The Anchoring Effect in Decision-Making with Visual Analytics

Isaac Cho* Ryan Wesslen† Alireza Karduni‡ Sashank Santhanam§ Samira Shaikh¶
Wenwen Dou‖

University of North Carolina at Charlotte

Abstract

Anchoring effect is the tendency to focus too heavily on one piece of information when making decisions. In this paper, we present a novel, systematic study and resulting analyses that investigate the effects of anchoring effect on human decision-making using visual analytic systems. Visual analytics interfaces typically contain multiple views that present various aspects of information such as spatial, temporal, and categorical. These views are designed to present complex, heterogeneous data in accessible forms that aid decision-making. However, human decision-making is often hindered by the use of heuristics, or cognitive biases, such as anchoring effect. Anchoring effect can be triggered by the order in which information is presented or the magnitude of information presented. Through carefully designed laboratory experiments, we present evidence of anchoring effect in analysis with visual analytics interfaces when users are primed by representation of different pieces of information. We also describe detailed analyses of users' interaction logs which reveal the impact of anchoring bias on the visual representation preferred and paths of analysis. We discuss implications for future research to possibly detect and alleviate anchoring bias.

Keywords: Visual Analytics, Anchoring Effect, Sense Making, Cognitive Bias, Interaction Log Analysis

Index Terms: K.6.1 [Management of Computing and Information Systems]: Project and People Management—Life Cycle; K.7.m [The Computing Profession]: Miscellaneous—Ethics

1 Introduction

Researchers in multiple fields, including psychology, economics and medicine have extensively studied the effect of cognitive biases on decision making [7, 8, 17]. Cognitive biases are rules of thumb or heuristics that help us make sense of the world and reach decisions with relative speed [28]. Decision making, the process of identifying solutions to complex problems by evaluating multiple alternatives [46] has been increasingly exacerbated due to explosion of big data [44]. To facilitate human decision-making processes on large and complex datasets, Visual Analytics (VA) combines automated analysis techniques with interactive visualizations to increase the amount of data users can effectively work with [31]. Evidently, the effectiveness of VA to support decision making is an area that warrants study. Our goal in this work is therefore to conduct a study which incorporates three complementary strands of research, given the premises that VA supports decision making, and that decision making is impacted by cognitive biases. Specifically, we investigate how users' decision making processes are impacted by cognitive biases when using VA systems to analyze large and complex datasets. Moreover, we explore if and how cognitive biases are reflected in the way that users interact with visual analytic interfaces.

In the context of VA research, many recent VA systems [9, 16, 21, 30, 34] designed to facilitate the decision making on large and complex datasets contain coordinated and multiple views (CMV). By presenting different visual representations that show various aspects of the underlying data and automatically coordinating operations between views, multiple coordinated views support exploratory analysis to enable insight and knowledge discovery [39]. In visual interfaces that employ CMV design, users often have choices on which views serve as primary vs. supporting views for their analysis and on the strategies to switch between different views.

The flexibility of visual interfaces with coordinated and multiple views make cognitive biases such as anchoring bias particularly relevant to study. People find cognitive biases to be useful heuristics when sorting through large amounts of information, when task constraints or instructions prime them to focus on specific types of information, or when asked to make quick decisions and analyses. This has been demonstrated for several biases and shown that biases affect decision-making processes in predictably faulty ways that can result in decision-making failures when information is discounted, misinterpreted, or ignored [29]. Additionally, the biases affect not only regular users, but also expert users, when thinking intuitively [29]. One type of bias, the anchoring effect describes the human tendency to rely too heavily on one/the first piece of information offered (the "anchor") when making decisions [17]. Research has demonstrated that individuals anchor on a readily accessible value and adjust from it to estimate the true value, often with insufficient adjustments. For instance, if a person is asked to estimate the length of the Mississippi River, following a question on whether the length is longer or shorter than 500 miles, their answer will be adjusted from the 'anchor' value of 500 miles and will underestimate the true length of the Mississippi River. The effect of such anchors have been extensively studied in multiple tasks in the laboratory and in the field (for a detailed review see [19]). However, the effect of anchoring in Visual Analytics interfaces have not been systematically studied. More importantly, the effect of anchoring bias on the strategies that users deploy to interact with the visual interface and their analysis outcomes remains an open question.

In this paper, we study the effect of anchoring on users exploration processes and outcomes. When interacting with visual interfaces employing CMV design, there is a possibility that users rely too heavily on one particular view. The reasons for such reliance include but are not limited to prior experience, familiarity with certain visualizations, and different ways they were trained to use the visual interface. The significance and impact of such anchoring is the subject of our study.

Prior work in the VA community provides empirical data on cognitive costs of visual comparisons and context switching in coordinated-multiple-view visual interfaces [11, 38]. Findings from these experiments inform design guidelines of CMVs. However, there is little research on how cognitive biases transfer to visualizations, in particular to visual interfaces with coordinated multiple views. MacEachren [33] argues that prior efforts in visualization of uncertainty deal with representation of data uncertainty, but do not

*e-mail:icho1@uncc.edu
†e-mail:rwesslen@uncc.edu
‡e-mail:akarduni@uncc.edu
§e-mail:ssantha1@uncc.edu
¶e-mail:sshaikh2@uncc.edu
‖e-mail:wdou1@uncc.edu

address the reasoning that takes place under these conditions. We therefore aim to investigate the impact of anchoring effect on human decision-making processes when using VA systems, because it has been shown to be overwhelmingly affect decision-making [17]. Our experiment design addresses several challenging requirements that are necessary to derive meaningful implications: first, the experiments need to be conducted using a VA system with tasks relevant to decision-making based on large and complex datasets; second, measures and experiment data that reflect users' decision making processes (beyond task completion time and accuracy) need to be collected; third, novel analyses methods need to be developed to tease out the effect of anchoring bias on decision making with VA systems. Accordingly, our work makes the following original contributions:

- To situate our study in complex decision making tasks with visual interfaces, the experiments are conducted with a sophisticated visual analytics system [10] with multiple coordinated views. The design of the visual analytics system enables the visual anchor on either geo or time related representation through tutorial/training.
- In order to study the effect of anchoring bias on the decision-making processes with greater nuance and granularity, we collect not only quantitative measures about users' performance, including questionnaire responses, but we also collect detailed interaction logs within the visual interface. The interaction logs capture the decisionmaking process at a action level. Significant differences in actions were found between subjects assigned to different visual anchors.
- In addition to running statistical tests on the quantitative measures collected through pre- and post-questionnaires, we apply two novel methods of analysis - graph analysis and structural topic modeling - to analyze the paths and patterns of users interactions and identify the effect of anchoring bias. Our analysis revealed that visual anchors impact users' decision-making processes while numerical anchors affect the analysis outcomes.

2 Background and Related Work

In this section, we describe literature in the areas relevant to our study.

2.1 Background on Anchoring Effect

Humans have the tendency to rely on heuristics to make judgments, which can lead to efficient and accurate decisions [22], however these heuristics may also lead to systematic errors known as cognitive biases [29]. Psychologists have long studied the presence of cognitive biases in human decision making process [29, 45]. The anchoring and adjustment bias, defined as *the inability of people to make sufficient adjustments starting from an initial value to yield a final answer* [45], is one of the most studied cognitive biases that can lead individuals to make sub-optimal decisions. In the classic study by Tversky and Kahneman [45], the authors found evidence that when individuals are asked to form estimates, they typically start with an easily accessible value or reference point and make adjustments from this value. While such an approach may not always lead to sub-optimal decisions, research has demonstrated that individuals typically fail to adjust their estimates away from their initial starting point the *anchor*. Research has shown that anchoring affects decision making in various contexts, including judicial sentencing [3], negotiations [32] and medical diagnoses [7]. Given this documented prevalence of anchoring bias in various contexts of decisionmaking activities, we hypothesize that such effects may also be present when individuals interact with data while using visual analytics.

2.2 Visual Analytics and Cognitive Biases

Sacha et al. [43] investigate how uncertainties can propagate through visual analytics systems and examine the role of cognitive biases in understanding uncertainties, and also suggest guidelines for the design of VA systems that may further facilitate human decision-making. Similarly, research in the detection of biased decision making with VA software is in the early stages [36]. Harrison et al. found through a crowd-sourcing experiment that affective priming can influence accuracy in common graphical perception tasks [23]. George et al. [20] examined robustness of anchoring and adjustment effect in the context of decision support systems. Although their study revealed the presence of anchoring bias in the user's decision making task of estimating the price of house, their decision support system did not contain a highly complex visual interface consisting of coordinated multiple view. Researchers have also investigated the role of various other biases such as confirmation bias [12] and attraction effect [14] in the context of visual analytics. Dimara et al. [14] studied attraction effect using crowdsourcing experiments to determine that attraction bias did in fact generalize to information visualization and that irrelevant alternatives may influence users choice in scatterplots. Their findings provide implications for future research on how to possibly alleviate attraction effect when designing information visualization plots but no study to date has explored the anchoring bias in visual interfaces. Additionally, no research to date has examined the interaction patterns and activities of users in decisionmaking while these users are explicitly anchored under controlled experimental conditions.

In the next section, we describe a novel approach to analyzing the users' interaction patterns which is grounded in the analysis of web log data.

2.3 Use of Topic Models for Analyzing Web Logs

For our analysis of the interaction logs, we employ a variant of topic models, structural topic modeling (STM), that facilitates testing the effect of document-level variables on topic proportions. By characterizing the temporal sequence of actions taken by the user during their interactions with the interface as a 'text document' and characterizing actions as 'topics', we are able to test the effects of several factors, which include not only demographic variables such as age and gender, but also the effects of anchoring bias on the user's actions, and hence their decision-making processes. Although topic models have been used to analyze web logs previously, our application of STM to user interaction logs is novel by providing a mechanism to test the effect of independent variables on actions (topic proportions). Early applications of topics models [13, 26] to analyze web log behavior used probabilistic latent semantic indexing (pLSI) [24], a predecessor model to LDA-based topic models [5]. In the case of analyzing web log data, the pLSI model has been helpful in capturing meaningful clusters of users' actions, and found to surpass state-of-the-art methods in generating user recommendations for Google News [13].

One limitation of this method was that it did not consider time or user-level attributes (independent variables) within the model. To address the issue of time, Iwata et al. [25] created a LDA-based topic model (Topic Tracking Model) to identify trends of individuals' web logs on two large consumer purchase behavior datasets. In their model, they created a time component to identify the dynamic and temporal nature of topics. As we will discuss in section 5, we address the same concern by creating a time component in our STM model. Further, we employ STM's flexible causal inference framework as a mechanism to test anchor bias by treating each anchor group as additional independent variables.

3 User Experiment

In this section, we first describe our research questions and provide a detailed description of the visual analytic system used in the exper-

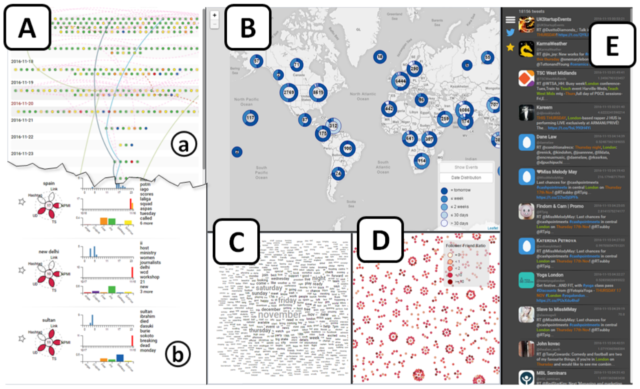

Figure 1: Crystalball interface: the interface has 4 main views: (A) calender view, (B) map view, (C) word cloud view, and (D) social network view. The calendar view shows the future event overview (a) by default. The event list (b) is shown when the user selects a subset of future events. The tweet panel (E) is shown when the user clicks the Twitter icon.

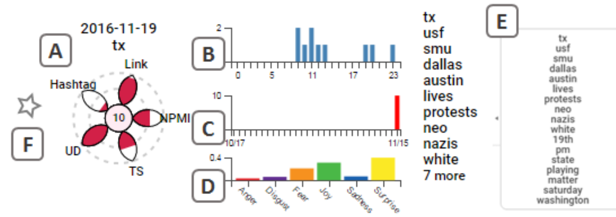

Figure 2: Event list. The flower glyph shows 5 measures of the future event and the number of tweets in the center (A). The three bar charts in the center show hourly distribution of tweet positing time (B), the number of tweets pointing to the event in last 30 days (C), and averages of emotion scores of the tweets (D). A list of keywords that summarize the tweets are displayed next to the bar charts (E). The user can bookmark the event by clicking the star icon (F).

iment. We then describe the experiment design rationale and tasks designed to elicit and test anchoring bias, and provide details about the experimental procedures and participants next.

3.1 Research Questions

Given that our research lies at the intersection of anchoring bias, decision making processes, and visual analytics systems, we designed two types of anchors, namely visual and numerical to evaluate their effects in the context of visual analytics systems. **The numerical anchor** is based on many psychology studies to test whether the participants can adjust away from the numerical anchor in their final answers. **The visual anchor** is designed specifically to prime people with different views in visual analytics interfaces with CMV design. The design of the numerical anchors is to evaluate if users are subject to anchoring bias when using visual analytics interfaces to aid decision making in a way similar to whats found by previous experiments conducted without the use of a visual analytics interface; while the design of the visual anchors is to test specifically whether users can be anchored visually and how that affects the analysis process and outcome. More specifically, we seek to answer the following research questions with respect to the impact of anchoring on decision-making activities using visual analytics systems:

- RQ1 - Visual Anchor: Can individuals be anchored on a specific view in a CMV?
- RQ2 - Numerical Anchor: Are the effects of numerical priming transferable to VA?
- RQ3 - Interaction Patterns I: How does anchoring influence the *paths* of interactions?
- RQ4 - Interaction Patterns II: Are there *systematic differences* in the interaction patterns and information seeking activities of individuals primed by different anchors?

To answer these questions, we designed and conducted a controlled experiment using a custom VA system, which is described next.

3.2 CrystalBall - a visual analytics system used for the experiment

To study the anchoring effect during complex decision making tasks performed in a visual analytics system, we conduct the experiment with Crystalball, a visual analytics system that facilitates users in identifying future events from Twitter streams [10]. Detecting future events from tweets is a challenging problem as the signals of future events are often overwhelmed by the discussion of on-going events. Crystalball is designed to detect possible future events from streaming tweets by extracting multiple features and enables users to identify potentially impactful events.

3.2.1 Analyzing large and noisy Twitter data

On average, around 500 million tweets are posted on Twitter per day by more than 300 million Twitter users [1]. However, many of them discuss past and ongoing events, and news headlines. To find, identify and characterize possible future events, the Crystalball system pipeline contains multiple components, including entity extraction, event identification and a visual interface.

The pipeline first extracts location and date from tweets. If the extracted date refers a future time and the extracted location is valid, then the tweet goes to the event identification component. Even if a tweet may mention a future time and valid location, it is possible that the tweet does not contain any informative content. Thus, in order to determine the quality of tweets as indicators of future events, we employ 7 measures: Normalized Pointwise Mutual Information (NPMI) of time-location pairs, link ratio, hashtag ratio, user credibility, user diversity, degree centrality and tweet similarity.

3.2.2 Multiple Coordinated views in the CrystalBall Interface and User Interactions

Figure 1 shows the Crystalball interface. The interface has 4 main views: a calendar view, map view, word cloud view and social network view. **The calendar view** displays a list of future events (Figure 1A). By default, it shows overview of future events (event overview, Figure 1a). The event overview shows all identified events and connections among them. Circles represent identified future events. The circles are grouped by dates. Events that have a same location are connected with a solid line and events that have same keywords are connected with a dotted line.

The event view shows detailed event information (event list) when the user selects a subset of the future events as shown in Figure 1b. Figure 2 shows enlarged image of Figure 1b. A flower glyph visualizes 5 of 7 measures of a future events with the number of tweets in the center (Figure 2A). The 5 measures are link and hashtag ratios, NPMI, user diversity and tweet similarity. Two timeline bar charts visualize distribution of tweet positing time (Figure 2B) and the number of tweets in last 30 days (Figure 2C). The bottom bar chart shows average emotion scores of the tweets (Figure 2D). Keywords that summarize tweets of the event is displayed on the right side of the view (Figure 2E). The event can be bookmarked as favorite by clicking the star icon (Figure 2F). The bookmarked events are stored in database so that the user can review them anytime.

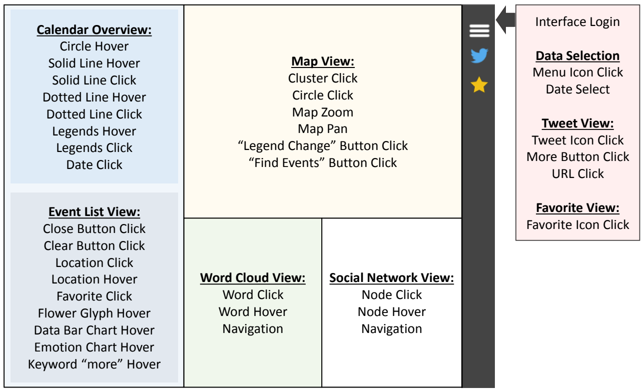

Figure 3: User interaction logs: the figure displays main user interaction logs of the Crystalball interface. Each view has different interaction logs based on its visual elements.

The map view shows identified events on the map to indicate where they will occur (Figure 1B). Events are aggregated based on the zoom level and are shown as rings. The color of ring proportions represent event dates ranging from tomorrow (dark blue) to more than a month (light blue). Clicking a ring will show its tweets as circles. Clicking a circle will show a tooltip showing the tweet.

There are two facilitating views to help users explore and further analyze the future events: **word cloud view** and **social network view** (Figure 1 C and D). The word cloud view shows keywords extracted from tweets of the identified events. The size of keywords represent frequencies of the keywords. The word cloud view is updated when selected events are changed. The social network view represents relationships between future events and Twitter users. Clusters in the view represent future events in same locations. In many cases, a cluster has several future events in a same location.

User Interactions: The highly interactive and exploratory nature of the Crystalball interface enables users to start their exploration and analysis of future events from any of the four main views present in the interface.

A user can start with the calendar view in order to know when the event will occur. Hovering the mouse over a circle in the event overview will highlight the corresponding events on the map, word cloud and social network view. The user can find events that share a same location or keywords by examining links. The user can select a particular date then the event list will be shown in the calendar view that shows all events of the date with detailed information. Other views will be automatically updated to show corresponding events in the views.

Alternatively, the user can start the analysis from the map view to make sense of where the event will occur first. The map view shows detailed evenets when zooming into a region of interest. When the zoom factor is lower than a zoom threshold, the calendar, word cloud and social network views are updated to show the events in the current map extent. The user can open the event list to show all the events in the map extent by clicking the "show events" button on the bottom right of the map view.

The interactions implemented in CrystalBall allows users to perform exploratory analysis to support decision-making tasks. Consequently, the decision making process is reflected by the actions participants take within CrystalBall. In our experiment, in order to analyze the effect of anchoring bias on a decision making task conducted in CrystalBall, we defined and logged 39 unique user interactions. Figure 3 lists 36 user interactions, situated in their corresponding views. The rest of the 3 interactions were used rarely but our participants during their interactions with CrystalBall. The interface records all user interaction logs with a timestamp and a user's name to database which are analyzed in Section 5 in order to show users' decision making process.

3.3 Design Rationale

The anchoring effect has been replicated in numerous studies in the laboratory and in the field [17]. Our experiment design is thoroughly grounded in these best practices of controlled experimental studies in that we use priming to elicit the anchoring bias. First, we focused the participants experience around a well-defined cognitively engaging decision-making task - we asked participants to estimate the number of protest events in a given period of time and in a given location. We conducted our experiment with the CrystalBall interface to predict and detect protest events from Twitter data. The *Calendar View* and the *Map View* as described in Section 3.2 serve as the **time** and **geo (visual)** anchor. In order to test our hypotheses, we followed a 2×2 between-subjects factorial design with two factors (numerical and visual) and each factor had two levels as described below.

3.4 Experimental Stimuli

The visual and numerical anchors for the experiments were devised in order to prime the participants in two ways. The numerical anchor primed participants on a number (High or Low) and the visual anchor primed participants on a specific view in the CrystalBall interface (map view, representing geo anchor or calendar view, which represented the time anchor). The decision-making task presented to the participants is one of the four choices presented below:

Geo + high/low anchor: Do you think that the number of protest events in the state of California <geo anchor first> was higher or lower than 152 (or 8) <high (or low) numerical anchor> between November 10, 2016 and November 24, 2016 <time anchor>?

Time + high/low anchor: Do you think that the number of protest events between November 10, 2016 and November 24, 2016 <time anchor first> was higher or lower than 152 (or 8) <high (or low) numerical anchor> in the state of California <geo anchor>?

As can be noted, the magnitude of the numerical anchor, either high or low, is subject to the experimental condition. These high and low numerical anchors were chosen based on the actual number of protest events present in the data (as determined by trained annotators). Additionally, the **order** of presentation of the visual anchors varies in the two questions. The visual anchors were further **reinforced** through custom training videos orienting the participants to the use of the CrystalBall interface.[1] The two training videos reinforce the visual anchors by starting and driving the analysis from either the map view (geo) or the calendar view (time).

3.5 Experiment Design

3.5.1 Procedures

The data collection for this study involved in-person laboratory participation. Participants were recruited via in class recruitment, email to listservs and the psychology research pool at our university. Sessions were conducted between February 10th, 2017 and March 15th, 2017. After signing up for the study, participants were assigned a unique code for secure identification. Associated with this code, was the random assignment to one of four experiment conditions (High/Geo, Low/Geo, High/Time, Low/Time). Participants were asked to come to the lab for the duration of one hour. The experimenter would first elicit their responses to informed consent. Next, the participants would view two training videos specifically designed for this experiment. The first video was a general training video (duration 5 minutes) which oriented them to the use of the Crystal Ball interface and its basic functionality (e.g., primary and

[1] Please refer to supplemental materials for the two training videos.

Table 1: Distribution of participants in 4 conditions. Row-Numerical anchor; column-Visual anchor.

	Geo	Time	Grand Total
High	20	21	41
Low	22	18	40
Total	42	39	81

supporting interactions). This video was shown to all the participants, regardless of experimental condition. Next, the participants were shown a priming video (duration of 3 minutes) based on their visual anchoring group. The priming video was designed to guide the users through a case scenario through either Geo or Time Visual Anchors.

Following the training, the participants were asked to complete a pre-test questionnaire. The pre-test questionnaire consisted of questions related to participant's demographics (age, gender, education), their familiarity with visual analytics systems and social media, and Big-5 personality questions [27]. The informed consent, training video and pre-test questionnaire typically took around 20 minutes to complete. The participants were then assigned the task, and asked to interact with CrystalBall for about 25 minutes. We designed and implemented interaction logging with the CrystalBall interface to capture their timestamped actions as they proceeded through the task. The interaction logging is transparent to the participants. At the end of their interaction, participants were asked to estimate the number of protest events based on their analyses within CrystalBall.

Next participants were asked to complete a post-test questionnaire. The completion of the post-test questionnaire ended their participation in the study. The post-test contained questions regarding the usability of the system (ease, attention, stimulation, likability), level of engagement during the task and questions to gauge their susceptibility to bias. The bias questions consisted of eight questions designed to measure the level of bias. Participants were compensated by either a $5 gift card or class credit assigned at the discretion of the class instructor willing to assign extra credit.

3.5.2 Participants

A total of 85 participants completed the study. We discarded the data for four participants due to usage of incorrect identification codes during the experiment. Distribution of participants across experiment conditions was relatively even and is represented in Table 1. Figure 4 shows a summary of the demographic characteristics of participants across factors including age, gender, education and major. We note that there is an even balance of participants across various demographic characteristics such as gender (male vs. female), age (different age ranges) and education background (computing vs. other majors), although there is some skewness in the data towards students pursuing Masters degrees. Participant demographic characteristics were also balanced across the four experiment conditions due to random assignment of participant to experiment condition. Males and females participants were uniformly distributed across experiment conditions (25% in each condition, $SD = 10\%$ for males, $SD = 6\%$ for females). Average ratio for males to females was 1.02 in each experiment condition. Average proportion of undergraduate, masters and PhD education level in each experiment condition was also 25% ($SD = 13\%$, 10% and 24% resp.).

4 Experiment Results: Analyzing Quantitative Measures

Two types of quantitative analyses are conducted to answer research questions RQ1 and RQ2 introduced in Section 3.1.

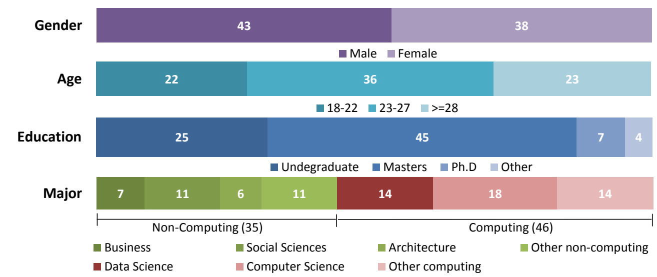

Figure 4: Demographic. A summary of demographic information of the participants based on gender, age, education and major.

4.1 RQ1 - Visual Anchor: Can individuals be anchored on a specific view?

To quantitatively evaluate whether participants can be anchored on a view in CrystalBall, we conducted two types of stastical analysis - two-way ANOVAs and Bonferroni-corrected pairwise t-tests. We extracted the overall time duration a participant spent in geo or time-oriented views from the interaction logs by taking into account the time stamp of each action occurred in a particular view.

A two-way ANOVA was conducted on the influence of two independent variables (numerical and visual anchor) on the amount of time spent in different views (map vs. calendar). The main effect for visual anchor was statistically significant and had an F ratio of $F(1,78) = 11.57, p < .001$. The main effect for numerical anchor indicated that the effect for numerical anchoring did not significantly affect the time spent in map vs. calendar view ($p > 0.05$). The interaction effect was not significant, $F(1,77) = 0.12, p > 0.05$.

Bonferroni corrected pairwise t-tests (α=0.05/4) were conducted to compare the duration of time spent in the different conditions with α=0.0125 level of significance. We found that visual anchor had significant effect on the time spent in map view vs. calendar view across both conditions ($p < 0.01$ in both cases) whereas the numerical anchor did not ($p > 0.05$ in both cases).

In Figure 5, we show the duration of time spent in each view for each participant across all four experimental conditions. The x-axis represents the time in minutes, with the blue bars representing duration in calender view and the black bar representing duration in map view; the y-axis are the participant ids in each condition. We see from charts labeled High/Geo and Low/Geo that participants spent significantly more time in the map view vs. the calendar view in the geo anchoring conditions. The charts labeled High/Time and Low/Time reveal that in the Time priming conditions, the time spent in each view was variable and no statistical trends can be observed. We have included four separate charts in Figure 5 to provide sufficient comparative detail across the experiment conditions.

4.2 RQ2 - Numerical Anchor: Are the effects of numerical priming transferable to VA?

The effect of numerical anchor on time spent within CrystalBall. As reported in Section 4.1, two-way ANOVAs conducted in order to determine the effects of numerical anchoring indicated the main effect for numerical anchor did not significantly affect the time spent in map vs. calendar view ($p > 0.05$). The interaction effect was also not significant, $F(1,77) = 0.12, p > 0.05$.

These findings indicate that being primed by a numerical anchor did not have an effect on the amount of time spent in map view compared to the calendar view. We discuss the implications of these findings further in the discussion section, and suggest that more investigation is needed to determine the cause of these effects.

The effect of numerical anchor on the decision-making outcome. To further asses the impact of numerical anchoring, we

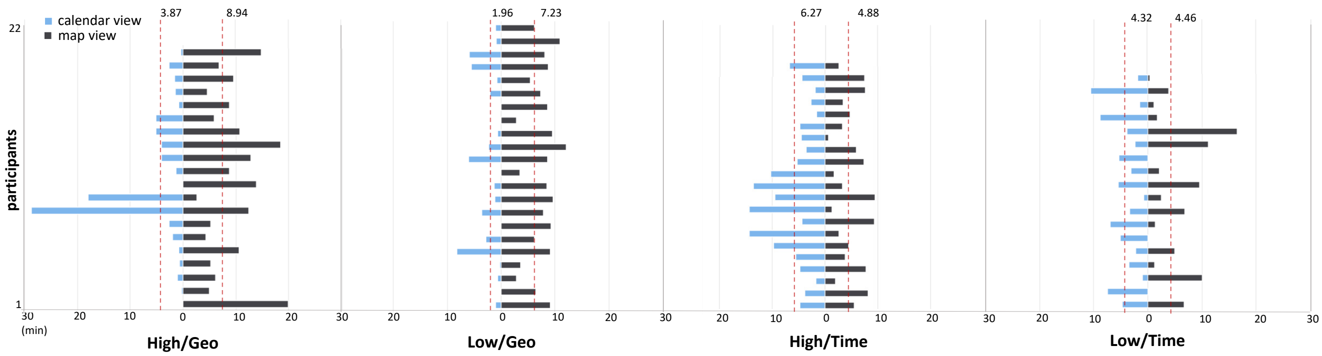

Figure 5: This figure provides a summary of the amount of time spent in calendar and map views on each of the four different conditions. The red dashed line is the mean of the amount of time spent in calendar and map view.

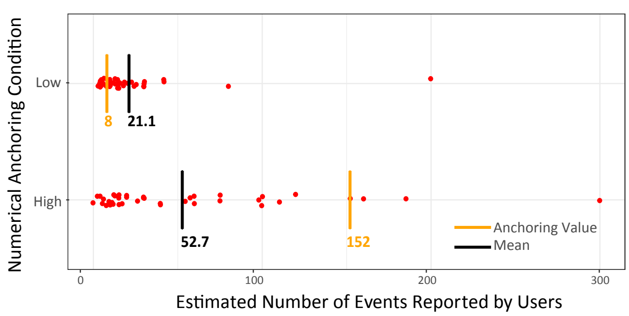

Figure 6: The estimated number of geo-political events reported by each participant is represented by red dots. The orange line represents the anchor value and black line is the mean of estimated number of political events.

analyzed the responses given by participants in the pre- and post-tests. The participants were asked to estimate the number of protest events, before and after their interactions with the data and the interface. The findings are shown in Figure 6. On the x-axis we show the two groups in the numerical anchoring condition High and Low. On the y-axis are each participant's estimates regarding the number of protest events, before the interaction (in orange) and after the interaction (in red). Mean post-test responses are in black. Our findings indicate that participants were consistently anchored on initial number presented to them in the framing of the questions ($p < 0.05$). Our findings suggest that the effects of the classic anchoring bias elicited by priming with numerical anchors in previous laboratory studies can be replicated in VA. We did not find any effects of the visual anchoring (geo vs. time) on the final outcome ($t(40) = 2.02, p > 0.05$), suggesting that the effects of the visual anchor may be more subtle than can be determined via post-test questionnaire responses. We conducted detailed analyses to capture these effects, which we shall describe in the next section.

5 Experiment Results: Analyzing Interaction Logs for User Activity Patterns

To test the hypothesis that user interaction logs reflect the participants' decision making processes, we applied two additional types of analyses on the logs in order to evaluate the impact of anchoring effect on the patterns of user interactions. The two analyses address research questions RQ3 and RQ4 (Section 3.1).

5.1 RQ3: How does anchoring effect influence the paths of interactions?

To analyze the paths users take during their analysis with CrystalBall and the effect of anchors on these paths, we developed a novel method to study the sequences of users interactions as a network of interaction nodes. We constructed the interaction network as follows: each interaction is logged with five attributes: time stamp of the interaction, the view it took place in, the type of interaction, and detailed description for each interaction (e.g., 12:56:35.56, Calendar view, Click, Zoom 89.55 36.00). As shown in Fig 3 there are 36 main interactions as well as 3 main secondary interactions over multiple views. Each of these interactions form the nodes in a network. The edges in the network are chronological pairs of interactions. For example, if a user has zoomed on the map and then hovered on a particular location in the calendar view, this would add an edge between the *Map Zoom* and *Calendar Location Hover* nodes. The edge weight would incrementally increase for each additional observed pair. For visualization purposes, we disregarded self-loops (i.e, repeated actions) because we are more interested in the relationship between different interactions and the paths of interactions taken by our users. This method yields a weighted directed graph which enables us to cluster interactions through community detection, rank each action by multiple centrality measures and compare aggregate user path differences controlling for each anchor. The network visualizations in this section were created with Gephi [2].

In this section, first we take actions of all users into account to get a complete picture of users' paths of interactions. We then analyze differences in users paths to detect different user strategies controlling for the two anchors (visual and numerical). We studied the network of all user logs, as well as our four experiment anchors. We did not find significant differences in the networks created from logs of users primed on the two numerical anchors. Hence in this discussion, we will focus on the three remaining networks: a full network (AllNetwork), a Geo-anchored network (GeoNetwork) and a Time-anchored network (TimeNetwork).

5.1.1 Analyzing the network of all interactions

By adopting an exploratory data analysis method, we started by analyzing different features of AllNetwork (39 nodes and 640 edges) (Figure 7). We first utilized the community detection algorithm developed by Blondel et al. [6], which resulted in 5 different communities of interactions. Most of these communities are comprised of interactions that occur within the same view or have a close semantic relationship to each other. The community detection results allowed us to categorize nodes in our network into three main groups that were in line with our initial system design strategies: preliminary interactions, primary interactions, and supporting interactions. The

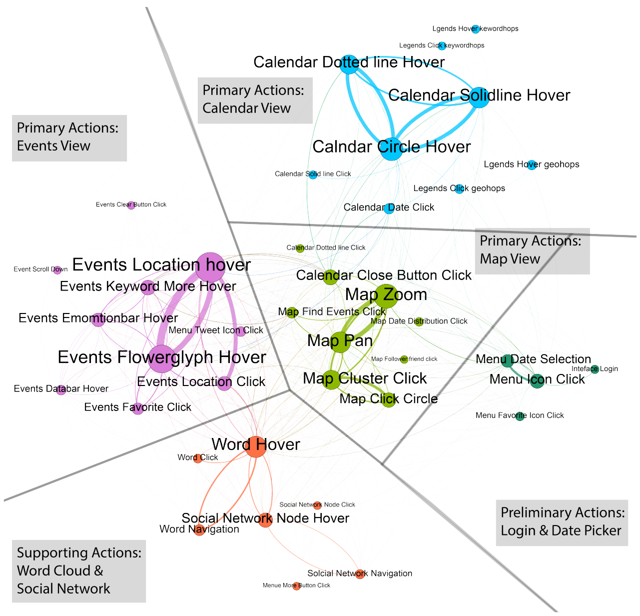

Figure 7: The directed network of all interactions. Nodes are interactions and edges are interactions that occur after each other. The size of nodes are proportional to Pagerank values and width of edges are proportional to the edge weights. Note, if a line is drawn between a start-node and an end-node, the outgoing edge from the start node is on the relative left side of that line.

primary interactions include those that users have to go through in order to find the events of interest. The supporting interactions are those that users perform in order to find supporting information to confirm the previously found events. The preliminary interactions such as login and clicking on the menu bar were used infrequently as they are not critical to the analysis process. Figure 7 shows the network colored and annotated based on the community detection results.

In order to measure the importance of different interactions, we utilized the Pagerank algorithm [37]. Since Pagerank takes into account the weight of edges between interactions, it is much more powerful than simply calculating the frequency of each interaction. Pagerank assigns probability distributions to each node denoting the importance of the node. These probability distributions are appropriate metrics for importance of the interactions in our system as they show the likelihood of a random surfer in the network to traverse to a specific end node. The top ranked interactions in our interface are all from the primary action communities with the exception of *Word Hover*. As seen in Figure 7, edge weights between important nodes in the same community are higher than ones between different communities. Furthermore, we can observe mutual higher weighted paths between these higher ranking interactions. Some exceptions to this observation is when the users are moving away from a view to another conduct more in depth analysis. For example, in AllNetwork, the edge between *Events Location Hover* to *Events Location Click* is weighted very strong, but the path of opposite direction is not. We can interpret *Events Location Click* as an interaction that drives users out of this community to others such as Word Cloud and Social Network to get complementary information of an event (See Fig 7). The AllNetwork and the analyses resulting from it serve as a reference for the comparisons we wish to make across the GeoNetwork and TimeNetwork.

5.1.2 Comparing interaction networks of Geo- vs. Time-anchored users

To answer our research questions of whether visual anchor has an effect on the way different groups of participants interact with the visual interface, we constructed two networks based on the actions of the geo and time-anchored groups. These networks consist of the same 39 action nodes but have different edges and weights allowing us to compare the interactions of participants primed on the two visual anchors through the lens of their respective networks.

Similar to our analysis of AllNetwork, we first started by detecting communities within GeoNetwork and TimeNetwork. Interestingly, the results show similar community structures to AllNetwork. However, there are subtle differences that point to the differences regarding the usage of CrystalBall between the two groups. For example, the action of *Favorite Icon Click* (through which users can save an event to view later in the Favorites Menu) in GeoNetwork is part of the preliminary actions community, but for TimeNetwork it is part of the Time related primary actions community. This subtle change could indicate that the time primed users had more interactions between saving an event as a favorite and then viewing the list of favorite actions in comparison to our Geo primed users.

We calculated Pagerank for interaction nodes in both these networks. Comparing these values would allow us to understand important interactions within each network and how they are affected by the visual anchor. Figure 8 illustrates significant differences of the two networks. In the GeoNetwork, the top nodes are a mixture of interactions from the Map and Event views, with the highest ranked interaction from the Map view. This pattern is consistent with the strategies shown in the Geo priming video. In contrast, in the TimeNetwork, the top ranked nodes are interactions within the Events view and the Calendar View, which is also consistent with the strategy shown in the time-anchor video. Other important but lower ranking actions in TimeNetwork are from the Map View. Furthermore, by observing the paths between the Events community (colored purple in Figure 8) in both GeoNetwork and TimeNetwork, we see that weight of the edge between *Events Location Hover* and *Events Location Click* is relatively higher in the GeoNetwork in comparison to the TimeNetwork. This could indicate that our Geo primed users use maps to explore and primarily use hovering and clicking on a location together to view more details in the map, word cloud, and social network views. Our time primed users on the other hand, utilize the hovering on locations to explore events. These differences show interesting behavioral variations in sequences of interactions between our two groups. These differences show that time primed users are more likely to use the Calendar and Events view actions as their primary exploratory tool. Figure 8 shows the comparisons between these two networks and two bar charts comparing the top 5 ranked nodes in each network.

Analyzing the interactions of our participants as a network has many benefits. It allows us to take into account the sequence of interactions, as well as the paths taken by users to arrive at the conclusion. The paths taken reflect the strategies users employ during the decision-making process. Furthermore, we can take an overview of all interactions within the CrystalBall interface and analyze what strategies need to be improved to make the interface more effective.

5.2 RQ4: Estimating the effects of anchoring bias on interaction patterns and information seeking activities

One drawback of the network analysis is that the estimated impact for each anchor is measured without standard errors to calculate the statistical significance of each result. To address this problem, we use structural topic modeling (STM) to measure the impact of the anchors and user-level attributes on users' actions [40]. Originally built for text summarization, STM is a generalized topic model

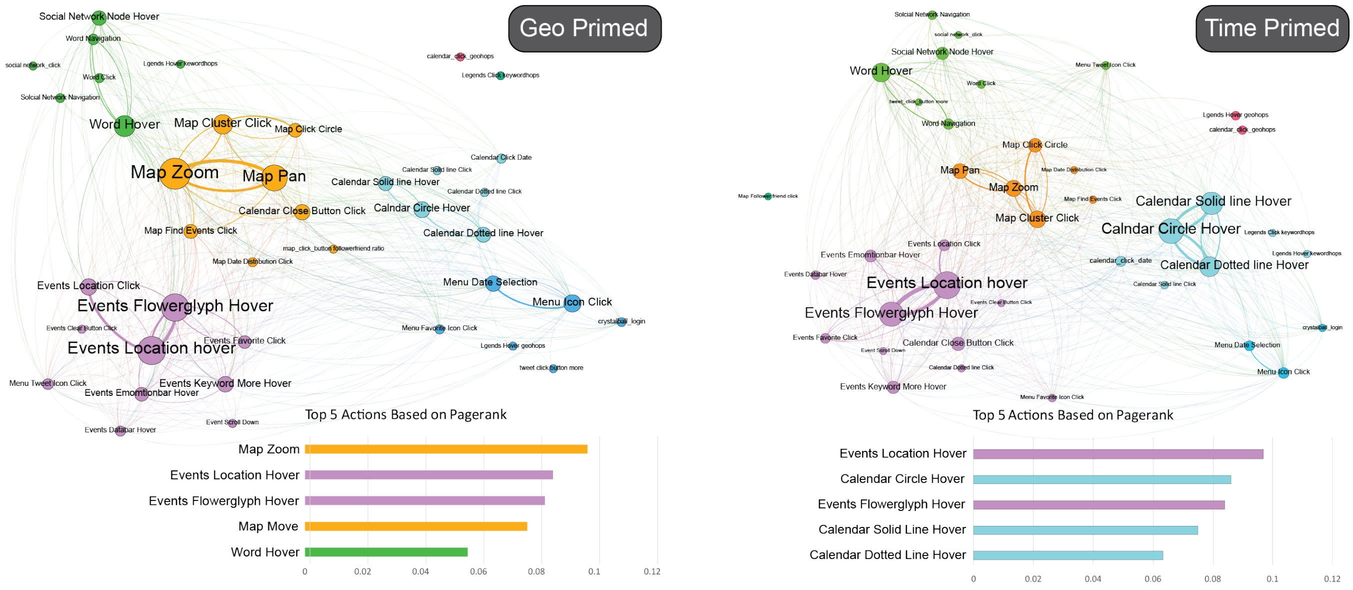

Figure 8: Side by side visualization of GeoNetwork and TimeNetwork. The size of nodes is proportional to Pagerank values of nodes in each graph, the color of nodes corresponds to the detected community of each node, and the width of each edges corresponds to the weight of that edges. The bar charts show the top 5 nodes based on their Pagerank value and is color coded based the community the nodes community.

framework for testing the impact of document-level variables. For our model, topics are clusters of interactions measured as probability distributions over the action space. We test our hypotheses of the effect of anchoring on the topic proportions through an embedded regression-component. STM is a consolidation of three predecessor models: the correlated topic model (CTM) [4], the sparse additive generative model (SAGE) [15] and the Dirichlet multinomial regression model (DMR) [35]. The CTM model introduces a correlation structure for topic distributions while the DMR and SAGE models provide mechanisms to estimate the effect of independent variables on either topic proportions (via DMR model) or word distributions for each topic (via SAGE model). [2]

Table 2: Independent Variables Tested

Type	Independent Variable	Level
Condition	Visual Anchor	Time / Geo
	Numerical Anchor	High /Low
Time	Percent of Actions	Action Deciles (b-spline)
Attribute	Gender	Make / Female
	Major	Computing / Non Computing
	Age	Under 23 / Over 23
	Education	Undergraduate / Graduate
Personality	Extroversion	High / Low
	Agreeableness	High / Low
	Conscientiousness	High / Low
	Openness	High / Low
	Neuroticism	High / Low

However, as with most topic models, STM is built from the bag-of-words (BoW) assumption that provides a key advantage and disadvantage in our analysis. The advantage is that it yields statistical properties (exchangability) that identifies topics as clusters of co-occurring interactions and facilitates statistical testing through

[2]We used the stm R package [41] for our analysis. This package includes additional tools for topic modeling including a spectral initialization process that aids in addressing the multi-modality problem (stability of the results).

Table 3: This table provides the seven actions with the highest probabilities for three sample topics: Map View, Calendar View and Event List (all tools). Action combinations (bi- or tri-grams) are denoted by the plus sign.

Rank	Map View (Topic 8)	Calendar Overview (Topic 6)	Event List: All Tools (Topic 3)
1	Map Zoom	Calendar Hover Circle	Event Keyword More Hover
2	Map Zoom + Map Zoom	Calendar Solid Line Hover	Event Keyword More Hover + Event Keyword More Hover
3	Map Pan	Calendar Dotted Line Hover	Event Keyword More Hover + Event Keyword More Hover + Event Keyword More Hover
4	Map Circle Click	Calendar Hover Circle + Calendar Hover Circle	Event Flower Glyph Hover
5	Map Zoom + Map Zoom + Map Zoom	Calendar Solid Line Hover + Calendar Hover Circle	Event Emotion Chart Hover
6	Map Cluster Click	Calendar Dotted Line Hover + Calendar Hover Circle	Event Favorite Click
7	Map Zoom + Map Pan	Calendar Solid Line Hover + Calendar Solid Line Hover	Event Flower Glyph Hover + Event Flower Glyph Hover

the DMR (GLM regression) component. On the other hand, a disadvantage of the BoW assumption is that it ignores the order of interactions. To address this issue, we made **two modifications**: extracting bi-/tri-grams and creating a session time variable by interaction deciles. First, we extracted every bi- and tri-gram as chronological action pairs and triplets from the interaction logs. Including bi- and tri-grams and the single actions, we had 237 unique features after removing sparse features. Second, we created a time variable that divided each user's session into ten evenly distributed groups (interaction decile). Given that each user's session averaged nearly 800 individual actions, each decile maintained sufficient interactions to facilitate topic inference. Additionally, inclusion of the time variable had the advantage of increasing our sample size (number of documents) from 81 to 810 as the document-level went from each user to a user's interaction decile (e.g. first 10% of user X's interactions).

To test the effect of anchoring bias on users' interactions, our

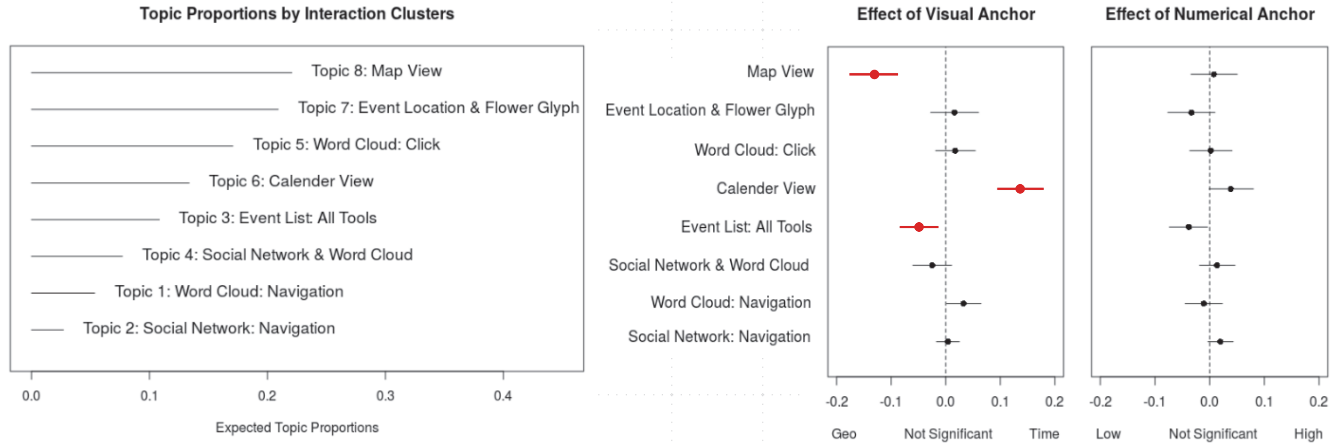

Figure 9: The figure on the left provides the expected topic (action-cluster) proportions with judgmental labels to aid in interpretation. The figures on the right provide the estimated effect of the visual and numerical anchors on each of the eight topics' proportions. The dot is the point estimate and the line represents a 95 percent confidence interval. The red dots/lines are topics that are significant with 95% confidence.

baseline model to explain topic proportions (dependent variable) incorporates three independent variables: the visual anchor, the numerical anchor, and time as interaction deciles. After analyzing the model, we tested other demographic attributes including gender, major, age, education level, and the Big-5 personalities. Table 2 above provides a list of the independent variables tested and the categorical levels. We binned the user attributes into binary levels. Similarly, we converted the Big-5 personality results into binary levels in which users who scored above the mean were categorized as High while users who scored below the mean were categorized as Low.

5.2.1 The effect of visual anchor on interaction patterns estimated by topic proportions

We find the visual anchor has a significant effect on the proportion of users' interactions as topics clustered automatically in view-based groups (e.g., map, calendar, events). Figure 3 provides the top seven interactions for three sample topics. We observe that the interactions tend to cluster into groups related to each interaction's associated view hierarchy as shown in Figure 3. For example, topic 8 includes interactions related to the map view including *Map Zoom*, *Map Pan*, *Map Circle Click* and *Map Click Cluster*. Therefore, we gave topic 8 the manual label of Map View since its interactions are all related to that view. Following this approach, we created manual labels for the other seven topics. Further, we find that the topics tend to cluster in groups consistent with our network communities found in Section 5.1. For instance, the four prominent interactions of the Map View topic (by probability) have the strongest connections as well as highest PageRank in the Map View community cluster (green nodes) in Figure 7.

Second, we find in Figure 9 that the Map View and Flower Glyph topics had the largest topic proportions. Alternatively, the social network and word cloud were the smallest topics. To test our number of topics, we followed the procedure recommended by [42] by considering multiple topic scenarios (5, 8, 10, 15, 20, 25, 30, 40) and comparing each model's held-out likelihood and average semantic coherence. We decided on an eight topic model given a high average semantic coherence and parsimony of topics (see supplemental materials).

We observe that the visual anchor had a significant effect on the Map View and Calendar View topics. Figure 9 provides the effect the anchors had on the topic proportions. In this plot, each dot is the estimated topic proportion difference for each topic by the two levels of each anchor. The line represents a 95% confidence interval around each estimate. From these figures, we find that the Map View and the Calendar View topic proportions have the most significant differences between the two groups. Consistent with our findings in sections 4.1 and 5.1, Geo primed users are anchored more to the view they were primed on while we see less of an effect in Time primed users. On the other hand, we found that the visual anchor had an unexpected effect with the Event List: All Tools (topic 3). Geo primed users tended to use tools like the *Keyword More* and *Emotion Bar* more than Time primed users. Alternatively, we find that the numerical anchor had only a marginal effect on two topics (Calendar View (topic 6) and Event List: All Tools (topic 3)). These results imply that the visual anchor had a more significant impact on the proportion of users' interactions than the numerical anchor. This is important as we observed opposite effect (numerical anchor was significant, visual anchor was not) in the users' estimation of the event outcome.

5.2.2 The effect of visual anchor on interactions used over time estimated by topic proportions

We find evidence of a temporal effect on the topic proportions. To measure this effect, we divided each user's interaction path into interaction deciles (see Section 5.1). To aid estimation, we used a b-spline to smooth the values. Figure 10 provides the effect of the visual anchor (line color) and time (x-axis) for the Map View and Calendar View topic proportions. We observe a significant impact of the time of the user's session on this topic proportions. For example, Map View topic proportion is nearly twice during the user's first twenty percent of interactions than users' remaining 80 percent of interactions. Moreover, we see this distinct drop for both visual anchor groups. This observation implies that users tended to use the Map View more in the beginning of the session as they were getting acclimated to the interface. Alternatively, the Calendar View topic trended down resulting in much lower use by session end (15% Time, single digits Geo). We found marginal effects of time for the other six topics, with most nearly flat given already low topic proportions (less than 10%).

5.2.3 The effect of demographic variables on interaction patterns estimated by topic proportions

To test other possible variables, we ran five additional model scenarios replacing the numerical anchor variable (as it showed only marginal significance) with the demographic variables (gender, major, student level and Big-5 personality). We found that none of

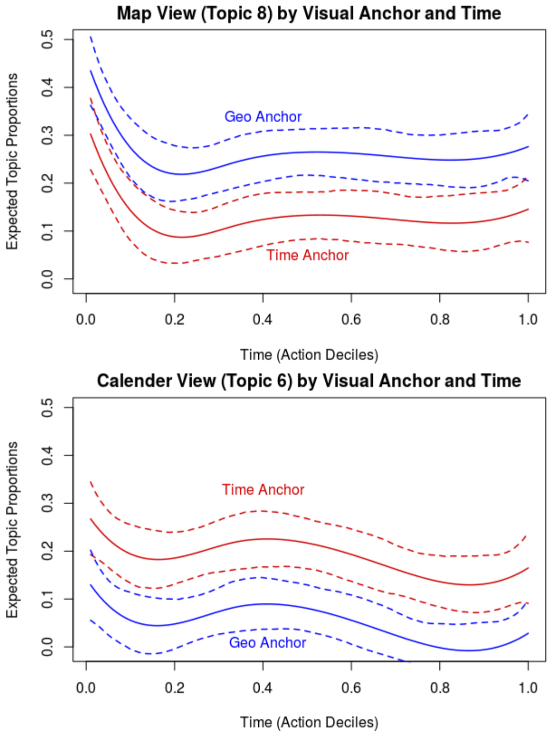

Figure 10: This figure provides two charts on the effect between the visual anchors (line color) and time as measured by interaction deciles (x-axis) for two topics (Map View and Calendar View). Each line is the estimated topic proportions across the session and controlling for the visual anchor. The solid line is the point estimate and the dotted line is a 95 percent confidence interval. For the interaction deciles (time), we divided users' sessions into ten evenly distributed groups. A b-spline was used to smooth the curve across the ten points.

the variables produced significant (95%) changes in topic proportions, although some produced marginally significant effects (see supplemental materials). For example, most variation occurs in the secondary view topics (Word Cloud, Social Network and an interaction topics).

6 Discussion and Limitations

In this section, we provide implications of our experiment results on anchoring effect in visual analytics, and point out possible limitations related to the study design and analysis.

Experiment implications As shown by our data analysis (section 4 & 5), our experimental results indicate that anchoring bias does transfer to visual analytics. Most interesting finding is that the visual anchor seems to significantly impact the decision-making **process**, while the numerical anchor has a significant effect on the decision-making **outcome**. The decision-making process reflects the way that users interact with CrystalBall; the outcome is the final answer that the participants provided at the end of the decision-making process.

Such findings have implications for user training on visual analytics systems with CMV, as well as how decision-making tasks are framed. With respect to training/tutorial, the visual analytic systems development team should provide multiple scenarios employing strategies that involve the use of different views as the primary visualization to drive the analysis. As of decision-making task framing, one should avoid accidentally anchoring the participants on an expected outcome or when possible, employ measures of cognitive bias (such as in our post-test) to evaluate the inherent cognitive bias of the users. As noted in Section §2.1, the tendency of humans to rely on heuristics to make judgments does often lead to efficient and accurate decisions. However, we need to determine when such heuristic decision-making is being applied, in order to ensure that the resulting decisions are optimal.

Experiment sample size limitation. As can be expected with any laboratory experiment, this research has limitations. One such limitation is the sample size of 81 participants in our experiment. However, the diversity of our sample with respect to gender, age, educational background and personality factors are steps we have taken to ensure the validity of our results. Our findings replicate the effects of anchoring that have been long studied in literature, further attesting to the validity of the experiment.

Experiment control limitation. Another limitation of our experiment is that we do not consider a control group , that is, participants who engage in the decision-making task without being primed by any anchors. While our initial study reported here was focused on determining whether the effects of anchoring are at all present and can be elicited in such experiments, our future work will be aimed at replicating these findings in more extensive experiments with larger sample size and will include control groups for comparison.

STM analysis limitations. Fong and Grimmer [18] note that topic models are susceptible to problems in estimating marginal effects due to the zero-sum properties of topic proportions. Further, topic models cluster only based on the count and ignore interaction duration (time spent). To address such limitation, the quantitative analysis in section 4 explicitly accounted for the duration of each interaction.

7 Conclusion

In this paper, we presented a systematic study and resulting analyses that investigate the effect of anchoring bias on decision-making processes and outcome using visual analytic systems. Our experimental results provide evidence on anchoring effect being transferable to visual analytics in that visual and numerical anchors affect the decision-making process and outcome respectively. The present study is a first step in an overarching research agenda of determining the use of heuristics in decision-making processes from the user interactions and if these decision-making processes can be reliably inferred then to automatically suggest ways in which to improve the process.

References

[1] Twitter usage statistics. `http://www.internetlivestats.com/twitter-statistics/`. Accessed: 2017-03-31.

[2] M. Bastian, S. Heymann, and M. Jacomy. Gephi: An open source software for exploring and manipulating networks, 2009.

[3] M. W. Bennett. Confronting cognitive anchoring effect and blind spot biases in federal sentencing: A modest solution for reforming a fundamental flaw. *J. Crim. L. & Criminology*, 104:489, 2014.

[4] D. M. Blei and J. D. Lafferty. A correlated topic model of science. *The Annals of Applied Statistics*, pp. 17–35, 2007.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

[6] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[7] J. S. Blumenthal-Barby and H. Krieger. Cognitive biases and heuristics in medical decision making: a critical review using a systematic search strategy. *Medical Decision Making*, 35(4):539–557, 2015.

[8] L. Cen, G. Hilary, and K. J. Wei. The role of anchoring bias in the equity market: Evidence from analysts earnings forecasts and stock returns. *Journal of Financial and Quantitative Analysis*, 48(01):47–76, 2013.

[9] I. Cho, W. Dou, D. X. Wang, E. Sauda, and W. Ribarsky. Vairoma: A visual analytics system for making sense of places, times, and events

in roman history. *IEEE transactions on visualization and computer graphics*, 22(1):210–219, 2016.

[10] I. Cho, R. Wesslen, S. Volkova, W. Ribarsky, and W. Dou. Crystalball: A visual analytic system for future event discovery and analysis from social media data. *Visual Analytics Science and Technology (VAST), 2017 IEEE Conference on*.

[11] G. Convertino, J. Chen, B. Yost, Y. S. Ryu, and C. North. Exploring context switching and cognition in dual-view coordinated visualizations. In *Proceedings International Conference on Coordinated and Multiple Views in Exploratory Visualization - CMV 2003 -*, pp. 55–62, July 2003. doi: 10.1109/CMV.2003.1215003

[12] M. B. Cook and H. S. Smallman. Human factors of the confirmation bias in intelligence analysis: Decision support from graphical evidence landscapes. *Human Factors*, 50(5):745–754, 2008.

[13] A. S. Das, M. Datar, A. Garg, and S. Rajaram. Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web*, pp. 271–280. ACM, 2007.

[14] E. Dimara, A. Bezerianos, and P. Dragicevic. The attraction effect in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):471–480, 2017.

[15] J. Eisenstein, A. Ahmed, and E. P. Xing. Sparse additive generative models of text. 2011.

[16] D. M. Eler, F. V. Paulovich, M. C. F. d. Oliveira, and R. Minghim. Coordinated and multiple views for visualizing text collections. In *2008 12th International Conference Information Visualisation*, pp. 246–251, July 2008. doi: 10.1109/IV.2008.39

[17] M. Englich. Anchoring effect. *Cognitive Illusions: Intriguing Phenomena in Judgement, Thinking and Memory*, p. 223, 2016.

[18] C. Fong and J. Grimmer. Discovery of treatments from text corpora. In *In Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 2016.

[19] A. Furnham and H. C. Boo. A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40(1):35–42, 2011.

[20] J. F. George, K. Duffy, and M. Ahuja. Countering the anchoring and adjustment bias with decision support systems. *Decision Support Systems*, 29(2):195–206, 2000.

[21] N. A. Giacobe and S. Xu. Geovisual analytics for cyber security: Adopting the geoviz toolkit. In *Visual analytics science and technology (VAST), 2011 IEEE Conference on*, pp. 315–316. IEEE, 2011.

[22] G. Gigerenzer and H. Brighton. Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1(1):107–143, 2009.

[23] L. Harrison, D. Skau, S. Franconeri, A. Lu, and R. Chang. Influencing visual judgment through affective priming. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pp. 2949–2958. ACM, New York, NY, USA, 2013. doi: 10.1145/2470654.2481410

[24] T. Hofmann. Probabilistic latent semantic indexing. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 50–57. ACM, 1999.

[25] T. Iwata, S. Watanabe, T. Yamada, and N. Ueda. Topic tracking model for analyzing consumer purchase behavior. In *IJCAI*, vol. 9, pp. 1427–1432, 2009.

[26] X. Jin, Y. Zhou, and B. Mobasher. A unified approach to personalization based on probabilistic latent semantic models of web usage and content. In *Proceedings of the AAAI 2004 Workshop on Semantic Web Personalization (SWP'04)*, 2004.

[27] O. P. John, E. M. Donahue, and R. L. Kentle. The big five inventoryversions 4a and 54, 1991.

[28] D. Kahneman. A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, 58(9):697, 2003.

[29] D. Kahneman. *Thinking, fast and slow*. Macmillan, 2011.

[30] D. Keefe, M. Ewert, W. Ribarsky, and R. Chang. Interactive coordinated multiple-view visualization of biomechanical motion data. *IEEE transactions on visualization and computer graphics*, 15(6):1383–1390, 2009.

[31] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon. *Visual Analytics: Definition, Process, and Challenges*, pp. 154–175. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. doi: 10.1007/978-3-540-70956-5_7

[32] D. D. Loschelder, R. Trötschel, R. I. Swaab, M. Friese, and A. D. Galinsky. The information-anchoring model of first offers: When moving first helps versus hurts negotiators. *Journal of Applied Psychology*, 101(7):995, 2016.

[33] A. M. MacEachren. Visual analytics and uncertainty: Its not about the data. 2015.

[34] H. A. Meier, M. Schlemmer, C. Wagner, A. Kerren, H. Hagen, E. Kuhl, and P. Steinmann. Visualization of particle interactions in granular media. *IEEE transactions on visualization and computer graphics*, 14(5):1110–1125, 2008.

[35] D. Mimno and A. McCallum. Topic models conditioned on arbitrary features with dirichlet-multinomial regression. *arXiv preprint arXiv:1206.3278*, 2012.

[36] A. Nussbaumer, K. Verbert, E.-C. Hillemann, M. A. Bedek, and A. Dietrich. A framework for cognitive bias detection and feedback in a visual analytics environment. In *Proceedings of the European Intelligence and Security Informatics Conference*, pp. 1–4. IEEE, 2016.

[37] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.

[38] M. D. Plumlee and C. Ware. Zooming versus multiple window interfaces: Cognitive costs of visual comparisons. *ACM Trans. Comput.-Hum. Interact.*, 13(2):179–209, June 2006. doi: 10.1145/1165734.1165736

[39] J. C. Roberts. State of the art: Coordinated multiple views in exploratory visualization. In *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*, pp. 61–71, July 2007. doi: 10.1109/CMV.2007.20

[40] M. E. Roberts, B. M. Stewart, and E. M. Airoldi. A model of text for experimentation in the social sciences. *Journal of the American Statistical Association*, 111(515):988–1003, 2016.

[41] M. E. Roberts, B. M. Stewart, and D. Tingley. stm: R package for structural topic models. *R package version 1.2.1*, 1, 2017.

[42] M. E. Roberts, B. M. Stewart, D. Tingley, C. Lucas, J. Leder-Luis, S. K. Gadarian, B. Albertson, and D. G. Rand. Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4):1064–1082, 2014.

[43] D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. A. Keim. The role of uncertainty, awareness, and trust in visual analytics. *IEEE transactions on visualization and computer graphics*, 22(1):240–249, 2016.

[44] J. M. Tien. Big data: Unleashing information. *Journal of Systems Science and Systems Engineering*, 22(2):127–151, 2013.

[45] A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making*, pp. 141–162. Springer, 1975.

[46] P.-L. Yu. *Multiple-criteria decision making: concepts, techniques, and extensions*, vol. 30. Springer Science & Business Media, 2013.