a dot or other small glyph. Different colors or glyph shapes differentiate the binary results, indicating whether a shot was made or missed. A significant limitation of *Shot* is overplotting when a large number of shots is present. Therefore, *Shot* is often employed to show a small number of shots (e.g., shots taken in a single game). The

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3326910

basic *Shot* technique is often straightforward and easy to understand for general audiences.

Cell (Proportional symbol maps): The *Cell* technique divides the court into small, predefined cells. It then aggregates the shots by cell and calculates different metrics on those aggregations. *Cell* typically uses some symbol (e.g., circle, rectangle, hexagon) to represent cell metrics, frequently employing symbol size to represent the shot frequency and color (usually diverging scale) to represent efficiency.

Zone (Choropleth maps): The *Zone* technique is commonly seen on data websites. It divides the court into larger, predefined zones, usually of varying sizes and shapes. While these divisions may vary across different websites, they typically follow domain conventions. *Zone* usually presents the aggregated efficiency for shots contained within each zone via the fill color of the region, but it does not represent the frequency. Consequently, *Zone* is easier for the general public to understand but holds less analytical value than *Cell*.

Others: Other less common shot map categories exist. For example, ***Contour line*** uses contour lines to depict the frequency of shots as a mountain landscape and ***Heatmap*** employs kernel density estimation to achieve the same, but with color intensity. Both *Contour line* and *Heatmap* usually present shot frequency but omit shot efficiency. They often leverage existing contouring and/or kernel density algorithms and can be aesthetically appealing. Such encodings are most useful in surfacing the most common places where shots were taken but may represent the data less precisely.

3.3 Task Analysis and Design Objectives

Since these different forms of shot charts have already become well-established industry standards, our goal is not to invent new chart variants but to focus on enhancing current capabilities and filling in the gaps of missing capabilities. In particular, we aim to: (1) enable more effective spatial comparisons, (2) support multiple analytical and comparative perspectives, and (3) communicate comparison within the context of original data.

Enable more effective spatial comparisons

Enable multivariate spatial comparison: The most notable analytical value of current shot maps lies in their ability to display two key metrics (*Frequency* and *Efficiency*) within a spatial context. We must similarly support exploration of these two key variables with respect to different spatial locations (i.e., positions on the court) for two different scenarios. We must support individual analysis of each scenario, as well as a comparison of differences and changes between the two scenarios. Ideally, viewers should be able to rapidly and clearly determine trends and patterns of the two variables across spatial locations.

Explicitly encode the differences: As reviewed earlier, Gleicher et al. [16, 17] categorize visual designs for comparison into three types: juxtaposition, superposition, and explicit encoding. People currently compare shot charts and similar sports data primarily by juxtaposing them as small multiples (e.g., [1,4]) beside each other and visually scanning back and forth. While this method is scalable, it can be cognitively challenging to compare multiple 2-D shot charts. The combination of many regions to examine and multiple variables to understand makes repetitive visual scanning to ascertain the differences quite difficult. We believe that explicitly encoding shot differences between two scenarios, that is, providing additional view(s) that only represent the delta/difference, can furnish deeper and more direct analytical insights, which are difficult to uncover by simply scanning between individual shot maps.

The above goals lead us to the first design requirement:

DR1 — Design explicit visual encodings that facilitate pre-attentive multivariate spatial comparison

This can be broken down into four sub-requirements. First, the locations from which shots are taken are often associated with players' skill sets and teams' game strategies. Therefore, we prioritize displaying areas where the number of shots has increased or decreased across scenarios (**DR1a**), as well as the specific increased/decreased quantities at each location (**DR1b**). Furthermore, the system should assist in detecting spatial patterns and trends in the data, which often signify a shift in team strategy or playing style (**DR1c**). Finally, although changes in efficiency may be inconsistent across neighboring locations, they can provide insights into a player/team's shot-making ability, i.e., whether or not they improve their shooting performance from certain areas. Thus, the system also must visualize values and patterns of quantitative efficiency (**DR1d**).

Support multiple analytical and comparative perspectives

Current shot maps mainly provide two types of insights: where shots are taken and how efficient a player/team is at shooting. Together, they can disclose information swuch as whether a player/team took more shots in highly efficient areas. Comparison, however, can further lead to a wider range of analytical insights based on different references and scenarios. For example, by comparing a player before and after he was traded to another team, one can analyze the change in his roles. Comparing two players directly can yield insights into their different play styles. To support multiple analytical and comparative perspectives, the system must enable users to create different comparison scenarios, which leads to the second design requirement:

DR2 — Enable users to create different comparison scenarios

We can model the different types of comparisons via a comparison cross-table ($entity \times scenario$). In this context, entity refers to players, teams, and opponents, while scenario involves lineups, shot types, time windows, and other variables. This can be simply described as *"comparing **entities** under different **scenarios**."* For instance, one could compare Michael Jordan's (entity) shooting performance with and without Scottie Pippen on the court (lineup scenarios), or compare Jordan's (entity) shooting pattern in his early career vs. later years (temporal scenarios). This design requirement can be divided into two sub-requirements: 1) obtain necessary data and structure the data to allow such filtering activities (**DR2a**); 2) design interactions to afford different entity selection and comparison scenario creation (**DR2b**).

Communicate comparison within the context of original data

An earlier goal emphasized the desire to explicitly encode comparison data. Only presenting an explicit comparison encoding (i.e., without the two source data views) can lead to decontextualization [17], however. Ideally, we seek to display comparison information *as well as* the original data from the two scenarios being compared. This consideration leads to the third design requirement:

DR3 — Present the two reference data sources (shot maps being compared) concurrently with comparison views to provide context and serve as devices for direct manipulation

We divide this requirement into three sub-requirements. First, the system should combine explicit encoding with juxtaposition and interactive connections to guide users' attention and simplify analysis (**DR3a**). Second, the system should incorporate more visualizations to provide further context and additional metrics (**DR3b**). The insights extracted from these supplementary views can stand alone, but more importantly, they should assist users in identifying interesting subsets of the data or provide extensive insights. Finally, we aim to allow users to reconfigure comparison scenarios by directly manipulating supplementary views (**DR3c**).

4 HoopInSight

Our basketball shot visualization system is called HoopInSight and is driven by multiple raw datasets that we retrieved from the NBA API [36]. We merged and transformed the collected data into shot datasets grouped by players, teams, and seasons. Each data item (an individual shot event) contains 31 attributes, including the court coordinates (spatial), the result (numeric), the type (categorical) of each shot, team rotation data (textual), a video link connecting to the specific shot, and metadata. The data is updated on a daily basis. We also fetched supplemental data, such as player/team information. This data pipeline (retrieval, aggregation, and partition) helps us fulfill **DR2a**.

Within the system, users can select to examine a particular NBA player's or team's data and which season to display. We also support a special third focus entity (Opponents) that aggregates all opponents' shooting data from games played against the selected team. This is a novel addition that enables users to analyze a team's defensive

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3326910

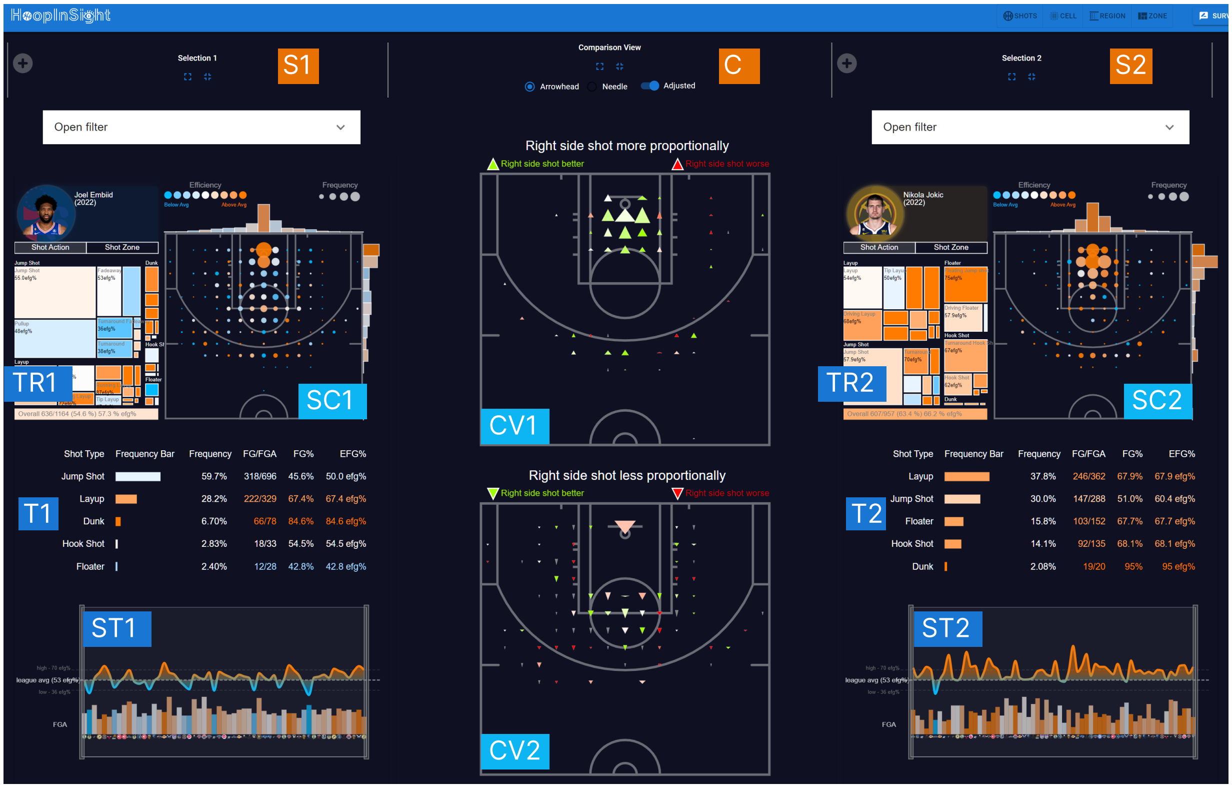


Fig. 1: HoopInSight user interface: It consists of three large columns — two *Selection Views* (S1&S2) and the *Comparison View* (C). Each *Selection View* has a *Shot Chart* (SC1&SC2) and multiple supplementary views (T, TR, ST). The *Comparison View* is divided into two sub-views (CV1&CV2).

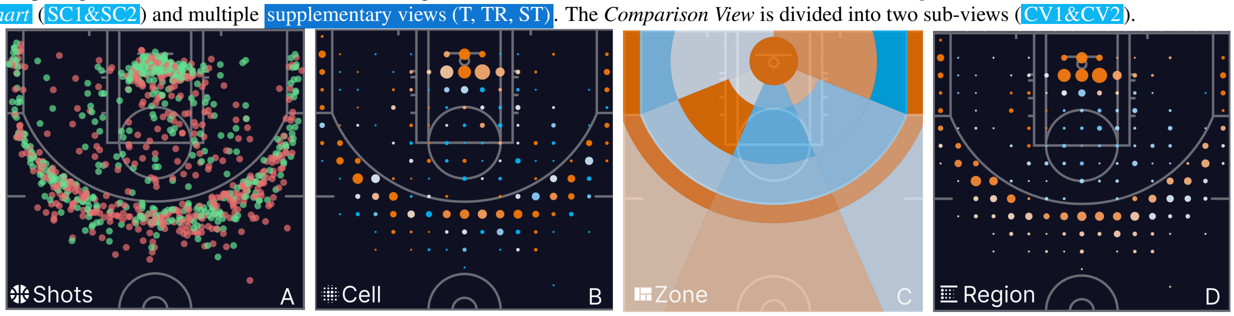


Fig. 2: The four visual encodings for a shot map (Stephen Curry 2022-23 season): (A) Shots, (B) Cell, (C) Zone, (D) Region

capabilities from the "shots allowed" perspective. Together, these capabilities partially fulfill **DR2b** (i.e., entity selection).

HoopInSight normalizes shot **frequency** by dividing the number of shots taken in a specific area by the total number of shots taken during the specified time period. We adopt the statistic, *eFG%* - Effective Field Goal Percentage [2, 34] to measure **efficiency**. eFG% adjusts for the fact that a 3-point field goal (shot made) is worth 50% more than a 2-point field goal. For example, if a player makes 4 out of 10 of his 3-point attempts, the traditional FG% (field goal percentage) is 40% while eFG% is 60%.

4.1 Interface and Visualizations

The user interface of HoopInSight (Figure 1) consists of three columns of views. Each column contains multiple visualizations. The two columns on the left and right of the interface are called a *Selection View* (Figure 1 S1&S2). Each *Selection View* centers around a relatively standard *Shot Chart* (Figure 1 SC1&SC2) that represents all the original shots taken by the selected entity under the specified scenarios. For example, *Shot Chart 1* may show all of a player's shots from the previous season, while *Shot Chart 2* shows all the shots from this season. In between the *Selection Views* (Figure 1 C), in the center, is the *Comparison View* that employs two stacked sub-views to represent a comparison of the two *Shot Charts*, i.e., the difference/delta.

Shot Chart (Figure 1 SC1&SC2): Fundamentally, this chart displays all shots taken by the selection, using one of four potential visual encodings that we implemented following the shot map typology categorized earlier in Section 3.2. The first, called the ***Shots*** encoding (Figure 2A), is a map of half the basketball court with a small point at each spot where a shot was taken. The color of the point indicates whether the shot was made (green) or missed (red). For the second encoding, ***Cell*** (Figure 2B), we divide the basketball half-court into 272 contiguous square cells (17 columns and 16 rows), each about 3ft (0.91m) wide and tall. We then aggregate all the shots taken within each cell and calculate the frequency and efficiency for that cell. We use circles to represent each cell, positioned at its center, with the size of the circle indicating the corresponding shot frequency and its color indicating the shot efficiency for that cell via a diverging color scale from orange (high efficiency) to blue (low efficiency). The third encoding, called ***Zone*** (Figure 2C), divides the half-court into meaningful zones known to basketball experts (e.g., left corner three, above the break 3, etc.). These zones are much larger than an individual cell of the ***Cell*** encoding and provide a chart roughly resembling a choropleth map. Again, all shots within a zone are aggregated. In this view, however, no circles are drawn. Instead, the fill color of the cell encodes the efficiency of shots

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3326910

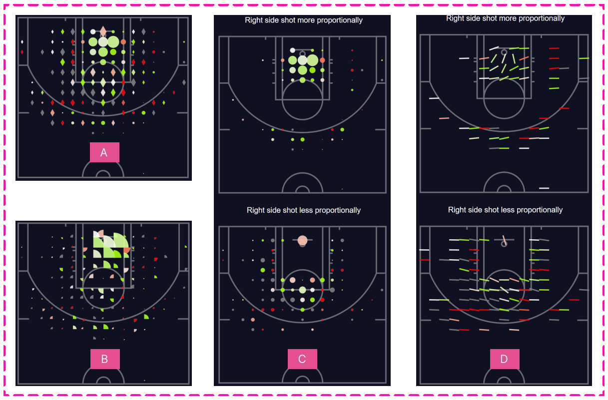


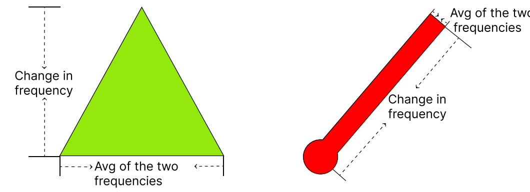
Fig. 3: Alternative comparison encodings we explored but did not choose. Designs A and B employ one consolidated view with different types of glyphs. Designs C and D employ two sub-views with the top showing cells of frequency increase and the lower cells of frequency decrease.

via the same diverging color scale as in the ***Cell*** encoding. The ***Zone*** encoding does not represent frequency. Finally, the fourth encoding, ***Region*** (Figure 2D), is a hybrid of the ***Cell*** and ***Zone*** encoding. It contains the same cell circles as the ***Cell*** encoding, with circle size indicating frequency. Each circle's color, however, is assigned based on the color of the zone its cell falls within. In other words, we use the zones as spatial filters to create smoother circle colors.

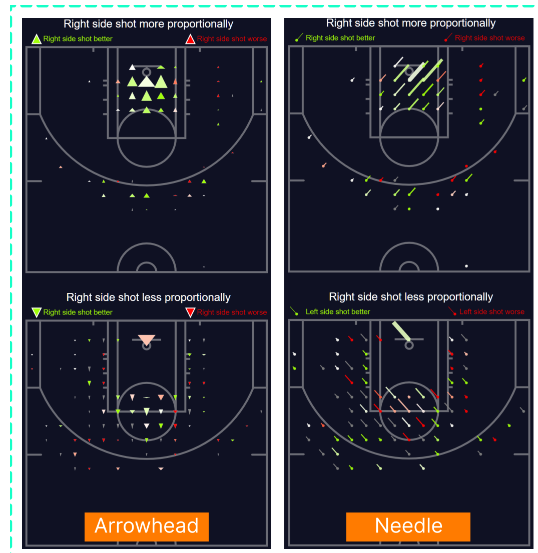
These four visual encodings provide insights at different levels of granularity: the ***Shots*** encoding allows users to see specific shots and their textual descriptions, as well as access a corresponding video clip. The ***Cell*** encoding offers visual representations of aggregated statistics in terms of efficiency and frequency. It also provides the basis for us to compare two selected targets. The ***Zone*** encoding allows viewers to focus only on efficiency — a simpler analytical task that fans tend to be more familiar with. The ***Region*** encoding provides a more balanced approach between ***Cell*** and ***Zone*** by still displaying frequency but with a smoother color encoding, as the efficiency tends to be more volatile across the cells. HoopInSight allows any of these four encodings to be used for the left and right *Selection Views*.

Comparison View (Figure 1C): The central *Comparison View* provides a visual comparison of the two displayed *Shot Charts* with a primary goal of communicating the differences (Δ) between the two selections, thus fulfilling design goals (**DR1**). As we sought to display the changes in both frequency and efficiency, we chose to compare the cells from the ***Cell*** shot encoding, and thus the *Comparison View* uses the same cell units. For each corresponding comparison cell, we use the data for S1 ($D1$) and the data for S2 ($D2$) to calculate the Δ (data for the comparison cell). The resulting Δ has three main attributes: its index on the court $\Delta_{(i,j)}$, the difference in frequency Δ_F, and the difference in efficiency Δ_E. $\Delta_{(i,j)}$ translates to the $[x,y]$ position on the court, and the Δ_E and Δ_F are two quantitative variables we encode for each cell.

Our first challenge was to design a visual encoding that would be placed at each cell to represent the difference between the two corresponding original shot aggregations for that position. We experimented with using glyphs of one or more graphical shapes, and we conducted multiple rounds of "brainstorming" feedback sessions with our local visualization research group. During the first round, the authors introduced the motivations and analytical tasks and presented initial designs to elicit feedback. Other researchers provided suggestions on visual encodings or low-fidelity sketches. Multiple candidate visual encodings emerged, and we selected six and implemented each through a high-fidelity visualization prototype. We presented these candidates (Figure 3 & Figure 4b) again to our local group and gathered second-round feedback. One design (Figure 3A) used one consolidated view with two different shapes: circles denoted cells where frequency increased from Shot Chart 1 to Shot Chart 2, and diamonds denoted cells where the frequency decreased. The size of the glyph denoted the amount of frequency change and its color represented the change



(a) Mappings of data attributes to glyph features. Color represents change in efficiency in both.



(b) The two chosen designs for representing comparisons in HoopInSight.

Fig. 4: Visual encodings used in the *Comparison View*

in efficiency between the two scenarios. A second design placed a four-quadrant axis on each cell, with a sector's appearance indicating one of the four results (increasing/decreasing $\times$ Δ_E/Δ_F) and the radius indicating the quantity. With both of these designs, viewers felt that each glyph was comprehensible, but it was relatively difficult to discern trends and patterns across the entire view of glyphs showing all cells (**DR1c**).

A third design (Figure 3C) employed two sub-views, one (upper) with glyphs for cells with increasing frequency between the scenarios, and the other (lower) for cells with decreasing frequency. The design used circles for the glyphs with size and color encoding deltas in frequency and efficiency, respectively. Viewers felt that this design was too similar to the circles in the *Shot Charts*, however, which caused confusion. Another design (Figure 3D) substituted "slopes", i.e., tilted lines, for the circles with angles of the line representing frequency differences. Although this design helped to show group patterns (**DR1c**), viewers felt it was too difficult to understand and decode the angles representing frequency differences **DR1b**. We also experimented with animations to denote the difference between two corresponding cells but found the animations unsuitable for analytical reasoning.

Ultimately, feedback helped us to narrow down to two promising visual encodings, which we now include in HoopInSight and allow users to select. We call the two designs that emerged *Arrowhead* and *Needle* (Figure 4). In each, a single glyph (a triangle or a sloped line, respectively) is drawn at each cell to encode the delta between the two selections. The first variable Δ_F, the change in frequency, is encoded in the *Arrowhead* via the height of each triangle, with taller triangles indicating larger changes in frequency. The width of each triangle's base represents the average of the two frequencies, thus showing wider triangles in regions where the player tends to take more shots. For example, in one cell, the number of shots taken in the two *Shot Charts* may have increased from 70 to 80, while in another, the shots increased from 10 to 20. Both of the corresponding triangles will have the same height because the delta between the two selections is +10 in both

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3326910

cases. However, the triangle in the first case will be wider and larger overall, helping to indicate that this is a region where the player shoots frequently. The change in efficiency between the two regions, Δ_E, is denoted by the color of the triangle. We employ a diverging color scale, with a greener fill color indicating increased efficiency from selection 1 to 2 and a redder one signifying decreased efficiency.

In the *Needle* glyph technique, the length of the sloped line denotes the change in shot frequency at the location and the thickness of the line denotes the average of the two frequencies, with thicker lines indicating regions where the player takes more shots. The color of each line encodes the difference in efficiency between the two cells.

Both designs also employ the idea of two sub-views, as was done in the earlier alternative designs shown in Figure 3C&D above. The top sub-view indicates all cells where the shot frequency increased from the first to the second selection, while the bottom sub-view indicates cells where the shot frequency decreased. Additionally, triangles in the upper *Arrowhead* sub-view (frequency increase from Selection View 1 to 2) are drawn pointing upwards, while those in the lower sub-view (frequency decrease) point downwards. Similarly, lines in the upper *Needle* sub-view slope to the upper-right, while those in the lower sub-view slope to the upper-left. We feel that this separated view makes the frequency and efficiency differences between the two periods more pre-attentively clear, enabling rapid determination of areas of shot increases (glyph locations on the court in the upper view) versus shot decreases (glyph locations in the bottom view), fulfilling one of our primary design goals (**DR1a**).

Supplementary Views: The two *Selection Views* also include multiple other visualizations that help to encode more variables about the shots in the two selections and serve as instruments for users to interact with and perform informed filtering, thus fulfilling **DR3b&c**.

On the top and right edges of the two *Shot Charts* are histograms showing aggregated statistics of the relevant row or column of regions on the half-court. The height of each bar represents the frequency, and color encodes the efficiency for the respective row or column.

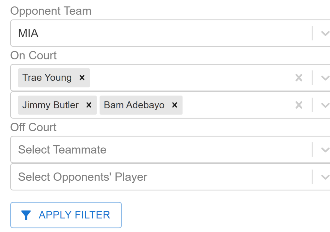
To the left of each *Shot Chart* is a zoomable treemap, *Shot Tree* (Figure 1 TR1&TR2), that displays the frequency and efficiency statistics for various shot types (such as pull-up jump shots, floaters, and layups) and different shot zones (such as the left corner three-pointer, midrange, and restricted area). Below each *Shot Chart* is a traditional statistics table (Figure 1 T1&T2) that presents a higher-level statistical summary of the treemap. Below the table is *Season Trend Chart* (Figure 1 ST1&ST2) displaying aggregated game-based statistics over the course of the selected season with time moving from left to right. The upper line chart shows efficiency throughout the season, while the lower bar chart shows the number of shot attempts in each game.

4.2 Interaction

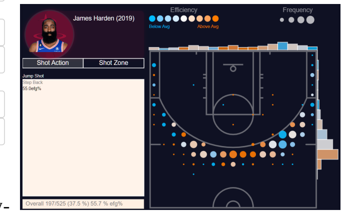
HoopInSight supports rich interactions that enable users to explore and create different comparison scenarios. Below we list some of the different styles of interaction [51] provided by the system.

Filter: The most prominent interaction in HoopInSight is filtering. Above the *Shot Chart* views are filter panels (Figure 5a) that enable users to filter shots (and thus make a selection) based on the opponent team, team lineups, and opponent lineups. For instance, users can apply filters to generate visualizations that represent one team's shot map with or without a particular player, an opponent's shooting patterns with or without a particular defensive player on the floor, or an individual player's shooting pattern when he plays against a particular team. The treemaps to the left of each *Shot Chart* also serve as categorical filters (Figure 5b) that support selection by shot actions and shot zones. For example, users can view the shot map of a particular player's "step-back" jump shots. The *Season Trend* charts below the *Shot Charts* also serve as a temporal filter (Figure 5c) using brushing, allowing users to select any portion of a season that interests them.

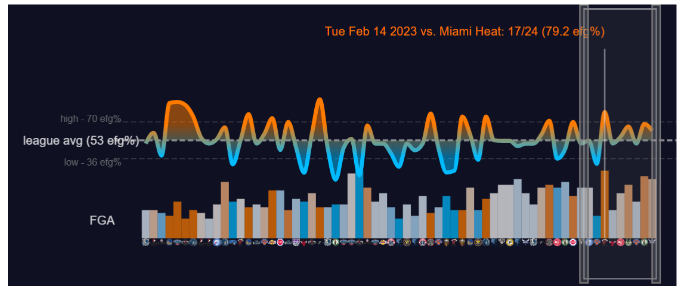
Another important filtering operation provided by the system is ***freehand/lasso selection*** (Figure 8a). When using the ***Shots*** encoding in a *Shot Chart*, users can freehand draw a polygon to select an area of interest. HoopInSight aggregates the shots within the polygon and calculates corresponding statistics. The background of the selected area



(a) The *Filter Panel* for teams and players



(b) The shot *Type filter*



(c) The *Temporal Filter* for selecting games during a season

Fig. 5: Primary filtering interactions supported by HoopInSight. (a) filtering by the opponent team and lineups for both teams. (b) filtering by different shot types (action type & zone type); the example here selects only Harden's step-back jump shots. (c) filtering by a portion of a season; the example here selects only the games late in the season.

transitions into a color reflecting the shot efficiency within it. Hovering over the selected area brings out an on-demand view that includes a shot action treemap, a sunburst chart for an assist breakdown of the selected shots (i.e., which players assisted the made shots taken from the selected area), and a statistical summary of shot actions. When the focus is on team shooting performance, this view additionally displays the top 5 players who took the most shots in the selected area, along with their statistical summaries. The filtering interactions fulfill the design goals to facilitate scenario creation (**DR2&DG3c**).

Select & Elaborate: HoopInSight generally provides details on demand through hovering. For instance, hovering over each cell in the *Comparison View* displays a tooltip that shows corresponding statistics regarding frequency and efficiency.

Connect: HoopInSight provides a connect style interaction to link related cells and zones. Hovering over cells/zones in one view will simultaneously highlight the identical areas in other views (**DR3a**).

Encode: HoopInSight allows users to toggle between the four visual encodings for the *Shot Charts* and the two visual encodings for the *Comparison Views*. Also, hovering over legends in the *Comparison Views* can highlight the cells where efficiency increased/decreased.

In summary, the visual technique enables users to "see the differences" between two scenarios, whereas the role of interactivity lies in expediting the construction of various analytical scenarios and accessing contextual information and corresponding details on demand.

5 Case Studies

This section presents two case studies to illustrate the analytic capabilities of HoopInSight.

5.1 Analyzing Nikola Jokic's Offensive Impact

Nikola Jokic is the star player of the Denver Nuggets and has been named the NBA's Most Valuable Player (MVP) twice. Recently, basketball analysts/journalists and thousands of fans have been engaged in a heated debate about whether he should win a third consecutive MVP award. While most model-derived advanced metrics point towards yet another historic performance from him, many traditional analysts/journalists have expressed skepticism toward such advanced metrics that often lack explainability and interpretability.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3326910

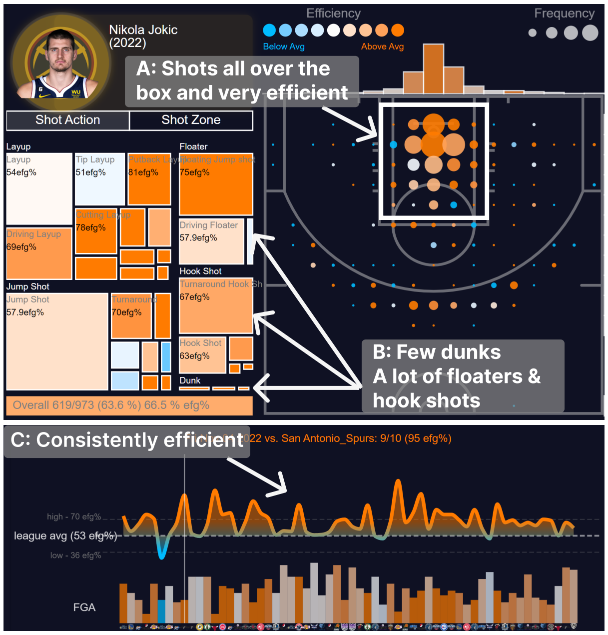


Fig. 6: Perspective I: Jokic's individual shooting performance. **Note**: We add external annotations using white text and arrows on the figures in this section to aid explanation. These annotations are not part of the system.

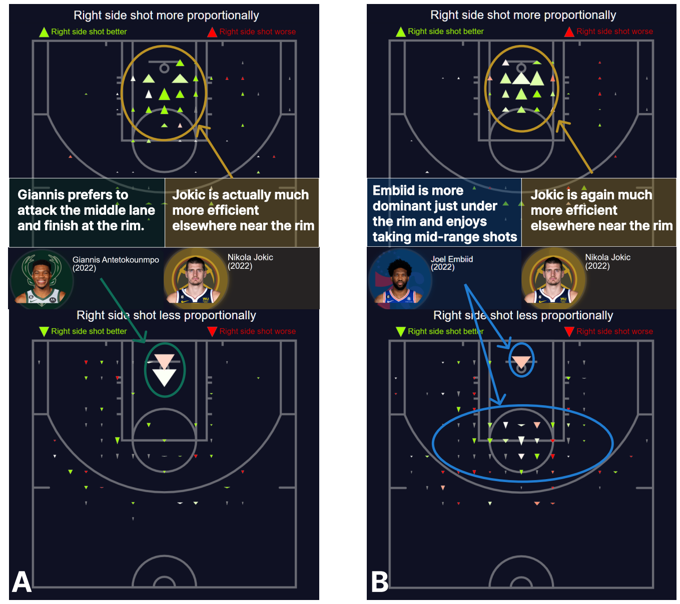


Fig. 7: Perspective II: Compare Jokic with counterparts. A - Jokic (S2 on right) compared to Antetokounmpo (S1 on left). B - Jokic (S2) compared to Embiid (S1)

Perspective I: Individual shooting performance — Jokic is extremely efficient and consistent; he is also versatile on offense

To understand Jokic's impact on offense, HoopInSight first supports analyzing his individual shooting performance. After selecting Jokic's shooting profile from the current season, the *Selection View* displays the pertinent metrics (Figure 6). Illustrated by all the orange in the views, it becomes apparent that Jokic is an extremely efficient scorer. The *Shot Map* (Figure 6A) provides insights into Jokic's most frequent shots — they are from all over the box, with a moderate number of three-pointers from above the break. The *Treemap* (Figure 6B) reveals that as a center, Jokic rarely ends the offense with a dunk but instead takes many floaters and old-fashioned hook shots, implying that Jokic relies heavily on his excellent touch rather than athleticism. The *Season Trend Chart* (Figure 6C) shows that Jokic is consistently efficient throughout the season, with the exception of a few games.

Perspective II: Comparing Jokic with his counterparts — Different styles and efficiency

Another angle to analyze Jokic's individual shooting performance is to directly compare him to his counterparts in the current MVP race — Giannis Antetokounmpo and Joel Embiid. Again, the *Selection Views* allow us to examine their overall efficiency throughout the season and their individual shooting performances. But it is the *Comparison View* that makes their distinctive shooting patterns extremely clear – Antetokounmpo is more dominant in the restricted area, where he shoots more frequently just under the rim and is even more efficient than Jokic, who shoots more frequently elsewhere and is significantly better away from the basket (Figure 7A). Jokic is also a better three-pointer shooter, especially from above the break. Compared to Embiid (Figure 7B), Jokic is not as dominant under the basket but is almost more dominant anywhere else in the box. While Embiid enjoys taking mid-range shots near the free-throw line, he surprisingly is not as accurate as Jokic.

Perspective III: Comparing on/off team shooting pattern — Jokic's team is taking more efficient shots when he is on the court

Jokic has been lauded as one of the best "team players" for his ability to create opportunities for his teammates. It is no secret that Jokic is an excellent passer — his average assists number speaks to that. HoopInSight supports an in-depth analysis of how exactly Jokic influences his team's offense. By applying the lasso filters to the Denver Nuggets' team shooting profile, the on-demand views (Figure 8a) illustrate that Jokic has the highest assist rate in almost every area. Furthermore, HoopInSight allows users to directly compare the Nuggets' shooting pattern with and without Jokic on the court. The green upward triangles, pointed out in Figure 8b, reveal that the Nuggets are taking more high-efficiency shots in the restricted area ($\Delta_F : 29.7\% \nearrow 36.1\%$) with higher efficiency ($\Delta_E : 65\% \nearrow 72\%$) with Jokic on the court.

Perspective IV: Comparing teammate's shooting pattern: how Jokic has transformed Aaron Gordon's game

HoopInSight can also be used to analyze Jokic's impact on specific teammates. Consider Aaron Gordon, a player who was traded to Denver two years ago and has thrived this season. The light green triangles (Figure 8c left) immediately reveal that Gordon has taken more shots in the restricted area with easy layups and dunks this year than he did two years ago with the Orlando Magic, indicating he has much better opportunities to score. One reason for this might be that Gordon has simply improved. To explore that notion, one can compare Gordon's same-season performance and filter the data based on whether Jokic is on the court. The large green upward triangles (Figure 8c right) quickly reveal that Gordon shoots more near the rim while playing alongside Jokic — the frequency of his dunks has doubled, largely due to Jokic's ability to create opportunities for him. This double-confirms Jokic's positive impact on Aaron Gordon's offensive metamorphosis.

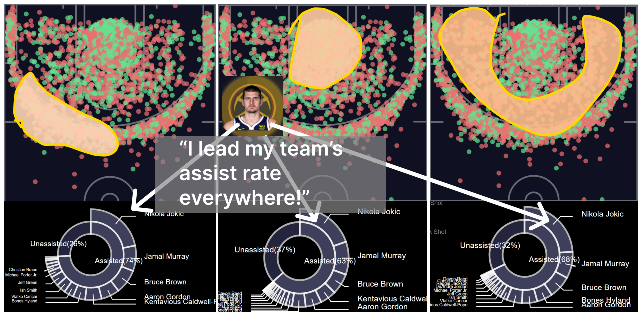
5.2 Compare Big Men's Defensive Influence

One area of analysis that existing shot maps rarely support is defense. Measuring the defensive impact of an individual player is notoriously difficult. Although advanced statistics and analytics can provide helpful aggregated information, they rarely allow for in-depth analysis, let alone insights into the intricate defensive schemes professional teams employ.

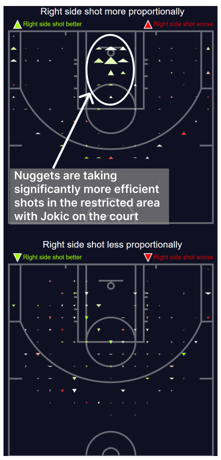
Although HoopInSight does not provide a direct answer, it can offer insights from the "shots allowed" perspective and enable analysts and coaches to test their hypotheses or ask insightful questions. One usage scenario is to compare opponents' shooting performance/patterns when a player is on the court versus when he is off the court. In this case (Figure 9), we compare the differences between teams' opponents when a big man (center position) is on or off the court. Here we select four different scenarios.

Figure 9A represents the difference in shots taken against the Utah Jazz during the 2020-2021 season with and without their Defensive Player of the Year award winner, Rudy Gobert, on the court. In this example, Selection 1 represents when Gobert was off the court and Selection 2 represents when he was on the court. Figure 9B-D follow the same procedure, but the focus teams are the Hawks, the Bucks, and the Mavericks from the current season, and the focus players are Clint

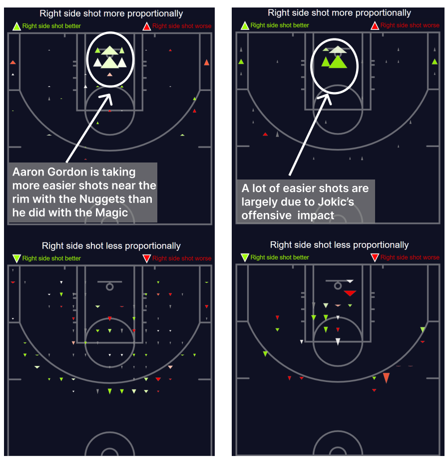
This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3326910



(a) Using the lasso selection to explore passing ability. The Sunburst chart reveals that Jokic assisted the most shots in each different area.



(b) Nuggets' performance with Jokic (S2 - right) vs. performance without Jokic (S1 - left)



(c) Left: Gordon's performance on the Nuggets (S2) compared to that on the Magic (S1). Right: Gordon with Jokic on the floor (S2) compared to Gordon without Jokic on the floor (S1)

Fig. 8: Perspective III & Perspective IV: Further exploration of Jokic's impact

Capela, Brook Lopez, and Christian Wood, respectively. The first three players are well-known rim protectors, while the fourth player, Wood, is arguably the opposite. The top four sub-views show where their opponents took more shots, while the bottom ones illustrate where their opponents took fewer shots. HoopInSight visually provides insights into how their presence changed opponents' shooting behaviors.

Insight I: When the first three players were on the court, their opponents took more mid-range shots and fewer and also less-efficient shots near the rim, as evidenced by the red, downward triangles. However, when Wood was on the court, opponents were seemingly encouraged to attack the rim. Additionally, when Capela was present, opponents took fewer shots from the two corners, suggesting that limiting penetration was effective. Conversely, when Wood was present, the Mavericks' opponents took more 3-pointers and fewer mid-range shots.

Insight II: While the first three players' presence led to more opponents' attempts in the midrange area, HoopInSight could reveal more nuanced insights. For example, the efficiency of the midrange shots Utah Jazz gave up decreased (red triangles), suggesting those shots were "forced" by Utah's defense with Gobert anchoring it. The green triangles in B, indicating that the Hawks' opponents had better opportunities in the mid-range area, raise questions about their other players' abilities to get through screens and contest the mid-range jump shots after their opponents' pick & roll actions [24].

6 Reflection and Discussion

6.1 Design Considerations For Spatial Comparison

Consideration I: Different comparison granularities

Drawing from Gleicher's insight that *"comparison is more than just finding differences"* [16], we argue that in our case, **spatial comparison**

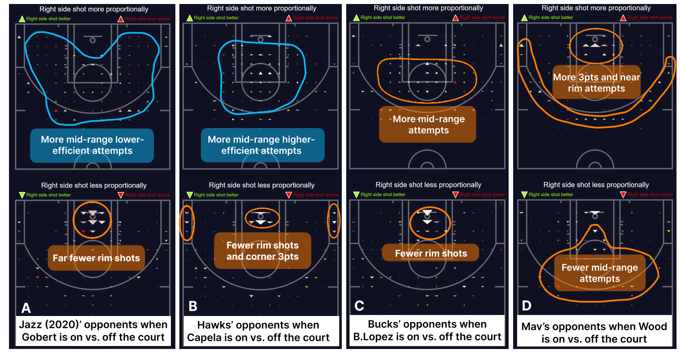


Fig. 9: How big men change their opponents' shooting behaviors. All four cases compare opponents' shooting behaviors with selected players on the court (S2 - right) vs. these players off the court (S1 - left)

is more than just finding individual spatial differences. Endert et al. [12] discussed three levels of visual aggregations (i.e., glyph level, multiple glyphs merge, and the aggregation of multiple glyphs into a field) on large displays and how different encodings would impact such aggregations. For HoopInSight's comparison view, our goal was to facilitate recognition of group trends and patterns, while still maintaining the interpretability of the two variables associated with individual cells.

This goal fundamentally determined how we chose glyphs. Superimposing glyphs on a map is a common way to visualize multivariate spatial data. [30, 46]. Some of our initial visual encoding designs, as well as later suggestions from colleagues, used more complex glyphs to encode more variables. For example, while the four-quadrant axis design we used (Figure 3B) can display individual differences and encode more variables in one cell, identifying group patterns can be problematic due to visual clutter. Similarly, using multiple color scales to encode different variables can also be visually confusing and impede our perception of group patterns.

We recommend that designers consider the granularity of insights they want to convey and the priority of different levels. In our case, if the primary goal was to compare the multiple variables between individual cells, complex glyphs would likely be more effective. Conversely, if our primary goal was to communicate group patterns without worrying about individual interpretability, it would be reasonable to use a metric that aggregates individual cells and employ more perceptually dominant encodings (e.g., colors) to illustrate the patterns.

Consideration II: Neighboring spatial data characteristics

Design choices also depend on the strength of the connections between the data associated with individual cells and their adjacent cells. These connections may be based on factors such as whether the cells in a neighborhood exhibit roughly similar values or whether the delta follows a certain pattern. A strong connection indicates the existence of a "flow" — the data itself is inherently spatially dependent (such as wind maps or ocean current maps).

Using U.S. election maps as an example [5, 14, 31], the voting behavior of a county tends to be strongly connected to the ones of its neighbors. The glyphs representing such data tend to merge together and form a group pattern. In such cases, occlusions are less likely to occur and less troubling. In fact, visualization designers can utilize this congruence to amplify patterns and even create affective influence, as demonstrated by The Washington Post's political wind map [31]. However, when the data is less consistent between neighbors, using the same visual encoding may result in greater occlusion, making it challenging to reveal patterns. It might be more appropriate to apply less occlusion-prone encodings or adopt our approach — using two separated sub-views for increases and decreases, respectively.

6.2 Scalability and Transferability

By combining shot location data with further relevant information, HoopInSight can easily extend to provide more cutting-edge analytical insights. For instance, coupling data about team strategies, either of-

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3326910

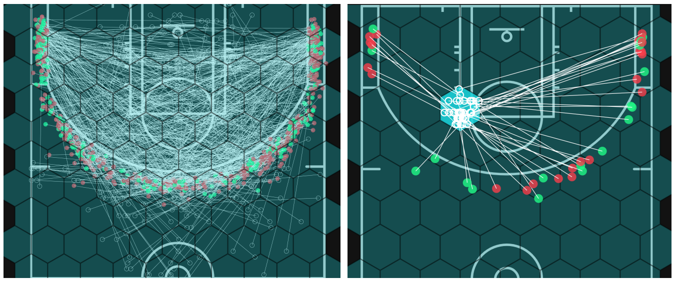


Fig. 10: A filter for passing-shooting spatial network analysis. Left: the entire spatial network. Right: a spatial filter is applied (via mouse hover) to view only the passes James made when he was inside the selected hexagon.

fensive or defensive ones, to existing datasets would lend HoopInSight the ability to provide insights into shooting performance/patterns with respect to such strategies (e.g., zone defense vs. man-to-man defense). By combining more temporal information about shots, including when they were taken during an offensive possession, HoopInSight could compare the shooting behaviors between early offense and late-clock shots in terms of spatial distribution and efficiency.

The system's spatial filtering capability can also be transferred to other scenarios, sports, or domains involving spatial analytics. Figure 10 showcases a variant of this technique — we developed a prototype using a limited experimental dataset that records LeBron James' passes to three-pointer shooters. The coordinates where James passes the ball were manually tracked and linked to the shots those passes led to. Together, they form a spatial network. The left figure displays the entire network, with white-outlined circles representing the *passing* coordinates and colored circles representing the *shooting* coordinates. The right figure showcases when a spatial filter is applied. The sizes of hexagons are adjustable, allowing users to explore at different levels of granularity. Although this is only a preliminary experiment, it has great potential to help coaches understand their players' passing patterns and habits and thereby adjust their game plans accordingly. We believe this interaction technique can benefit other sports and disciplines involving spatial data by enabling domain experts to create a spatial filter leveraging their prior domain knowledge.

6.3 Reflection on Designing as Domain Experts

Earlier, we outlined the reasons for adopting the Designing as Domain Experts (DaDE) method, such as adequate domain knowledge and available and accessible data. We believe that this method can yield significant benefits, but it does also incur dangerous pitfalls. Here, we further share our reflections on adopting DaDE.

Benefits

More manageable process: While a productive collaboration can bring many benefits, in reality, managing external interactions with collaborators is rarely easy. Sedlmair et al. enumerate a dozen pitfalls that may occur in the *winnow* and *cast* stage of a design study [41]. DaDE can provide a faster design process by avoiding the significant time external collaborations require. DaDE starkly contrasts our previous collaboration with two sports data companies. First, the available and accessible data allows for early idealization and data sketches. This is similar to what Oppermann and Munzner described in [35]. Second, collaborators, especially sports analytics companies, tend to care more about "just visualizing my data" (engineering) than "what visualization can do differently" (research). Adopting DaDE allows us to manage data acquisition and problem formation and reduce communication costs between multiple roles.

Focusing on novel solutions: In basketball, shot maps are a well-established solution. As previously noted, innovations in this field are mainly focused on two directions: the modeling approach that focuses on developing novel metrics and the narrative approach that emphasizes audience engagement. However, we believe that the addition of interactive features and new visual comparison techniques can foster more in-depth and comprehensive analysis. It also possesses great potential to be integrated with other approaches. We observed that the affordance and capabilities of interactive visualization have yet to be fully grasped by stakeholders. They also tend to exhibit conventional tool attachment [47]. As such, a traditional design study involving stakeholders tends to result in projects that facilitate existing tasks rather than challenging the common approaches. DaDE enables us to focus particularly on analytical tasks that are beyond existing approaches.

Synergy of domain knowledge and visualization expertise: *Expert designers* can use their knowledge in both fields to advance domain tasks through visualization solutions. In our case, our domain knowledge includes an understanding of high-level domain problems (what tasks are valuable to support) and a familiarity with the characteristics and availability of underlying data (what tasks can be supported). It lends us the flexibility to enter a design iteration from either *data abstraction* or *task abstraction* and rapidly match data and tasks. Our expertise in visualization serves as a catalyst for connecting tasks with data, creating a synergistic effect. It helps break down high-level domain goals into low-level tasks that can be supported by interactive visualization, as well as extract feasible tasks from multiple datasets and transform the data into a visualizable form. We feel such synergy is valuable, especially when the goal is to expand existing task space.

Pitfalls

Narrow domain perspectives: The term *domain expert* can be heterogeneous in real life. Simply differentiating *domain expert* from *casual user* or even *visualization expert* does not take adequate consideration of the heterogeneity of domain experts. In our case, different types of *domain experts* exist, such as sports journalists, sports analysts (data-driven or traditional "eye-test"), coaches/scouts, etc. While they share similar goals in analyzing shooting performance, their backgrounds and interests can vary drastically, which may affect their perception and adoption of such tools. Expert designers should be aware of their own positional biases and avoid a narrow domain perspective.

Prioritizing experts over casual users: Although prioritizing domain experts as end users may not necessarily be a pitfall, it is worth noting that DaDE naturally caters to them if the aim is to solve complex domain challenges. The resulting visualization interface will likely be sophisticated. Furthermore, as evidenced by our case studies, the insights provided by the visualization require strong domain knowledge to perceive and interpret. For example, casual users who are less familiar with the term "pick & roll" and different strategies to defend against it are less likely to acquire insights from Figure 9.

Task/data explosion: Task/data explosion can occur when the domain problem is unconstrained and the domain is data-rich. In such cases, the iterative cycle of data and tasks may spin out of control. This means that expert designers may keep pursuing additional tasks and data to address the domain problem in a more exhaustive way. While this may not necessarily be a pitfall, if the goal is to develop a functional product to address the domain problem, researchers should be aware of (visualization research) contribution saturation when new features are mere engineering.

Although DaDE is not a common method for design study, we believe our situation is not a lone case. Domain experts may play varying roles in the visualization design study process. We advocate for research to explore these benefits and pitfalls more systematically and provide methodological guidance for successful design endeavors.

7 Conclusion

This work aimed to enhance and expand the analytical capabilities of the shot chart, a well-known visual technique for basketball spatial analytics. To achieve this goal, we utilized our expertise in both visualization and basketball analytics and developed a novel visual technique for spatial comparison, along with an interactive visualization system called HoopInSight. The system supports a wide range of comparative analyses by allowing convenient selection of comparison entities and rapid creation of various comparison scenarios. Ultimately, we shared our reflections on and discussion about design considerations for visual-spatial comparison, the scalability and transferability of our approach, and adopting designing as domain experts (DaDE) methodology.

Supplemental Materials

Submitted supplemental materials for this paper include a document with enlarged figures, a supplemental glossary explaining the domain terms, and two videos: a main system introduction video with one in-depth case study and a secondary video for more case studies. The system is available at: https://hoopinsight.netlify.app/

References

[1] G. Andrienko, N. Andrienko, G. Anzer, P. Bauer, G. Budziak, G. Fuchs, D. Hecker, H. Weber, and S. Wrobel. Constructing spaces and times for tactical analysis in football. *IEEE Transactions on Visualization and Computer Graphics*, 27(4):2280–2297, 2021. doi: 10.1109/TVCG.2019.2952129 2, 3

[2] Basketball-Reference.com. Glossary - Effective Field Goal. https://www.basketball-reference.com/about/glossary.html, 2022. 4

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

[4] P. Beshai. Buckets: Visualizing NBA shot data. Technical report, University of British Columbia, 2014. 2, 3

[5] L. Buchanan, F. Fessenden, A. Parlapiano, K. Yourish, A. Pearce, T. Wallace, J. Keller, and J. Williams. How Trump reshaped the election map. New York Times: https://www.nytimes.com/interactive/2016/11/08/us/elections/how-trump-pushed-the-election-map-to-the-right.html, Nov 2016. 8

[6] D. Cervone, A. D'amour, L. Bornn, and K. Goldsberry. POINTWISE: Predicting points and valuing decisions in real time with NBA optical tracking data a new microeconomics for the NBA. In *Proceedings of the MIT Sports Analytics Conference*, 2014. 2

[7] W. Chen, T. Lao, J. Xia, X. Huang, B. Zhu, W. Hu, and H. Guan. GameFlow: Narrative visualization of NBA basketball games. *IEEE Transactions on Multimedia*, 18(11):2247–2256, 2016. doi: 10.1109/TMM.2016.2614221 2

[8] Z. Chen, Q. Yang, J. Shan, T. Lin, J. Beyer, H. Xia, and H. Pfster. iBall: Augmenting basketball videos with gaze-moderated embedded visualizations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pp. 841:1–841:18. ACM, 2023. doi: 10.1145/3544548.3581266 2

[9] Z. Chen, S. Ye, X. Chu, H. Xia, H. Zhang, H. Qu, and Y. Wu. Augmenting sports videos with VisCommentator. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):824–834, 2022. doi: 10.1109/TVCG.2021.3114806 2

[10] D. H. Chung, M. L. Parry, I. W. Griffiths, R. S. Laramee, R. Bown, P. A. Legg, and M. Chen. Knowledge-assisted ranking: A visual analytic application for sports event data. *IEEE Computer Graphics and Applications*, 36(3):72–82, 2016. doi: 10.1109/MCG.2015.25 2

[11] M. Du and X. Yuan. A survey of competitive sports data visualization and visual analysis. *Journal of Visualization*, 24(1):47–67, 2021. doi: 10.1007/s12650-020-00687-2 1

[12] A. Endert, C. Andrews, Y. H. Lee, and C. North. Visual encodings that support physical navigation on large displays. In *Proceedings of Graphics Interface 2011*, GI, pp. 103–110. Canadian Human-Computer Communications Society, 2011. 8

[13] Y. Fu and J. Stasko. Supporting data-driven basketball journalism through interactive visualization. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pp. 598:1–598:17, 2022. doi: 10.1145/3491102.3502078 2

[14] L. Gamio and D. Keating. How Trump redrew the electoral map, from sea to shining sea. Washington Post: https://www.washingtonpost.com/graphics/politics/2016-election/election-results-from-coast-to-coast/, 2016. 8

[15] J. Gels. Basketball terms - terminology. https://www.coachesclipboard.net/BasketballTerminology.html, 2020. 2

[16] M. Gleicher. Considerations for visualizing comparison. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):413–423, 2018. doi: 10.1109/TVCG.2017.2744199 2, 3, 8

[17] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts. Visual comparison for information visualization. *Information Visualization*, 10(4):289–309, Oct. 2011. doi: 10.1177/1473871611416549 2, 3

[18] K. Goldsberry. CourtVision: New visual and spatial analytics for the NBA. In *2012 MIT Sloan sports analytics conference*, vol. 9, pp. 12–15, 2012. 1, 2

[19] K. Goldsberry. *Sprawlball: A visual tour of the new era of the NBA*. Mariner Books, 2019. 1

[20] K. Goldsberry. The death of midrange. https://twitter.com/kirkgoldsberry/status/1227976828326797314?lang=en, 2020. 2

[21] D. Hickson and L. A. Waller. Spatial analyses of basketball shot charts: An application to Michael Jordan's 2001-2002 NBA season. In *Proceedings of the Hawaii International Conference on Statistics and Related Fields*, 2003. 2

[22] G. Hu, H. Yang, and Y. Xue. Bayesian group learning for shot selection of professional basketball players. *Stat*, 2021. doi: 10.1002/sta4.324 2

[23] J. Jiao, G. Hu, and J. Yan. A bayesian marked spatial point processes model for basketball shot chart. *Journal of Quantitative Analysis in Sports*, 17(2):77–90, 2021. doi: doi:10.1515/jqas-2019-0106 2

[24] J. V. Krause and C. Nelson. Basketball skills & drills. https://www.google.com/books/edition/Basketball_Skills_and_Drills/xDWx6uxC7DkC?hl=en&gbpv=1&pg=PA30&printsec=frontcover, 2019. 8

[25] P. A. Legg, D. H. S. Chung, M. L. Parry, M. W. Jones, R. Long, I. W. Griffiths, and M. Chen. Matchpad: Interactive glyph-based visualization for real-time sports performance analysis. *Computer Graphics Forum*, 31:1255–1264, 2012. doi: 10.1111/j.1467-8659.2012.03118.x 2

[26] T. Lin, Z. Chen, J. Beyer, Y. Wu, H. Pfister, and Y. Yang. The ball is in our court: Conducting visualization research with sports experts. *IEEE Computer Graphics and Applications*, 43(1):84–90, 2023. doi: 10.1109/MCG.2022.3222042 2

[27] T. Lin, Z. Chen, Y. Yang, D. Chiappalupi, J. Beyer, and H. Pfister. The quest for Omnioculars: Embedded visualization for augmenting basketball game viewing experiences. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):962–971, 2023. doi: 10.1109/TVCG.2022.3209353 2

[28] T. Lin, R. Singh, Y. Yang, C. Nobre, J. Beyer, M. A. Smith, and H. Pfister. Towards an understanding of situated AR visualization for basketball free-throw training. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pp. 461:1–461:13. ACM, 2021. doi: 10.1145/3411764.3445649 2

[29] A. G. Losada, R. Therón, and A. Benito. Bkviz: A basketball visual analysis tool. *IEEE Computer Graphics and Applications*, 36(6):58–68, 2016. doi: 10.1109/MCG.2016.124 2

[30] L. McNabb and R. S. Laramee. Multivariate maps—a glyph-placement algorithm to support multivariate geospatial visualization. *Information*, 10(10), 2019. doi: 10.3390/info10100302 8

[31] T. Meko, J. Fox, and D. Keating. The political winds in the U.S. are swirling. The Washington Post: https://www.washingtonpost.com/graphics/2020/elections/electorate-changes-2016-election-vs-2020/, Nov 2020. 8

[32] A. Miller, L. Bornn, R. Adams, and K. Goldsberry. Factorized point process intensities: A spatial analysis of professional basketball. In *Proceedings of the 31st International Conference on Machine Learning*, vol. 32 of *ICML*, pp. 235–243, 2014. 2

[33] S. Natarajan and S. Natarajan. Meet the data visualization king of basketball Twitter: Todd Whitehead. https://nightingaledvs.com/meet-the-data-visualization-king-of-basketball-twitter-todd-whitehead/, Nov 2022. 2

[34] NBA. Advanced stats: eFG%. https://www.nba.com/bucks/features/boeder-120917, Sep 2012. 4

[35] M. Oppermann and T. Munzner. Data-first visualization design studies. In *Evaluation and Beyond - Methodological Approaches to Visualization*, BELIV, pp. 74–80, 2020. doi: 10.1109/BELIV51497.2020.00016 9

[36] S. Patel. An API client package to access the APIs for NBA.com. https://github.com/swar/nba_api, 2022. 3

[37] C. Perin, R. Vuillemot, and J.-D. Fekete. SoccerStories: A kick-off for visual soccer analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515, 2013. doi: 10.1109/TVCG.2013.192 2

[38] C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, and M. S. T. Carpendale. State of the art of sports data visualization. *Computer Graphics Forum*, 37, 2018. doi: 10.1111/cgf.13447 1

[39] B. J. Reich, J. S. Hodges, B. P. Carlin, and A. M. Reich. A spatial analysis of basketball shot chart data. *The American Statistician*, 60(1):3–12, 2006. doi: 10.1198/000313006x90305 2

[40] D. Sacha, F. Al-Masoudi, M. Stein, T. Schreck, D. Keim, G. Andrienko,

and H. Janetzko. Dynamic visual abstraction of soccer movement. *Computer Graphics Forum*, 36:305–315, 06 2017. doi: 10.1111/cgf.13189 2

[41] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012. doi: 10.1109/TVCG.2012.213 2, 9

[42] L. Sha, P. Lucey, Y. Yue, X. Wei, J. Hobbs, C. Rohlf, and S. Sridharan. Interactive sports analytics. *ACM Transactions on Computer-Human Interaction*, 25(2):1–32, 2018. doi: 10.1145/3185596 2

[43] B. Strauss. Kirk Goldsberry once mapped food deserts and flood data. now his maps are changing the NBA. The Washington Post:https://www.washingtonpost.com/sports/2019/04/29/kirk-goldsberry-once-mapped-food-deserts-flood-data-now-his-maps-are-changing-nba/, Apr 2019. 1

[44] J. Wang, J. Wu, A. Cao, Z. Zhou, H. Zhang, and Y. Wu. Tac-miner: Visual tactic mining for multiple table tennis matches. *IEEE Transactions on Visualization and Computer Graphics*, 27(6):2770–2782, 2021. doi: 10.1109/TVCG.2021.3074576 2

[45] T. Whitehead. Steph Curry's record-breaking 2974 career threes. https://twitter.com/CrumpledJumper/status/1471938165061554178, 2021. 2

[46] H. Wickham, H. Hofmann, C. Wickham, and D. Cook. Glyph-maps for visually exploring temporal patterns in climate data and models. *Environmetrics*, 23(5):382–393, 2012. doi: 10.1002/env.2152 8

[47] Y. L. Wong, K. Madhavan, and N. Elmqvist. Towards characterizing domain experts as a user group. In *Evaluation and Beyond - Methodological Approaches for Visualization*, BELIV, pp. 1–10, 2018. doi: 10.1109/BELIV.2018.8634026 9

[48] Y. Wu, D. Deng, X. Xie, M. He, J. Xu, H. Zhang, H. Zhang, and Y. Wu. OBTracker: Visual analytics of off-ball movements in basketball. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):929–939, 2023. doi: 10.1109/TVCG.2022.3209373 2

[49] Y. Wu, X. Xie, J. Wang, D. Deng, H. Liang, H. Zhang, S. Cheng, and W. Chen. ForVizor: Visualizing spatio-temporal team formations in soccer. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):65–75, 2019. doi: 10.1109/TVCG.2018.2865041 2

[50] N. Yau. Goodbye, mid-range shot | flowingdata. https://flowingdata.com/2019/01/15/goodbye-mid-range-shot/, 2019. 2

[51] J. S. Yi, Y. a. Kang, J. Stasko, and J. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231, 2007. doi: 10.1109/TVCG.2007.70515 6

[52] Q. Zhi, S. Lin, P. Talkad Sukumar, and R. Metoyer. GameViews: Understanding and supporting data-driven sports storytelling. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pp. 269:1–269:13. ACM, 2019. doi: 10.1145/3290605.3300499 2