



City Research Online

City St George's, University of London

Citation: Koytek, P., Perin, C., Vermeulen, J., Andre, E. & Carpendale, S. (2017). MyBrush: Brushing and Linking with Personal Agency. IEEE Transactions on Visualization and Computer Graphics, 24(1), pp. 605-615. doi: 10.1109/tvcg.2017.2743859

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/18383/>

Link to published version: <https://doi.org/10.1109/tvcg.2017.2743859>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

MyBrush: Brushing and Linking with Personal Agency

Philipp Koytek, Charles Perin, Jo Vermeulen, Elisabeth André, and Sheelagh Carpendale

Abstract— We extend the popular brushing and linking technique by incorporating personal agency in the interaction. We map existing research related to brushing and linking into a design space that deconstructs the interaction technique into three components: source (what is being brushed), link (the expression of relationship between source and target), and target (what is revealed as related to the source). Using this design space, we created MyBrush, a unified interface that offers personal agency over brushing and linking by giving people the flexibility to configure the source, link, and target of multiple brushes. The results of three focus groups demonstrate that people with different backgrounds leveraged personal agency in different ways, including performing complex tasks and showing links explicitly. We reflect on these results, paving the way for future research on the role of personal agency in information visualization.

Index Terms—Brushing, linking, personal agency, coordinated multiple views, interaction, design space, information visualization

1 INTRODUCTION

Starting from the 1987 application to scatterplots [6], brushing and linking has become a prevalent technique for interactively exploring data relations among several visualizations in coordinated multiple views (CMV) [65]. Hearst defined brushing and linking as “*the connecting of two or more views of the same data, such that a change to the representation in one view affects the representation in the other views as well*” [37]. This broad definition reflects the great variety of ways brushing and linking has been incorporated into visualizations. It has become one of the most implemented interaction techniques in infovis and is frequently included in research, popular toolkits like D3.js [9], and commercial products like Tableau [72].

Although many brushing and linking variations exist and different techniques may be useful at different times during data exploration, visualizations typically provide just one form of brushing and linking between views. Common examples include hovering over a data point to highlight related data points in other views (e.g., [11, 27, 76]), and drawing a shape to select data points (usually a rectangle) and highlight the corresponding data points in other views (e.g., [6, 12, 53, 71]). These implementations do not provide access to the full power of brushing and linking. In response, we explore whether the incorporation of personal agency can augment brushing and linking interactions and further empower exploration of data. In using Coyle et al.’s definition of personal agency, “*a persons innate sense of being in control of their actions and through this control of being responsible for, or having ownership of, the consequences of these actions*” [17], we recognize the importance of supporting an individual’s expression of their intentions by providing them with the ability to make personal choices and feel in control of their actions. We explore personal agency in brushing and linking by: i) providing direct access to brush components and making these configurable; ii) offering choice in degree of personal agency so that people can adapt the configuration complexity to their needs and expertise; and iii) supporting options for increasing complexity of personal agency through personal actions.

Examining brushing and linking from the perspective of personal agency led us to develop a design space that describes and compares the

variations of brushing and linking in the literature. This design space deconstructs brushing and linking into three components: *source* (what is being brushed), *link* (the expression of relationship between source and target) and *target* (what is revealed as related to the source). Leveraging this design space, we integrate several brushing and linking techniques to create an interface called MyBrush. MyBrush incorporates personal agency in configurable brushing and linking, letting people decide how data points will be brushed and linked, and providing the freedom to choose the technique suited to their preferences and current exploration task. We conclude by discussing the results of three focus groups which show that MyBrush offers personal agency by letting people of different backgrounds adapt brushing and linking to preferences and needs.

2 BACKGROUND

Before deconstructing existing brushing and linking techniques, we briefly overview coordinated multiple views with brushing and linking, and frame personal agency in the context of interactive interfaces.

2.1 Coordinated Multiple Views and Brushing and Linking

The concept of CMV emerged in the late 1990s [64] and has since become a core body of research in infovis. The underlying idea of CMV is that people “*understand their data better if they interact with the presented information and view it through different representations*” [65]. Although guidelines exist for CMV systems [78], Roberts [65] debate whether CMV are a “*solved problem*” or if the visualization community has “*only just scratched the surface of the subject*.” The continuing emergence of new CMV applications (e.g., [14, 15, 21, 68, 74, 76]), and the exploration of new design spaces (e.g. of composite visualizations [43] and meta-visual view-relations [45]) suggest the latter.

Brushing and linking is a powerful interaction technique for exploring different facets of a dataset [44], is integrated in most modern CMV systems [65], and is widely adopted by visualization tools such as Tableau [72] and SpotFire [4], and D3.js [9]. Since the introduction of brushing and linking for scatterplot matrices [6], researchers have proposed many variations of the technique. Martin and Ward [51, 79] provide brushes with the same dimensionality as the attribute space (N-dimensional brushing), operations on brushes, and logical combinations of multiple brushes. Angular brushing [36] is a specialization for parallel coordinates. Roberts et al. [66] add menus and axes to brushing. Further notable variations include: multi-resolution brushing and linking [82], smooth brushing [19] for fuzzy selection of data points, structure-based brushes to navigate hierarchies in the data [25], and constrained brushing that is augmented with quantitative statistics [63].

Despite its numerous variations, possible generalizations of brushing and linking have been restricted to creating models and toolkits that enable custom coordination of views, with the goal of supporting coordinating unforeseen combinations of different visualization types (e.g., [10, 13, 39, 57, 80]). These models, which are mostly based on visual programming paradigms or formalisms, are challenging to use in practice because: i) creating custom coordinations requires people

- Philipp Koytek is with University of Calgary and Augsburg University. E-mail: philipp.koytek@me.com.
- Charles Perin is with City, University of London and University of Calgary. E-mail: charles.perin@city.ac.uk.
- Jo Vermeulen and Sheelagh Carpendale are with University of Calgary. E-mail: jo.vermeulen@ucalgary.ca, sheelagh@ucalgary.ca.
- Elisabeth André is with Augsburg University. E-mail: andre@informatik.uni-augsburg.de.

to spend significant amounts of time configuring the views, leading to a temporal separation between coordinating views and exploring the data; and ii) these models require advanced levels of expertise to understand the underlying paradigms. Pantazos et al. [58] reviewed 13 popular visualization toolkits and found that little attention is paid to supporting non-experts in constructing advanced visualizations.

To address these challenges, we look into personal agency i) to support in-place view coordination by providing flexibility over brushing and linking interactions between multiple views; and ii) to enable people of different backgrounds and visualization expertise to quickly adapt brushing and linking behavior to their differing needs.

2.2 Interactive Interfaces and Personal Agency

Shneiderman’s [69] seventh golden rule of interface design emphasizes the importance of supporting individuals’ internal locus of control, as people “*strongly desire that they are in charge of the interface and that the interface responds to their actions*”. In cognitive neuroscience, this is called *personal agency* and refers to the experience of being in control of one’s own actions and, through them, affecting the external world [32]. The degree of personal agency is directly linked to the freedom of choice [5], thus increases with the number of available actions. See Moore’s [54] work for a review of personal agency.

Bridging the gap between cognitive neuroscience and HCI, Coyle et al. [17] found that personal agency is an important factor for how people experience technology. They suggest that “*interaction techniques that facilitate a high sense of personal agency are likely to have a strong empowering effect for users*”. In particular, personal agency will allow people interacting with technology to instinctively sense that *they* did something, as opposed to “the system did something” [17], thus providing a sense of control. Our approach consists of leveraging the notion of personal agency by providing people with the freedom to choose their own actions in brushing and linking interactions.

3 DECONSTRUCTION OF BRUSHING AND LINKING

To leverage personal agency in brushing and linking, we first present a design space based on a review of the literature and a deconstruction of the technique into three components: *source*, *link*, and *target*.

3.1 Source, Link, and Target

The interactive process of brushing and linking can be described in three steps. First, data points of interest are selected, for example by drawing a rubberband rectangle or clicking their visual marks. Second, the selection propagates to other views of interest that provide an alternate perspective on the data. Third, this propagation changes the views to make data points, related to the selected data points, distinguishable. From this, we derive the three core components of brushing and linking:

1. **Source** is the set of one or more selected data points in a view.
2. **Link** is the expression of relationship between the source and the related data points in other views (target).
3. **Target** is the set of data points that are related to the source.

Figure 1 illustrates three variants of these components. In all three, the **source** consists of the four green data points within the rectangular brush in the view on the left. The **target** is a set of bars in a barchart that *aggregates* data points together in categories. In (A), the **link** is implicit, i.e. links between source and target are not explicitly represented. In (B), links are explicit using lines to connect source to target data points, and a bar of the barchart can be the target of several source data points. In (C), the target is *de-aggregated* creating a one-to-one mapping between source and target and identifying individual, previously aggregated, data points. This deconstruction of brushing and linking into source, link, and target opens up a new design space for describing, comparing and generating brushing and linking techniques.

3.2 Design Space of Brushing and Linking

To examine the core brushing and linking papers in infovis, we used a snowballing collection approach. That is, we defined our relevancy criteria, collected a group of possibly important seed papers, applied our criteria, used forward and backward referencing to collect more

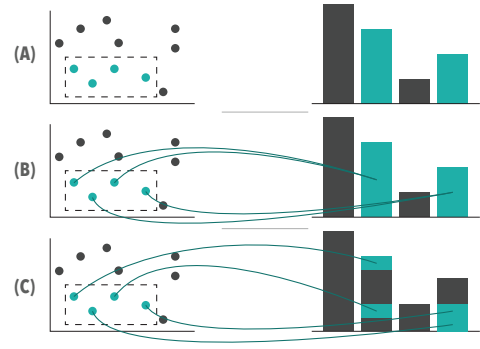


Fig. 1. Conceptual examples of brushing and linking techniques with different source, link and target configurations: (A) green source highlight, no visual link, green target highlight; (B) green visual links added; and (C) green source, visible links and de-aggregated target.

papers, to which in turn we applied our criteria and repeated until we did not find any new aspects to the design space. Relying on existing citation patterns favours papers that the community considers important. To fill in gaps, we used keyword searches for “brushing and linking”, “infovis toolkit”, and “coordinated multiple views” on Google Scholar.

We based our criteria on Hearst’s [37] definition of brushing and linking: we considered techniques that provide selection of *source* data points (brushing) and expression of relationship between source and target with two or more coordinated views (linking). In line with Hauser’s [35] categorization of brushing as an *explicit, on-view* focusing technique, we focused on brushing through direct manipulation. We relaxed our definition of views to include techniques which are similar to brushing and linking in CMV but use a different terminology for views, e.g., we interpret the blocks in Domino [27] as views.

We applied this recursive process starting from 17 seed papers about meta visualization and CMV. This resulted in a list of 31 papers, which although not exhaustive, is well suited to laying the groundwork of our design space. We categorized the 31 papers in Table 1. Rows r_i are brushing and linking papers, or “techniques”. Columns c_j represent attributes of source, link, and target that we found in our review. For example, the category “temporality” of the source includes the attributes *transient*, *temporary* and *persistent* (c_{10} to c_{12}). We use a 4-point scheme in Table 1: no glyph means that the technique in row does not feature the attribute in column; a light-grey glyph means it is unclear whether or not it features the attribute; a dark-grey glyph means it features the attribute; and an orange glyph means that the attribute can be interactively configured. We structure our discussion of the design space according to the horizontal order of the columns in Table 1.

3.2.1 Source: What is Being Brushed

The source is the selection of data points created by brushing. Columns c_1 to c_{16} describe the attributes of source.

Visual Attributes (c_1 – c_8) are used to emphasize the source data points. They include visual variables [7] such as color and shape, as well as visual marks such as labels and containers. While any visual attribute can be used to create emphasis [33], only a handful of visual attributes have been used to emphasize the source [67, 81]. 24/31 techniques use the *fill color* of the source data points, and six of them provide interactive choice of color [11, 13, 22, 28, 42, 55]. Other visual attributes that have been used are *outline color* (5/31), *size* (Compound Brushing [13], r_{31}), *shape* (2/31), focus and blur (GeoViz Toolkit [34], r_{10}), and *transparency* (2/31). Nine techniques show *labels* or tooltips for the source, and four enclose the source data points using *containers* such as rectangular frames, bubblesets, linesets and convex hulls. Three techniques do not use visual attributes of the source; Caleydo [48] (r_6) and Matchmaker [49] (r_7) show visual links to point at the source data points, and in Show Me The Invisible [26] (r_5) the source is a search term entered by a person, with search hits (target) being linked across multiple windows using attributes for link and target, but not source.

Group Selection (c_9) is available in 24/31 techniques in addition to single-item selection, which is featured by all. While single-item

		SOURCE																LINK										TARGET											
		VISUAL ATTRIBUTES								GROUP SELECTION	TEMPORALITY		MULTIPLE SELECTIONS		LOGICAL COMBINATION	DEGREE OF INTEREST FUNCTION	VISUAL ATTRIBUTES						ROUTING	SELECTIVE LINKING OF VIEWS	VISUAL ATTRIBUTES						DE-AGGREGATION								
		FILL COLOR	OUTLINE COLOR	SIZE	SHAPE	FOCUS AND BLUR	TRANSPARENCY	CONTAINER	LABEL	Mouse Rectangle Circle Polygon Lasso Angle Line	TRANSIENT	TEMPORARY	PERSISTENT		BINARY	NON-BINARY	STROKE COLOR	COLOR GRADIENT	THICKNESS Thin Ribbon Variable	CURVATURE Straight Stepwise Curved	TRANSPARENCY	LINK STUBS	ANIMATION	CONTEXT PRESERVING	BUNDLING	View to view View to view Brush to view Brush to view	FILL COLOR	OUTLINE COLOR	SIZE	SHAPE	FOCUS AND BLUR	TRANSPARENCY	HIDE UNSELECTED	CONTAINER	LABEL				
	YEAR	1980	2000	2020	1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1	VISUAL LINKS ACROSS APPLICATIONS [77]																																						
2	ENTOURAGE [47]																																						
3	DOMINO [27]																																						
4	INFOVIS TOOLKIT [24,62]																																						
5	SHOW ME THE INVISIBLE [26]																																						
6	CALEYDO [48]																																						
7	MATCHMAKER [49]																																						
8	GRAPHTRAIL [21]																																						
9	CONNECTEDCHARTS [76]																																						
10	GEOVIZ TOOLKIT [34]																																						
11	VISLINK [15]																																						
12	CONTEXT-PRESERVING VISUAL LINKS [70]																																						
13	ROLLING THE DICE [22]																																						
14	BRUSHING OF ATTRIBUTE CLOUDS [42]																																						
15	INFOSCOPE / CITY-O-SCOPE [11]																																						
16	TREEJUXTAPOSER [55]																																						
17	PARALLEL TAG CLOUDS [16]																																						
18	BRUSHING SCATTERPLOTS [6]																																						
19	XGOBI [71]																																						
20	FOCUSING AND LINKING [12]																																						
21	GEOGRAPHIC BRUSHING [53]																																						
22	COLOR M-AND-N PLOTS [52]																																						
23	CONSTRAINED BRUSHING [63]																																						
24	MONDRIAN [73]																																						
25	BRUSHING FUNCTION GRAPHS [46]																																						
26	INTERACTIVE FEATURE SPECIFICATION [18]																																						
27	ANGULAR BRUSHING [36]																																						
28	XMDVTOOL [51,79]																																						
29	SMOOTH BRUSHING [19]																																						
30	WEAVE [28]																																						
31	COMPOUND BRUSHING [13]																																						
32	MYBRUSH																																						

Table 1. Categorization of selected brushing and linking techniques in terms of source, link and target. Techniques were ordered by overall similarity using Bertifier [60]. The last row shows MyBrush, the implementation we present in this paper.

selection is usually available through clicking or hovering over a data point (e.g., [11, 27, 76]), the mechanisms for group selection vary more. 5/31 techniques use a mouse mechanism. Four of these feature incremental selection building by clicking on data points, and Color M-and-N-plots [52] (r_{22}) allow for selecting data points in the vicinity of the mouse pointer. In most cases, however, group selection is done by drawing a shape in a view to select enclosed data points. The shape is often a rectangle (e.g., [6, 12, 53, 71]), but it can also be a circle, a polygon or a free-form lasso — Attribute clouds [42] (r_{14}) lets people select among three shapes. Constrained brushing [63] (r_{23}) provides rectangular and circular percentile brushes that always contain a fixed number of data points. A special case is the Mahalanobis brush, which dynamically adapts its elliptical shape to the underlying data distribution as it is moved. Angular Brushing [36] (r_{27}) features group selection that specifies an angle between two parallel coordinates to collect line segments whose slopes fall within the specified angle. Konhya et al. [46] (r_{25}) proposed line brushes that act as persistent crossing interactions [3] to select function graphs that intersect the line.

Temporality (c_{10} – c_{12}) of source refers to the lifetime of a brush. We saw three types of temporality (*transient*, *temporary* and *persistent*). *Transient* brushes (12/31) are only active (i.e. visible in the interface) as long as a cursor is pointing at the source. They are typically created by hovering the mouse over a data point, and removed by moving the mouse away from the data point. This interaction, called “probing” in InfoScope [11] (r_{15}), is dedicated to viewing more details of an object and learning the relationships between the different views. *Temporary* brushes (18/31) remain active until a new brush is created. As the replacement is done automatically, there cannot be more than one active brush in the application (e.g., [15, 53, 77]). *Persistent* brushes (16/31) remain active until explicitly removed. In Brushing Scatterplots [6] (r_{18}), using “lasting paint” creates a brush that stays until undone.

Multiple Selections (c_{13}) are possible only with persistent brushes. 12/31 techniques support the creation of multiple brushes. WEAVE [28] (r_{30}), for instance, provides three different highlighting colors for people to create three visually distinct brushes.

Logical Combination (c_{14}) is available in 8/31 to form compound selections using logical operations. The strategy for realizing logical combination differs between techniques. For example, Mondrian [73] (r_{24}) provides five logical operations to either replace the current selection or refine it with AND, OR, XOR, and NOT operations. Thus, as with Brushing Function Graphs [46] (r_{25}), each new brush iteratively can only extend the latest selection state. In contrast, Doleisch et al. [18] (r_{26}) maintain a tree of logical brush combinations. Finally, Compound Brushing [13] (r_{31}) provides a diagrammatic language to build logical combinations between abstractions of selections.

Degree-of-Interest Functions (c_{15} – c_{16}) refer to how data points can be brushed. 30/31 techniques provide a *binary* function: a data point is either part of the selection or not. *Non-binary* functions allow for multi-degree selection of data points. Color M-and-N-plots [52] (r_{22}) uses a ternary DOI-function using three colors to indicate a point's distance to the cursor; and 4/31 techniques allow for defining smooth selection boundaries, ie. non-discrete degree-of-interest functions.

3.2.2 Link: *The Relationship between Source and Target*

The link is the expression of relationship between source and target across different views. Columns $c_{17}-c_{26}$ describe attributes of the link.

Visual Attributes (c_{17} – c_{22}) are more sparsely used for links than sources because links are often implicit. 11/31 techniques represent links explicitly, as in integrated views [43], usually using a solid *stroke color* to make connections visible. VisLink [15] (r_{11}) uses a *color gradient* to show directionality, and alpha-blending (i.e., *transparency*) to encode the number of lines in a bundle. Line *thickness*, varied from

thin, to thick, to variable, where the width of the link narrows in the middle and expands at the endpoints. Variations of link *curvature*, were usually either straight or curved, with the exception of Domino [27] (r_3) which uses angled, or stepwise, links. Entourage [47] (r_2) uses *link stubs* to reduce visual clutter. Note that 4 techniques differ in that they permanently show visual links between corresponding data points and only change the stroke color upon brushing [21, 27, 49, 76] (r_3, r_7 – r_9).

Animation (c_{23}) is used only in Visual Links Across Applications [77] (r_1), which lets people choose to animate links to fade out after some time to reduce visual clutter and fade back in when needed.

Routing (c_{24} – c_{25}) is used to optimize the route of links. *Context-preserving visual links* [70] (r_{12}) are routed to avoid crowded regions. Link bundling can also minimize occlusion by bundling links together to take common routes (4/31). Caleydo [48] (r_6) first bundles all connection lines per view and then joins all link bundles.

Selective Linking of Views (c_{26}) enables people to specify which views are linked (6/31). This can be done using visual coordination schemata external to the data visualizations (The Infovis Toolkit [24, 62], r_4 and Compound Brushing [13], r_{31}), or directly in the CMV system through menus (The GeoViz Toolkit [34], r_{10}) and by linking adjacent views that can be reorganized via e.g., drag and drop (Domino [27], r_3 , ConnectedCharts [76], r_9 , and VisLink [15], r_{11}).

3.2.3 Target: Data Related to the Source

The target is the set of data points in other views that have a relationship to the source. Columns c_{27} to c_{36} describe the attributes of target.

Visual Attributes (c_{27} – c_{35}) of targets are almost identical to visual attributes of sources, since they are usually assigned the same attribute values. In Brushing Scatterplots [6] (r_{18}), target data points are highlighted using the same fill color as their sources. An exception is *hide unselected*. The shadow highlight mode in Brushing Scatterplots [6] (r_{18}) hides unrelated data points in target views, and Parallel Tag Clouds [16] (r_{17}) also hides non-selected data points in combination with magnifying the *size* of selected ones. A few techniques also apply separate visual attributes to the target than to the source. Similarly to highlighting with a text marker, Show Me The Invisible [26] (r_5) uses rectangular frames, i.e. *containers*, with *fill* and *outline color* to mark occurrences of a search term. In VisLink [15] (r_{11}), the *transparency* of target nodes increases the further away their view plane is from the source. In Smooth Brushing [19] (r_{29}) for non-discrete DOI functions, *color*, *size* and *transparency* of objects can optionally be modulated according to their degree of interest.

De-Aggregation (c_{36}) refers to the ability to de-aggregate an aggregated target (Figure 1). 6/31 techniques de-aggregate barcharts and histograms. The bars of these charts are partially highlighted (typically with color) to provide a one-to-one mapping between source and target.

3.3 From Deconstruction to Personal Agency

Deconstruction of brushing and linking has provided us with three core components — source, link, and target — whose attributes can be varied in many ways, although existing techniques rarely support interactive configuration of these attributes. Supporting personal agency in an interactive tool suggests providing people with greater ability to make choices with the tool (to offer more freedom of choice on the appearance, behaviour and mechanics of the tool). We combine this deconstruction with the definition of personal agency to frame three design goals for supporting personal agency in brushing and linking:

- DG1** *Provide direct access to brush components.* Making the brushes' components — source, link, and target — explicitly visible and configurable for each brush should offer the possibility of combining these components in different ways.
- DG2** *Offer choice in degree of personal agency.* Since the concept of interacting with personal agency might be unfamiliar, offering freedom of choice should allow people to use a little personal agency, to incrementally increase their use of it, and to delve into increasing complexity at will.
- DG3** *Support complex personal agency.* More complex agency, leading to more powerful interactions, should be supported through multiple independently configurable brushes.

4 MYBRUSH: CONFIGURABLE BRUSHING AND LINKING

Based on our design goals, we designed and implemented MyBrush, a CMV environment that provides personal agency in brushing and linking. MyBrush supports various types of visualizations and multiple simultaneous brushes. For each brush, the source, link, and target are configurable using local menus. To show the functionality of MyBrush, we use a dataset of the attributes and skills of the 50 most valuable soccer players from the EA Sports FIFA video game (see sofifa.com).

4.1 Basic Functionality of MyBrush

Figure 2 shows 10 soccer players in two views: a scatterplot, showing attacking and defending skills of the players represented by circles, and a barchart, showing the number of players in a given position on the soccer field. Note that while we describe MyBrush with touch-based interaction, it is also compatible with point-and-click mouse interaction.

Figure 2 row (A). Tapping on a data point in View 1 changes its fill color to red, as well as the fill color of all related data points in View 2 (default brushing and linking interaction). Because the bar chart aggregates data points, the full bar that contains the selected player is red. Tapping on an empty area resets the fills to grey.

Figure 2 row (B). Dragging a finger across View 1 draws a rectangular selection [6]. The attached recycle bin icon makes removal simple. The source, link, and target of the brush are not yet configured.

Figure 2 row (C). The brush is configured using three local menus for source, link, and target (see Figure 3). Tapping on the left-most menu  expands the *source* menu. Setting the fill color of the source to orange  sets the fill color of all data points within the brush to orange. The source menu's icon also changes to orange (visual feedback). In most cases, having the same configuration of attributes for source and target is sufficient. By default, the target is coupled to the source, as visually represented by the target menu's center turning the same color as the source and by the chain link icon on the target menu. The menu's center can be used to move or close the menu. Thin lines visually connect the menus to their associated brush (see Figure 3).

Figure 2 row (D1). The second menu configures the *link*. Tapping on the crosshair icon opens a sub-menu containing all views that can be linked to the brush. Here, View 2 is linked to the brush , which applies orange fill to all bars related to the source. In **row (D2)**, link stroke color is set to green  and curvature to straight lines . This shows explicit links (green lines) between source and target data point.

Figure 2 row (E). The third menu configures the *target*. Changing any attribute in the target menu breaks the default coupling with the source. This is shown by a broken chain link symbol. Tapping on the chain link symbol re-couples the target and source. Choosing “de-aggregate”  shows the one-to-one mapping between source and target by highlighting parts of the bars. We arbitrarily chose to order data points within aggregated views by value, i.e. the most-valuable players are at the bottom of the bars. Other target attributes can be configured to better discriminate individual players within the aggregated data, such as setting the outline color of the target to brown .

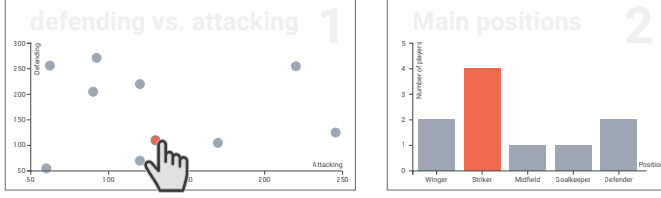
4.2 Combining Brushes in MyBrush

MyBrush allows for performing advanced iterative queries with multiple brushes. Figure 4 shows a three-brush example, with three views of the soccer dataset. View 1 shows the distribution of positions on the soccer field in a bar chart, View 2 is a scatter plot of the defending and attacking skills of the players, and View 3 shows the player names.

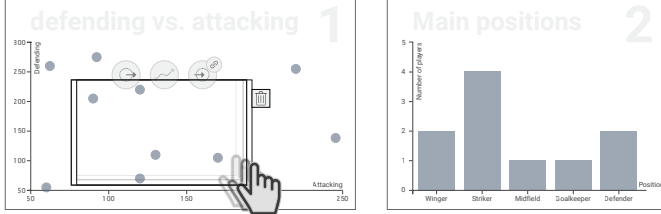
Figure 4 (A). Suppose that Ann is interested in finding out who the best strikers are. She creates a brush enclosing the striker bar in View 1. She colors the source green and links the brush to both View 2 and View 3. As the target and source are coupled, the strikers are highlighted in green in View 2 and View 3.

Figure 4 (B). Looking at the scatter plot in View 2, Ann notices that four strikers are not only good at attacking, but also reasonably good at defending. To discover who these players are she draws a new brush around the four strikers with the highest defending skills. She then links the brush to View 3, uses the link menu to draw straight blue lines between source and target, and sets the outline color of the target to blue. By doing this, Ann visualizes the intersection of the two brushes

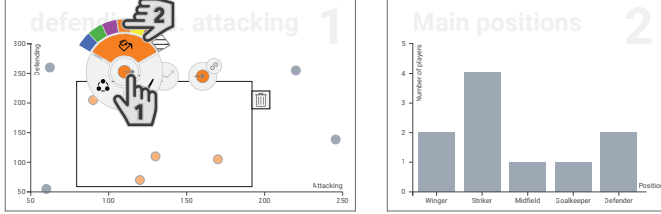
(A) TAP DATA POINT Tapping a data point in view 1 highlights the data point in all other views. The bar in view 2 that contains the selected data point is highlighted.



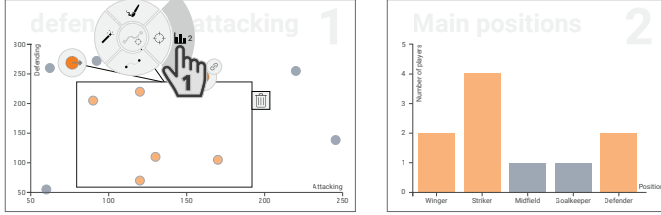
(B) BRUSH DATA POINTS Dragging the finger across view 1 creates a rubber-band brush rectangle to select data points.



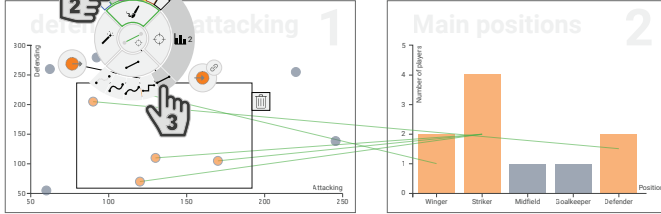
(C) CONFIGURE SOURCE (1) Expanding the source menu and (2) setting the fill color to orange, fills selected data points with orange color. This is reflected by the source menu root (now orange filled circle).



CONFIGURE LINK (1) In the link menu, linking the brush to view 2 highlights related bars in view 2 with orange fill color. The highlighted bars each contain at least one target data point.



(D2) Setting (2) the stroke color to green and (3) the line curvature to straight shows straight green lines between source and target data points.



(E) CONFIGURE TARGET (1) De-aggregating the target and (2) setting its outline color to brown shows the individual target data points within the aggregated bars. The chain link symbol indicates that the coupling with the source is broken.

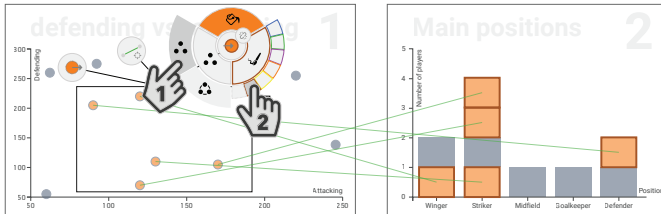


Fig. 2. The basic functionality of MyBrush illustrated using a scatterplot of attacking vs. defending skills in which each circle represents a soccer player (View 1, left) and a barchart of player positions on the soccer field (View 2, right). We demonstrate support for personal agency in the creation and configuration of a single brush in View 1.

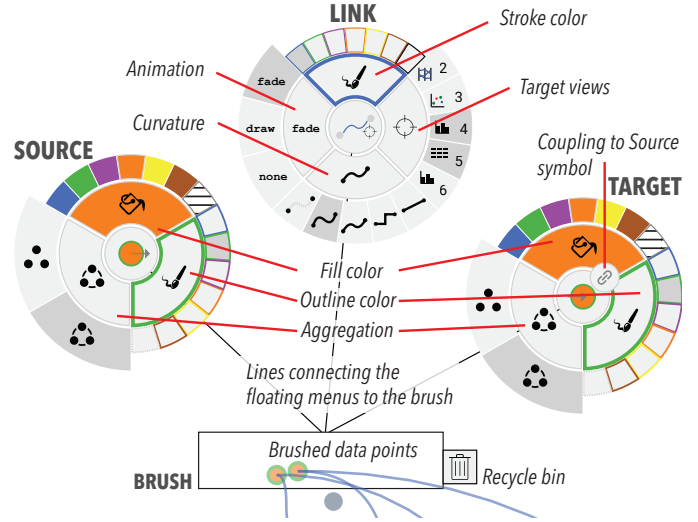


Fig. 3. Three hierarchical menus configure the source, link and target of a brush. They are collapsed by default but can be expanded and re-positioned. The source and target menus configure fill color, outline color, and aggregation mode of data points. Here the source is configured with orange fill and green outline colors. Since the target is coupled to the source, indicated by the chain link symbol, their configurations are the same. The link menu configures stroke color, animation, and curvature of explicit links. It also configures target views to be linked to the brush. Here the link is configured with a blue stroke color, a fade-in fade-out animation, a smooth curvature, and with Views 4 and 5 as target views.

in View 3 by using two non-conflicting attributes (fill and outline colors of the rectangular containers around the player names): players with a green fill are strikers; the ones with a blue outline are reasonably good defenders, such that the ones with a green fill *and* a blue stroke are the best defending strikers: L. Suárez, T. Müller, W. Rooney and D. Costa.

Figure 4 (C). Ann notices two non-highlighted, grey data points in the bottom-right corner of View 2. They have very high attacking skills but are not strikers. She is intrigued and creates a third brush to investigate these players. She sets the fill color of the source to purple and links the brush to View 3. This highlights two players in View 3: L. Messi and Cristiano Ronaldo.

5 EXPLORING THE ROLE OF PERSONAL AGENCY IN BRUSHING AND LINKING

To investigate the potential of providing personal agency with brushing and linking, we conducted an observational qualitative study consisting of three focus groups where participants freely explored a multi-dimensional sports dataset using MyBrush.

5.1 Study Design

Methodology: The goal of our study was to investigate if personal agency would be used by people exploring data, and if so, whether it offered any advantages. We chose the focus group study format because it encourages spoken, exploratory discussions between participants in an informal setting where time pressures seem much less apparent than in a controlled setting [50]. This setting supports natural discussion, which can be quite informative because the participants need to help the group understand each other's actions in context of the problems they are working on [23, 75]. It is also analogous to real-world situations since data analysis is often a collaborative group activity [61]. Our intention was to gain a better understanding of whether the need for personal agency would arise in self-driven data exploration and how the participants might use the provided personal agency. We were particularly interested in exploring how the need for and use of personal agency might vary according to people's background and knowledge. For example: will people familiar with traditional brushing make use of personal agency? Is personal agency more useful for people who are interested in the data, or for people less familiar with either the data or traditional brushing techniques?

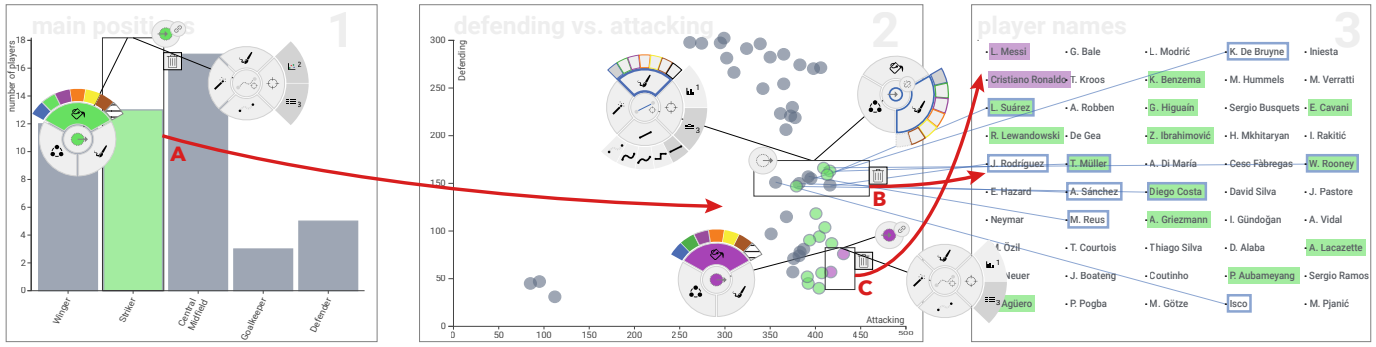


Fig. 4. Creating multiple brushes to iteratively refine explorations. (A) Strikers are highlighted in green in View 2 and View 3. (B) Strikers with good defending skills now also have a blue outline color in View 3. (C) The two players who are not strikers but have high attacking skills are highlighted in purple in View 3. *Note:* Menus are fully expanded here to better understand the figure. All menus can be collapsed to free up screen real estate.

Participants: We recruited twelve participants (5 F, 7 M) using campus-wide posters, mailing lists and word of mouth. We formed three focus groups of participants with different backgrounds: the *vis* group (2F, 2M: V1–V4) contained participants who were particularly familiar with visualization and brushing and linking; the *sports* group (4M: S1–S4) included people who were particularly enthusiastic about sports; and the *mixed* group (3F, 1M: M1–M4) included people less familiar with either the data or brushing and linking. In this mixed group participants had varying backgrounds, as shown in Figure 5.

Rationale for Data and Task: For the study, we used the *sofi fa.com* dataset of the 50 most valuable soccer players (Section 4). The FIFA video game series is the best selling sports video game franchise in the world [1] and has sold over 150 million copies [59]. The *sofi fa.com* website provides player statistics extracted from the video game. It reaches over 13 million visitors per month [2] and is used by gamers to find high potential soccer players and compare their skills. The data of the video game is so accurate that professional soccer players compare their ratings in new editions of the game and teams use the data to scout for potential signings or prepare for upcoming matches [59]. For these reasons, our focus groups reflect a realistic scenario in which people explore the FIFA data to discuss their favourite teams, decide who is the best player for certain positions or style of play, or which players a team should buy or sell. Using this dataset also made it possible to recruit participants that know and are interested in the data (sports group). We selected the most valuable players to increase the likelihood that participants would recognize these world-class players.

Setup: We configured MyBrush with six views (1 to 6; see Figure 6) and visualized the FIFA dataset as described above. The six views are:

- Views 1 and 3 are scatterplots showing player attributes. View 1 shows players’ attacking and defending skills. View 3 shows weekly income in US dollars and body height in centimeters.
- View 2 is a parallel coordinates view with seven skill attributes of soccer players. The parallel coordinates present a special case as a brush rectangle cannot be drawn freely within the view, but has to be drawn along one axis (see Figure 6).
- Views 4 and 6 are barcharts. View 4 shows the count of players within each club, while View 6 shows the count of players for each position on the field.
- View 5 is a text list showing the names of the 50 players.

We positioned our display (65” multi-touch SMART Board 6000 series with 3840×2160px resolution) in a quiet corner of a room. There was a three meters empty space in front of the screen to allow participants to freely move around. One camera was placed high in the opposite corner of the room to capture the dynamics and activities of the whole group. The second camera provided a side angle within two meters of the display to capture participants’ interactions with the touch screen. We also screen-captured the MyBrush interface. To clearly record voices, we installed a microphone on a shelf next to the display.

Procedure: Each focus group lasted approximately 60 minutes and consisted of four phases: prelude, training, exploration, and wrap-up.

Prelude (~10min). Participants were greeted, filled a consent form and a

demographic questionnaire covering experience with touch interaction, data analysis and visualization, and interest in sports and soccer.

Training (~10min). The experimenter demonstrated the capabilities of MyBrush through simple scenarios. Participants were encouraged to interrupt at any time, ask questions, and propose exploration tasks.

Exploration (~30min). The study transitioned from demonstration by the experimenter to active discussions and interactive exploration of the data by the participants. This exploration phase was the core of the study, where participants formulated their own data queries and interacted freely with MyBrush to find answers to their questions. If participants did not have their own questions, the experimenter suggested examples to encourage them to explore the data. The experimenter was available to answer questions and to give guidance if needed.

Wrap-up (~10min). We stopped the exploration after 50 minutes to keep the study within one hour and participants filled out post-study questionnaires to give feedback and received \$20 for their participation.

Data Analysis: For each focus group, we synchronized the video and audio streams and selected the speaking and interaction parts from the exploration phase. This resulted in a 20 minute long video for the *vis* group, 35 minutes for the *sports* group, and 30 minutes for the *mixed* group. We transcribed all discussions and described all observed actions on paper snippets. Led by the first author, we used affinity diagramming [8] to spatially analyze and cluster the observation snippets. This process was combined with frequent discussions among all authors to interpret affinity clusters and identify emerging themes.

5.2 Study Results

All three groups performed brushing and linking exploration tasks with different configurations for source, link, and target. All participants were strongly positive regarding their experience with MyBrush in the post-study questionnaire (see Figure 5). Although MyBrush “*had a bit of a learning curve*”(S2), they found that “*it was easy to learn*”(V4) and “*very fun*”(S2), that the brush menus were “*nicely done*”(V2) and “*very helpful*”(M1), and they liked that they could “*so easily connect this many views*”(V3). They were largely positive regarding the value of personal agency, e.g., “*I really like the many possibilities cause every person is gonna try differently*”(S1), which allowed them to answer complex data questions (e.g., “*I could learn about the players [...] and was able to search for answers to the questions I had*”(S1)). All but M1 stated that they would use such a tool to explore other data from “*video games*”(N4), “*other sports or business*”(M3), “*stock market*”(M4), and for “*data analysis in general*”(M1).

The sports group was notably more active than the other groups. We observed 22 scenarios (i.e., sequences of interactions that participants performed to achieve an exploration task) with the sports group, 7 with the *vis* group, and 9 with the *mixed* group. It was also apparent that the complexity of tasks had an influence on the type of interaction and the brush configurations that participants used. Below, we provide illustrative examples for simple use of available defaults, for small explorations that led to using little personal agency, and more complex examples where participants made considerable use of personal agency.

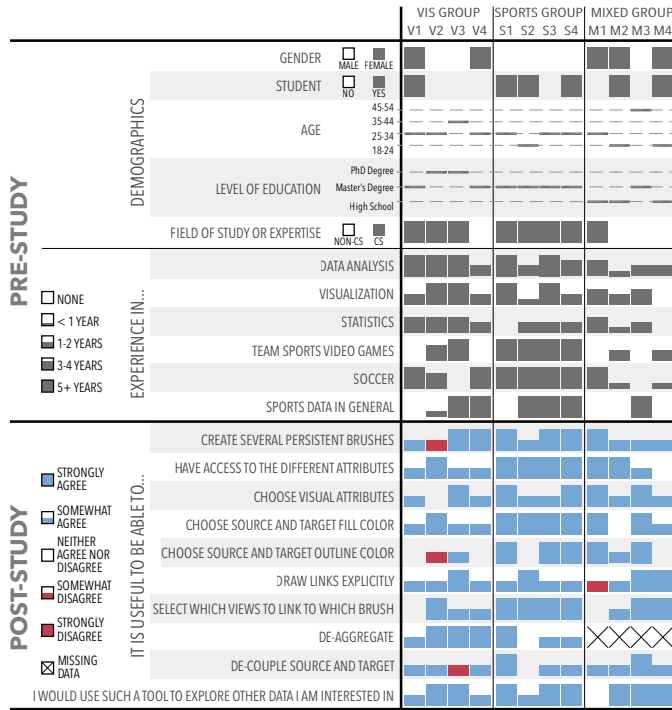


Fig. 5. Demographics of participants to the three focus groups, along with pre-study and post-study questionnaire answers.

5.2.1 Using the Brushing Defaults

The simplest brushing and linking interaction in MyBrush is accessed by tapping. Tapping a data point highlights it and the related data points in all other views. Participants used tapping to achieve simple tasks and answer simple data questions. It was the first interaction all groups used and was a way of discovering the interface and initiating their exploration. Tapping was sometimes used by itself, but was often included in a more complex scenario. When using tapping to explore *filtered outliers*, S3 highlighted the (de-aggregated) positions of the fastest players in View 6. In View 6, all wingers but one were highlighted, so S3 *tapped* the non-highlighted (slowest) winger, saying “*you didn’t know you wanted to find out about this outlier [before filtering and de-aggregating], but now you do!*”. *Obvious outliers* are data points that are spatially separated from the rest of the data points in the visualization, such as minimum or maximum data points. For example, V2 tapped in View 3 to find out “*who is making the most money?*”. Finally, *data points of personal interest* are players recognized or known by participants and therefore especially interesting to them. For example, M1 tapped the player M. Özil in the names list (View 5) after M2 had told the group that she knows and likes the player.

5.2.2 Moderately Simple Uses of Personal Agency

All groups created basic brushes with minor variations: 3 times for the sports group and once for the other two groups. These basic scenarios explored aspects of a single data subset, i.e. involved a single brush, and participants used a small subset of the available attributes for source, link, and target. They often used the source menu to choose a unique highlight color to avoid conflicts with other brushes on the screen, and the link menu simply to select views to be linked with the brush.

Extending the number of linked views. The sports group investigated players of Bayern Munich since S3 wanted to “*see who these highest paid players are on this team*”. By choosing a fill color and linking to View 3, S3 found Bayern Munich’s highest paid player: A. Robben. Next, to find out how Bayern Munich players perform in terms of attacking and defending skills, S3 decided to make them “*show up in this view (View 1)*”. To do this, S3 simply added View 1 to the same brush’s link menu “*to find out how their stats compare*”. Selective linking gave participants control over which views were linked and over the pace in which they were linked (one at a time or all together). This

externalized their activities, providing a means for developing common ground and raising an awareness of their activities, which is critical to successful collaboration [20]. It enabled the other group members to follow the exploration process and to contribute their own thoughts which often led to additional explorations. For instance, this finding from S3 intrigued S4 to further explore high-paid players.

Making Links Explicit. Research has shown that connection lines are as effective as color to link information across multiple views [29]. However, only the mixed group used connection lines between source and target. For example, M2 created two brushes in View 4 to compare the wages and attacking skills of the players from Arsenal and Manchester United. For each brush, she configured lines with different curvatures and colors. Then, on realizing that “*Arsenal is just connected to View 3, but the other one is connected to two views*”, M2 updated the brush on Arsenal to add the other view as linked view. The mixed group used the lines as additional visual guidance to be aware of which views are linked to which brush and to “*track back to the other view*” (N1). They even asked whether they could also draw lines between target views of a brush to make their relationship more salient. In contrast, the sports enthusiasts said lines caused visual clutter (“*there are too many lines. I can’t tell the individual connections*” (S2)). The vis group did not use lines during the exploration phase at all.

De-aggregating Targets. Both the sports group and the vis group used de-aggregation of target to explore the distribution of data attributes. For example, the sports group investigated the positions of fast players. S3 created a brush highlighting the best players in the movement axis of View 2 and linked it to the positions in View 6: all positions were highlighted except for the goalkeepers. S3 wondered if he selected too much data with the brush in View 2, but, after de-aggregating, said “*we immediately see that wingers are fast*” as most of the wingers and only a few players in other positions were highlighted. The sports group used de-aggregation in subsequent scenarios that involved aggregated views, e.g. to examine the positions of average-height players. The vis group used de-aggregation several times in their exploration. In fact, three of the four people in the vis group noted the “*idea of de-aggregating bar charts*” as a favourite feature (see Figure 5–“de-aggregate”).

5.2.3 Increasing Power through Personal Agency

As the data questions became more complicated, participants made increasing use of personal agency and of multiple brushes.

Drilling Down with Consecutive Brushes. When participants found something interesting, they often drilled down further, performing a series of exploration tasks in sequence, with each new task being motivated by the previous findings, leaving behind an exploration trail of related brushes. For example, the mixed group created a brush in View 1 selecting the three goalkeepers of the dataset and linked View 2. On observing that goalkeepers have a weak “*mentality*”, M1 decided to find out the positions of players with strong mentality by brushing them in View 2 and linking to View 6 to show the de-aggregated positions. Persistent brushes have the advantage over transient brushes in that they can be modified and updated over time. We found, however, transient and persistent brushes are complementary. Transient brushes (tapping on data points) are useful for quick interactions such as exploring outliers, and persistent brushes to answer more detailed data questions and inspire follow-up exploration.

Complex Intersecting Brushes. Persistent brushes were mostly used for compound brushing. Participants combined multiple brushes with different configurations either to compare two or more subsets of players regarding specific attribute distributions, or to combine two brushes to visualize the intersection of different subsets of players.

Figure 6 shows an exploration scenario from the sports group that combines brushes to compare and intersect subsets of players. After creating a brush in View 4 to highlight Chelsea’s players, S2 wondered which player could replace Fabregas if he left Chelsea? S2 limited the search to players with the same position as Fabregas by brushing the central midfield (CM) bar in View 6, saying “*That’s one criteria.*”. He linked this brush to the skills plot in View 2 to analyze the skills of CM players, setting source and target to green. Then, S3 helped S2 by tapping Fabregas to compare his skills to the skills of other CM players

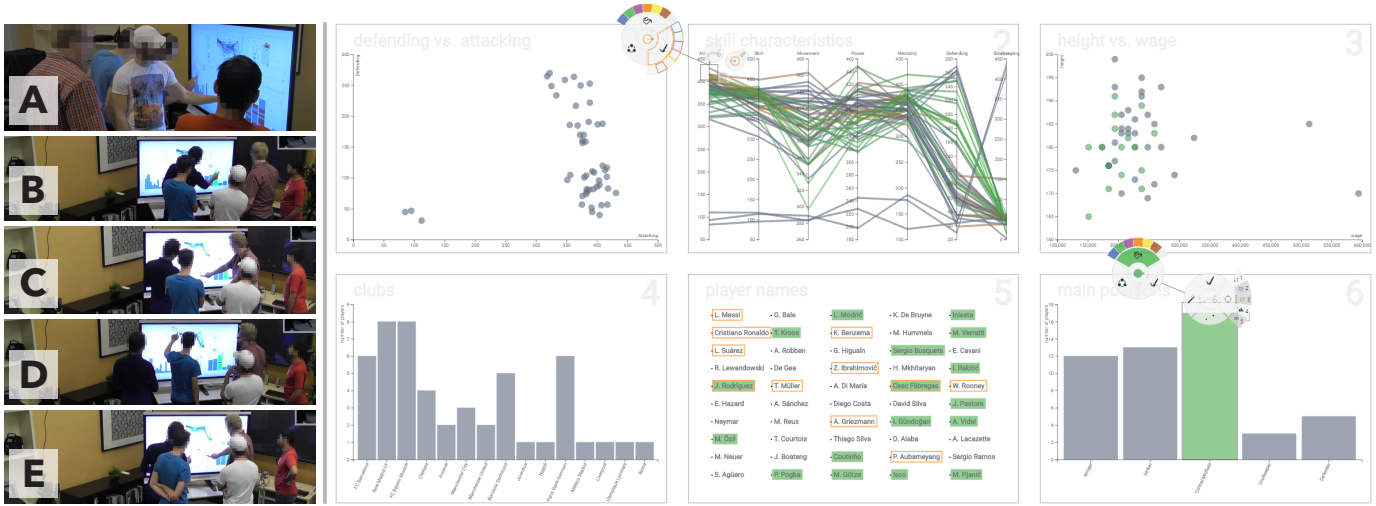


Fig. 6. (A) In the back, S2 creates a brush on “Chelsea” in View 4 to highlight Chelsea’s players and taps on “Fabregas” in View 5. (B) S2 brushes the central midfield players in View 6. He configures it with a green fill color for both source and target, and links View 2 to the brush in order to analyze central midfield skills. (C) S3 helps S2 by tapping Fabregas to highlight the player in View 2. (D) After the experimenter explains how to create brush intersections, participants show the names of central midfield players with a green fill color, and those of players with high attacking skills with an orange outline (as shown in the screenshot). (E) S2 and S3 simultaneously create two brushes, one for each of the two players that have a green fill color and an orange container: Fabregas and Rodriguez. They agree that Rodriguez is the best substitute for Fabregas.

in View 2. Since “wage is also important”, S2 linked the brush to View 3. S3 then tapped the CM player closest to Fabregas’ wage in View 3 to highlight his name in View 5: Iniesta. However, based on their knowledge and personal preferences, they did not consider him to be a good match. Then they created a new brush in View 2 to select players with similar attacking skills as Fabregas. They used non-conflicting visual attributes to show the names of CM players with green fill, and the players with high attacking skills with an orange outline (see Figure 6, View 5). They realized that only two players were highlighted with both green fill and orange outline: Fabregas and Rodriguez. S2 and S3 simultaneously created two brushes in View 5 to highlight Rodriguez and Fabregas across all views. The group discussed and agreed that Rodriguez is the best substitute for Fabregas: he has “same attacking skill”(S2), “same price”(S1), “better movement”(S2). Rodriguez “is worth it actually, compared to Fabregas”(S2) who is slower, “by a lot actually”(S3). “Yeah, so get Rodriguez!”(S1).

The vis group also combined brushes to compare the positions across views. They created four brushes in the positions barchart (View 6) and configured each with a unique source fill color, linked it to all other views, and de-aggregated the target for the clubs barchart in View 4. The mixed group used two brushes to compare player wages of Arsenal and Manchester United as described earlier. With a little hint by the experimenter, they created a separate brush for both clubs and found out that the Arsenal players earn more than the Manchester players.

While the sports group posed the most complex data questions (inspired by practical concerns that soccer clubs face, such as looking for a replacement player), all groups combined multiple brushes to solve advanced exploration tasks — and all participants increasingly leveraged personal agency as their tasks gained in complexity.

6 DISCUSSION

We discuss our findings about the role of personal agency when exploring interactive visualizations, and describe challenges and opportunities in designing for and evaluating personal agency.

6.1 The Case for Personal Agency

To provide direct access to brush components (DG1), MyBrush features expandable, contextual menus. People configured these components in their own ways (personal agency). For example, through the simple action of controlling the rate at which selected views were linked to a brush, the sports group made the exploration of player statistics a group conversation. Choosing the rate of information disclosure, is an example of using personal agency.

To offer choice in degree of personal agency (DG2), MyBrush provides reasonable defaults. All groups used the default tapping interaction for their first queries, showing that not being required to fully form a brush before first use made getting started easier. While the sports group noted the clutter of visual links, the mixed group, who had less experience with vis or the data, made use of visual links and appreciated the added clarity. Being able to choose the degree of personal agency, participants could learn by doing. This supported gentle entry and eased transitioning to complex use while still feeling in control.

Our focus group studies demonstrated that MyBrush supports complex personal agency (DG3). The sports group’s quest to find a replacement player for Fabregas is an example of a more complex query than we had envisioned that was put together relatively easily. We observed personal agency unfolding as participants became more familiar and started asking deeper questions. We watched different groups decide on different preferences (e.g., implicit and explicit links) that better supported the group’s needs, background and experience. The vis group, who mastered traditional brushing and linking were faster than the other groups to understand the mechanisms of MyBrush. The sports group, deeply interested in the data, created the most brushes among all groups to answer the many data questions that they had. They were able to use MyBrush to design a complex set of brushes to successfully answer their complex query about player replacement. The mixed group, finally, was able to design brushes that worked for the people with less visualization and data experience.

6.2 Designing for Personal Agency

In the context of designing for personal agency, we discuss scalability, sensible defaults, personal agency in real-world scenarios, conflict resolution, and support for personal agency in collaborative interfaces.

Scalability. There are well known scalability issues in terms of data size. MyBrush suffers from the same limitations as any brushing and linking techniques with large datasets (e.g., overplotting and difficulty of selecting data points). Some visual attributes such as making links explicit become more limited in use as the number of data points increases. However, basing a brushing and linking interface on personal agency makes it intrinsically modular, making it possible to implement techniques dedicated to these scalability issues. For example, in the case of visual links for large data sets, options for link bundling [41] and routing [70] could be added as choosable visual attributes. A promising avenue for future work is to investigate how to provide people with personal agency over highly configurable attributes such as bundling and routing algorithms, and over scalability issues in general.

MyBrush is designed to support the interactive configuration of any of the attributes in Table 1. While one of the vis experts said that “*there are already so many cool features*” (V1), participants also suggested adding more configurable attributes: visual attributes such as “*textures in addition to color*” (V3) and “*a way to hide elements from a view*” (S3), and more complex ones such as “*other shapes for selection besides rectangles*” (V3). As adding more features raises interface scalability issues, one may look into ways of increasing the number of items radial menus can contain (e.g., [30, 40]). However, we need to balance the number of features with the complexity of the interface, occlusion of the brush and data points, and discoverability of the interface functionality. Clutter is another challenge. While we envisioned people would delete brushes they did not need, we noticed that participants only removed old brushes when they were in their way and when they wanted to reclaim display real estate. Finally, we designed MyBrush with large screens in collaborative settings in mind (e.g., smartboards, walls, and tables). An interesting challenge is to expand our approach to mobile devices, applying insights from existing mobile interactions with CMV [68].

Sensible defaults. One of the vis experts suggested “*have default selecting behaviour (highlighting) for drawing rectangles*” (V4), i.e., to link all views to any brush by default and reduce configuration overhead. However, we observed that selective linking supported collaborative engagement through increased activity awareness. Four participants (M1, M3, S1, V2) asked about the order of aggregated items in bars, such as M1 complaining that there was “*no ordering of items in the bar chart/bars/player names*”. This interest showed that participants found it useful to be able to drill down into aggregated views on demand through brushing and linking. It also raises interesting challenges for future work, such as how to make the ordering of items in a de-aggregated view explicit and transparent, and how to provide personal agency over the order of items in a de-aggregated view. A future direction is to explore sensible default behaviours and their role in the experience of interactive visualization with personal agency.

Personal agency in real-world scenarios. There is likely a potential trade-off between accomplishing tasks efficiently and having a sense of agency and control over one’s actions in the interface. While these do not need to be at odds, one may wonder whether people would be willing to spend the time and effort to configure brushes when dealing with different data sets in different situations, for example, when under time constraints. Although we did not study the efficiency of MyBrush in different scenarios, our findings provide some initial insights that can serve as directions to explore in future deployments and studies. Eleven out of twelve participants stated that they would use a tool like MyBrush to explore other data sets. This suggests that MyBrush was perceived to be useful for some real-world situations. Participants projected themselves into real-world scenarios as they made suggestions of features that they would like to see implemented, such as detail-on-demand to “*make a selection and view a data label*” (V4) and “*[a] way to see a fuller info card on each player*” (M3), and expert shortcuts such as “*selecting multiple [views] at once*” (M2).

Conflict resolution. While MyBrush supports combinations of brushes by using several visual attributes, it does not address configuration conflicts. A configuration conflict occurs when two brushes overwrite the same attribute of a data point (e.g., fill color blue by brush A and fill color red by brush B). A significant future challenge is to consider how to handle these conflicts without interfering with personal agency. A simple solution to this challenge would be to highlight conflicts so that people could correct their own configurations. However, there are other promising alternatives. One of the vis people noticed this issue, and suggested using “*color striping or similar for conflicting colors*” (V3) to show both conflicting fill colors simultaneously, similarly to the blending of conflicting colors used in WEAVE [28].

Support for personal agency in collaborative interfaces. Participants appreciated the collaborative possibilities of MyBrush, e.g., “*I really liked the multi-touch so that multiple people could work at the same time*” (V3). A deeper look into design considerations for collaborative visual analytics brings design challenges such as making brushes shareable, either synchronously or asynchronously [38]. In the real world, collaboration is characteristic of most visual analytics

work [61] and commonly involves people with different backgrounds and levels of expertise, such as scientists, analysts, and policy makers. Our observations suggest that personal agency in MyBrush encouraged collaboration. Explicit visibility of completed and ongoing brush configurations provided activity awareness within the groups, helped participants to externalize their activities, and provided a means for developing a common ground about the analysis tasks at hand. All these aspects are critical to successful collaboration [20, 31]. Moreover, we found that personal agency supported the needs of people with different backgrounds and levels of expertise. Participants increasingly leveraged personal agency as their tasks increased in complexity, delving into more complex functionality of the interface as they gained expertise, suggesting that personal agency may facilitate acquiring expertise gradually, which may prove beneficial in real-world analysis tasks.

6.3 Evaluating Personal Agency in Visualizations

The participants’ verbal comments during and after interacting with MyBrush reveal a high level of personal agency. Their discussions on the soccer domain while exploring the views presented by MyBrush indicate that they understood the effects of their actions and that the system’s behavior appeared transparent to them even with increasing task complexity – they felt in control. Participants’ statements also demonstrate that they felt as initiators of an action (as opposed to the system). For example, S2 referred explicitly to his own exploration goals rather than interface functionality: “*I could find an alternative to a player of my favourite club who might leave the team*” (S2).

While our study made it possible to observe personal agency in a pseudo real-world scenario, in the future we need to establish strategies for evaluating to which degree participants make use of personal agency to achieve their goals. One possibility is to look into linguistic analysis to compare the frequency of statements referring to people as actors (“I did”) with the frequency of statements referring to the system as an actor (“The system did”). Another possibility is to evaluate whether people have a temporal bias towards systems supporting personal agency. According to neuroscientists [54], agency is reflected by a temporal bias with the time interval between the expressed intent and the resulting effect appearing shorter to people. Some of our participants’ statements hint at this (e.g., “*updates work quickly*” (V3)) but controlled quantitative studies are needed to investigate this question.

7 CONCLUSIONS

Letting people configure the brushing and linking interaction by direct manipulation of its source, link, and target attributes is a way of opening up the usually closed box that contains the interaction mechanisms. Providing people with the opportunity to unwrap and *personalize* interactions is a promising approach to designing more transparent and more engaging interactive visualization tools — as one sports enthusiast said, “*every person is gonna try [things] differently*” (S1). Making the effects of a configuration instantly visible allows for rapid and incremental actions. By providing direct feedback and making all options explicitly visible in the source, link, and target menus, MyBrush supports discoverability and visibility [56] and encourages people to explore the space of possibilities. While this approach likely has a cost in terms of people’s efficiency initially, we are confident that the benefits – in terms of discoverability, learnability, transferability, personalization, agency, and sense of control – are much higher, and that the cost in terms of efficiency will diminish as people learn to use the interface.

SUPPLEMENTAL MATERIALS

See <http://innovis.cpsc.ualgary.ca/supplemental/MyBrush/> for supplemental material, including a demo and the source code of MyBrush.

ACKNOWLEDGMENTS

We thank the reviewers for their thoughtful comments and Richard Pusch for proofreading the paper. This research was supported in part by: Alberta Innovates - Technology Futures (AITF); Natural Sciences and Engineering Research Council of Canada (NSERC); SMART Technologies; and German Academic Exchange Service (DAAD). We used icons from Freepik on flaticon.com.

REFERENCES

- [1] EA Sports FIFA Soccer Franchise Sales Top 100 Million Units Lifetime, Nov. 2010. <http://www.businesswire.com/news/home/20101104006782/en>, Last accessed on June 21, 2017.
- [2] Sofifa.com traffic (w3snoop.com), June 2017. <http://sofifa.com.w3snoop.com/>, Last accessed on June 21, 2017.
- [3] J. Accot and S. Zhai. More Than Dotting the I's — Foundations for Crossing-based Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, pages 73–80, New York, NY, USA, 2002. ACM.
- [4] C. Ahlberg. Spotfire: An Information Exploration Environment. *SIGMOD Rec.*, 25(4):25–29, Dec. 1996.
- [5] Z. Barlas and S. Obhi. Freedom, choice, and the sense of agency. *Frontiers in Human Neuroscience*, 7:514, 2013.
- [6] R. A. Becker and W. S. Cleveland. Brushing Scatterplots. *Technometrics*, 29(2):127–142, May 1987.
- [7] J. Bertin. *Semiology of Graphics*. University of Wisconsin Press, Wisconsin, US, 1983.
- [8] H. Beyer and K. Holtzblatt. *Contextual Design*. Morgan Kaufmann Publisher, Inc., San Francisco, 1998.
- [9] M. Bostock, V. Ogievetsky, and J. Heer. D3 Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, Dec. 2011.
- [10] N. Boukhelifa and P. J. Rodgers. A Model and Software System for Coordinated and Multiple Views in Exploratory Visualization. *Information Visualization*, 2(4):258–269, Dec. 2003.
- [11] D. Brodbeck and L. Girardin. Design Study: Using Multiple Coordinated Views to Analyze Geo-referenced High-dimensional Datasets. In *Proceedings of the Conference on Coordinated and Multiple Views In Exploratory Visualization*, CMV '03, pages 104–111, Washington, DC, USA, 2003. IEEE Computer Society.
- [12] A. Buja, J. A. McDonald, J. Michalak, and W. Stuetzle. Interactive Data Visualization Using Focusing and Linking. In *Proceedings of the 2Nd Conference on Visualization '91*, VIS '91, pages 156–163, Los Alamitos, CA, USA, 1991. IEEE Computer Society Press.
- [13] H. Chen. Compound Brushing. In *Proceedings of the Ninth Annual IEEE Conference on Information Visualization*, INFOVIS'03, pages 181–188, Washington, DC, USA, 2003. IEEE Computer Society.
- [14] J. H. T. Claessen and J. J. van Wijk. Flexible Linked Axes for Multivariate Data Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2310–2316, Dec. 2011.
- [15] C. Collins and S. Carpendale. VisLink: Revealing Relationships Amongst Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1192–1199, Nov. 2007.
- [16] C. Collins, F. B. Viegas, and M. Wattenberg. Parallel Tag Clouds to explore and analyze faceted text corpora. In *2009 IEEE Symposium on Visual Analytics Science and Technology*, pages 91–98, Oct 2009.
- [17] D. Coyle, J. Moore, P. O. Kristensson, P. Fletcher, and A. Blackwell. I did that! Measuring users' experience of agency in their own actions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 2025–2034, New York, NY, USA, 2012. ACM.
- [18] H. Doleisch, M. Gasser, and H. Hauser. Interactive Feature Specification for Focus+Context Visualization of Complex Simulation Data. In *Proceedings of the Symposium on Data Visualisation 2003*, VISSYM '03, pages 239–248, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.
- [19] H. Doleisch and H. Hauser. Smooth Brushing for Focus+Context Visualization of Simulation Data in 3D. In *The 10-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision '2002*, WSCG 2002, pages 147–154, 2002.
- [20] P. Dourish and V. Bellotti. Awareness and Coordination in Shared Workspaces. In *Proceedings of the 1992 ACM Conference on Computer-supported Cooperative Work*, CSCW '92, pages 107–114, New York, NY, USA, 1992. ACM.
- [21] C. Dunne, N. Henry Riche, B. Lee, R. Metoyer, and G. Robertson. GraphTrail: Analyzing Large Multivariate, Heterogeneous Networks While Supporting Exploration History. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1663–1672, New York, NY, USA, 2012. ACM.
- [22] N. Elmqvist, P. Dragicevic, and J.-D. Fekete. Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1539–1148, Nov 2008.
- [23] J. K. Esser. Alive and Well after 25 Years: A Review of Groupthink Research. *Organizational Behavior and Human Decision Processes*, 73(2):116 – 141, 1998.
- [24] J.-D. Fekete. The Infovis Toolkit. In *Proceedings of the IEEE Symposium on Information Visualization*, INFOVIS '04, pages 167–174, Washington, DC, USA, 2004. IEEE Computer Society.
- [25] Y. Fua, M. O. Ward, and E. A. Rundensteiner. Navigating Hierarchies with Structure-Based Brushes. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'99)*, pages 58–64, 1999.
- [26] T. Geymayer, M. Steinberger, A. Lex, M. Streit, and D. Schmalstieg. Show Me the Invisible: Visualizing Hidden Content. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 3705–3714, New York, NY, USA, 2014. ACM.
- [27] S. Gratzl, N. Gehlenborg, A. Lex, H. Pfister, and M. Streit. Domino: Extracting, Comparing, and Manipulating Subsets Across Multiple Tabular Datasets. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2023–2032, Dec 2014.
- [28] D. L. Gresh, B. E. Rogowitz, R. L. Winslow, D. F. Scollan, and C. K. Yung. WEAVE: A System for Visually Linking 3-D and Statistical Visualizations, Applied to Cardiac Simulation and Measurement Data. In *Proceedings of the Conference on Visualization '00*, VIS '00, pages 489–492, Los Alamitos, CA, USA, 2000. IEEE Computer Society Press.
- [29] A. L. Griffin and A. C. Robinson. Comparing Color and Leader Line Highlighting Strategies in Coordinated View Geovisualizations. *IEEE Transactions on Visualization and Computer Graphics*, 21(3):339–349, March 2015.
- [30] F. Guimbretière and T. Winograd. FlowMenu: Combining Command, Text, and Data Entry. In *UIST '00*, pages 213–216, New York, NY, USA, 2000. ACM.
- [31] C. Gutwin and S. Greenberg. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Comput. Supported Coop. Work*, 11(3):411–446, Nov. 2002.
- [32] P. Haggard and M. Tsakiris. The Experience of Agency. *Current Directions in Psychological Science*, 18(4):242–246, 2009.
- [33] K. W. Hall, C. Perin, P. G. Kusalik, C. Gutwin, and S. Carpendale. Formalizing Emphasis in Information Visualization. *Computer Graphics Forum*, 2016.
- [34] F. Hardisty and A. C. Robinson. The Geoviz Toolkit: Using Component-oriented Coordination Methods for Geographic Visualization and Analysis. *Int. J. Geogr. Inf. Sci.*, 25(2):191–210, Feb. 2011.
- [35] H. Hauser. *Generalizing Focus+Context Visualization*, pages 305–327. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.
- [36] H. Hauser, F. Ledermann, and H. Doleisch. Angular Brushing of Extended Parallel Coordinates. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'02)*, INFOVIS '02, pages 127–130, Washington, DC, USA, 2002. IEEE Computer Society.
- [37] M. A. Hearst. User Interfaces and Visualization. In B. Ribeiro-Neto, editor, *Modern Information Retrieval*, chapter 10, page 260. Addison Wesley Longman, 1999.
- [38] J. Heer and M. Agrawala. Design Considerations for Collaborative Visual Analytics. *Information Visualization*, 7(1):49–62, Mar. 2008.
- [39] J. Heer, S. K. Card, and J. A. Landay. Preface: A Toolkit for Interactive Information Visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 421–430, New York, NY, USA, 2005. ACM.
- [40] T. Hesselmann, S. Flöring, and M. Schmitt. Stacked Half-Pie Menus: Navigating Nested Menus on Interactive Tabletops. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '09, pages 173–180, New York, NY, USA, 2009. ACM.
- [41] D. Holten. Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):741–748, Sept. 2006.
- [42] H. Jänicke, M. Böttinger, and G. Scheuermann. Brushing of Attribute Clouds for the Visualization of Multivariate Data. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1459–1466, Nov. 2008.
- [43] W. Javed and N. Elmqvist. Exploring the Design Space of Composite Visualization. In *Proceedings of the 2012 IEEE Pacific Visualization Symposium*, PACIFICVIS '12, pages 1–8, Washington, DC, USA, 2012. IEEE Computer Society.
- [44] D. A. Keim. Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, Jan. 2002.
- [45] S. Knudsen and S. Carpendale. View Relations: An Exploratory Study

- on Between-View Meta-Visualizations. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, NordiCHI '16, pages 15:1–15:10, New York, NY, USA, 2016. ACM.
- [46] Z. Konyha, K. Matkovic, D. Gracanin, M. Jelovic, and H. Hauser. Interactive Visual Analysis of Families of Function Graphs. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1373–1385, Nov. 2006.
- [47] A. Lex, C. Partl, D. Kalkofen, M. Streit, S. Gratzl, A. M. Wassermann, D. Schmalstieg, and H. Pfister. Entourage: Visualizing Relationships Between Biological Pathways Using Contextual Subsets. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2536–2545, Dec. 2013.
- [48] A. Lex, M. Streit, E. Kruijff, and D. Schmalstieg. Caleydo: Design and evaluation of a visual analysis framework for gene expression data in its biological context. In *2010 IEEE Pacific Visualization Symposium (PacificVis)*, pages 57–64, March 2010.
- [49] A. Lex, M. Streit, C. Partl, and D. Schmalstieg. Comparative Analysis of Multidimensional, Quantitative Data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1027–1035, Nov. 2010.
- [50] C. Marshall and G. Rossman. *Designing Qualitative Research*. SAGE Publications, 1999.
- [51] A. R. Martin and M. O. Ward. High Dimensional Brushing for Interactive Exploration of Multivariate Data. In *Proceedings of the 6th Conference on Visualization '95*, VIS '95, pages 271–278, Washington, DC, USA, 1995. IEEE Computer Society.
- [52] J. McDonald. *Interactive Graphics for Data Analysis*. PhD thesis, Calif. Univ. Stanford, Aug 1982.
- [53] M. Monmonier. Geographic Brushing: Enhancing Exploratory Analysis of the Scatterplot Matrix. *Geographical Analysis*, 21(1):81–84, 1989.
- [54] J. W. Moore. What Is the Sense of Agency and Why Does it Matter? *Frontiers in Psychology*, 7:1272, 2016.
- [55] T. Munzner, F. Guimbreti re, S. Tasiran, L. Zhang, and Y. Zhou. TreeJuxtaposer: Scalable Tree Comparison Using Focus+Context with Guaranteed Visibility. *ACM Trans. Graph.*, 22(3):453–462, July 2003.
- [56] D. A. Norman. *The Design of Everyday Things: Revised and Expanded Edition*. Basic Books, 2013.
- [57] C. North and B. Shneiderman. Snap-together Visualization: A User Interface for Coordinating Visualizations via Relational Schemata. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '00, pages 128–135, New York, NY, USA, 2000. ACM.
- [58] K. Pantazos and S. Lauesen. Constructing Visualizations with InfoVis Tools - An Evaluation from a user Perspective. In *GRAPP & IVAPP 2012: Proceedings of the International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization Theory and Applications*, Rome, Italy, 24-26 February, 2012, pages 731–736, 2012.
- [59] S. Parkin. Fifa: the video game that changed football, Dec. 2016. <https://www.theguardian.com/technology/2016/dec/21/fifa-video-game-changed-football>, Last accessed on June 21, 2017.
- [60] C. Perin, P. Dragicevic, and J. Fekete. Revisiting Bertin matrices: New Interactions for Crafting Tabular Visualizations. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):2082–2091, 2014.
- [61] W. A. Pike, J. Stasko, R. Chang, and T. A. O'Connell. The Science of Interaction. *Information Visualization*, 8(4):263–274, Dec. 2009.
- [62] R. M. Pillat and C. M. D. S. Freitas. Coordinating Views in the InfoVis Toolkit. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '06, pages 496–499, New York, NY, USA, 2006. ACM.
- [63] S. Rado , R. Splechna, K. Matkovi , M. Duras, E. Gr ller, and H. Hauser. Towards Quantitative Visual Analytics with Structured Brushing and Linked Statistics. *Computer Graphics Forum*, 35(3):251–260, 2016.
- [64] J. C. Roberts. On Encouraging Multiple Views for Visualisation. In *Proceedings of the International Conference on Information Visualisation*, IV '98, pages 8–14, Washington, DC, USA, 1998. IEEE Computer Society.
- [65] J. C. Roberts. State of the Art: Coordinated & Multiple Views in Exploratory Visualization. In *Proceedings of the Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization*, CMV '07, pages 61–71, Washington, DC, USA, 2007. IEEE Computer Society.
- [66] J. C. Roberts and M. A. E. Wright. Towards Ubiquitous Brushing for Information Visualization. In *10th International Conference on Information Visualisation*, IV 2006, pages 151–156, 2006.
- [67] A. C. Robinson. Highlighting in Geovisualization. *Cartography and Geographic Information Science*, 38(4):373–383, 2011.
- [68] R. Sadana and J. Stasko. Designing Multiple Coordinated Visualizations for Tablets. *Computer Graphics Forum*, 35(3):261–270, 2016.
- [69] B. Shneiderman and C. Plaisant. *Designing the User Interface: Strategies for Effective Human-Computer Interaction (4th Edition)*. Pearson Addison Wesley, 2004.
- [70] M. Steinberger, M. Waldner, M. Streit, A. Lex, and D. Schmalstieg. Context-Preserving Visual Links. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2249–2258, Dec. 2011.
- [71] D. F. Swayne, D. Cook, and A. Buja. XGobi: Interactive dynamic data visualization in the X Window System. *Journal of Computational and Graphical Statistics*, 7:113–130, 1998.
- [72] Tableau. Tableau Public, 2017. <https://public.tableau.com/s/>, Last accessed on March 29, 2017.
- [73] M. Theus. Interactive Data Visualization using Mondrian. *Journal of Statistical Software*, 7(11):1–12, 2002.
- [74] M. Tobiasz, P. Isenberg, and S. Carpendale. Lark: Coordinating Co-located Collaboration with Information Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1065–1072, Nov 2009.
- [75] M. E. Turner and A. R. Pratkanis. Twenty-Five Years of Groupthink Theory and Research: Lessons from the Evaluation of a Theory. *Organizational Behavior and Human Decision Processes*, 73(2):105 – 115, 1998.
- [76] C. Viau and M. J. McGuffin. ConnectedCharts: Explicit Visualization of Relationships Between Data Graphics. *Comput. Graph. Forum*, 31(3pt4):1285–1294, June 2012.
- [77] M. Waldner, W. Puff, A. Lex, M. Streit, and D. Schmalstieg. Visual Links Across Applications. In *Proceedings of Graphics Interface 2010*, GI '10, pages 129–136, Toronto, Ont., Canada, Canada, 2010. Canadian Information Processing Society.
- [78] M. Q. Wang Baldonado, A. Woodruff, and A. Kuchinsky. Guidelines for Using Multiple Views in Information Visualization. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '00, pages 110–119, New York, NY, USA, 2000. ACM.
- [79] M. O. Ward. XmdvTool: Integrating Multiple Methods for Visualizing Multivariate Data. In *Proceedings of the Conference on Visualization '94*, VIS '94, pages 326–333, Los Alamitos, CA, USA, 1994. IEEE Computer Society Press.
- [80] C. Weaver. Building Highly-Coordinated Visualizations in Improvise. In *Proceedings of the IEEE Symposium on Information Visualization*, INFOVIS '04, pages 159–166, Washington, DC, USA, 2004. IEEE Computer Society.
- [81] G. J. Wills. Selection: 524,288 Ways to Say "This is Interesting". In *Proceedings of the 1996 IEEE Symposium on Information Visualization (INFOVIS '96)*, INFOVIS '96, pages 54–60, Washington, DC, USA, 1996. IEEE Computer Society.
- [82] P. C. Wong and R. D. Bergeron. Multiresolution Multidimensional Wavelet Brushing. In *Proceedings of the 7th Conference on Visualization '96*, VIS '96, pages 141–148, Los Alamitos, CA, USA, 1996. IEEE Computer Society Press.