HindSight: Encouraging Exploration through Direct Encoding of Personal Interaction History

Mi Feng, Cheng Deng, Evan M. Peck, Lane Harrison

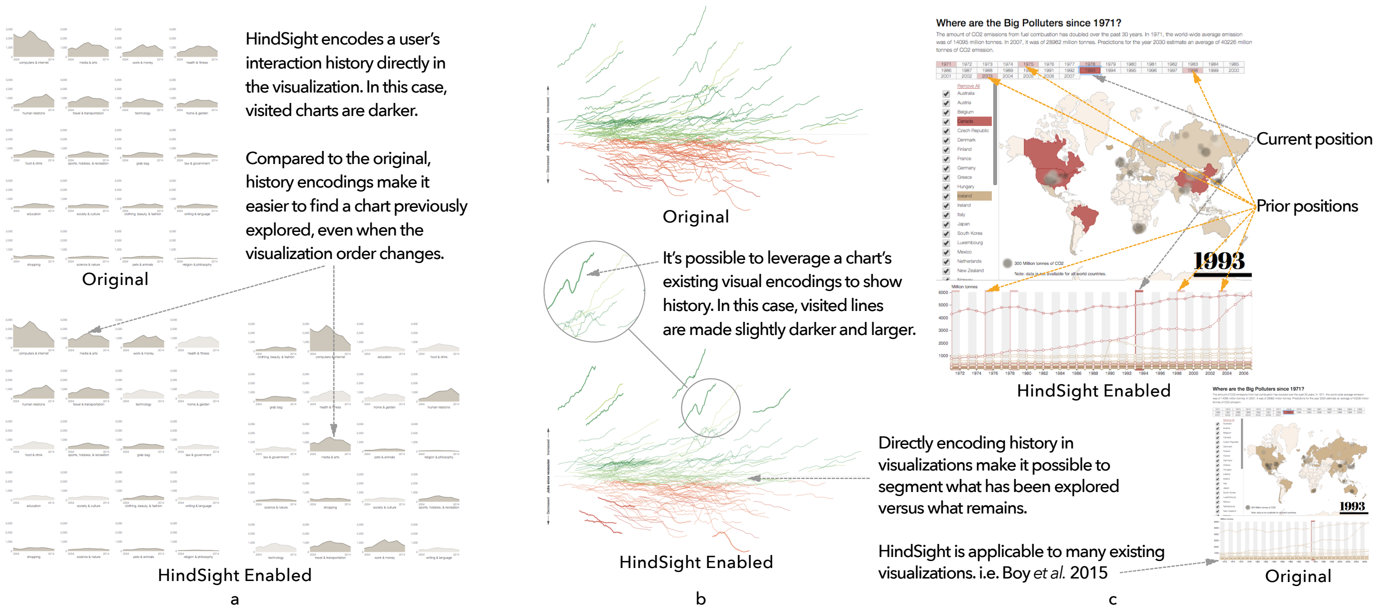

Fig. 1: Visually encoding a user's interaction history – a technique we call "HindSight" – can be easily implemented in many existing visualizations and is shown to significantly impact both exploration and insights. Here we show the three visualizations from our experiment, encoding interaction history through: a) chart opacity, b) line width and opacity, c) color (red highlighting), and "shadows" of previous marker positions.

Abstract— Physical and digital objects often leave markers of our use. Website links turn purple after we visit them, for example, showing us information we have yet to explore. These "footprints" of interaction offer substantial benefits in information saturated environments – they enable us to easily revisit old information, systematically explore new information, and quickly resume tasks after interruption. While applying these design principles have been successful in HCI contexts, direct encodings of personal interaction history have received scarce attention in data visualization. One reason is that there is little guidance for integrating history into visualizations where many visual channels are already occupied by data. More importantly, there is not firm evidence that making users aware of their interaction history results in benefits with regards to exploration or insights. Following these observations, we propose *HindSight* – an umbrella term for the design space of representing interaction history directly in existing data visualizations. In this paper, we examine the value of HindSight principles by augmenting existing visualizations with visual indicators of user interaction history (*e.g.* How the Recession Shaped the Economy in 255 Charts, NYTimes). In controlled experiments of over 400 participants, we found that HindSight designs generally encouraged people to visit more data and recall different insights after interaction. The results of our experiments suggest that simple additions to visualizations can make users aware of their interaction history, and that these additions significantly impact users' exploration and insights.

Index Terms—Visualization, Interaction, History.

1 Introduction

During exploratory data analysis (EDA), people navigate through unseen data for an indeterminate amount of time until an unknown insight is discovered. As a result, EDA aligns with some of the fundamental goals of information visualization. Data Exploration is generally defined in the context of scientific workflows, yet it is quickly becoming a part of peoples' day-to-day lives through news organizations and broadly accessible analysis tools.

Exploration takes time, however, creating a tension with our biological capacity for memory – a tension that is not supported by the visualization itself. Our memory's capacity to remember recent interactions is severely limited in both amount and decay [21, 18]. As a result, even when a visual design is aligned with our perceptual abilities, we struggle to remember and track parts of the data we have encountered, creating a barrier to exploration and engagement. These limitations suggest that a refinement of visualization techniques to support memory in interactive contexts may have broad impact in supporting exploratory data analysis.

- *Mi Feng, Worcester Polytechnic Institute. mfeng2@wpi.edu.*
- *Cheng Deng, Worcester Polytechnic Institute. cdeng@wpi.edu.*
- *Evan M. Peck, Bucknell University. evan.peck@bucknell.edu.*
- *Lane Harrison, Worcester Polytechnic Institute. lane@cs.wpi.edu.*

Manuscript received 31 March 2014; accepted 1 August 2014; posted online 13 October 2014; mailed on 4 October 2014.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

The call to support *history* operations in data visualization is not new. Many systems leverage formal representations of visualization state to capture and analyze scientific provenance [4, 16, 14]. Shneiderman identified *history* as an important visualizaton task to "allow users to retrace their steps" [28]. Gutwin realized Shneiderman's hypothesis, showing that indicators of exploration history helped users identify which parts of the data they have seen [10, 11, 12]. Collaborative analysis has also been a focus, where users are shown a history of operations from their collaborators to support situational understanding [1, 19]. Despite these advances, interaction history is not common in visualization systems today. One reason for this scarcity is that there is currently little guidance on how interaction history can be incorporated into the visualization itself. More importantly, however, there is little evidence for the possible benefits making users aware of their history, beyond supporting a user's ability to retrace their steps.

To uncover new opportunities in this space, we applied Wexelblat and Maes' interaction history framework [36] to the current state-of-the-art in visualization. Wexelblat and Maes identified six design properties – *proxemic vs. distemic, active vs. passive, rate/form of change, degree of permeation, personal vs. social, kind of information* – that can be used to characterize interaction history systems, or in this case, shed light on unexplored regions of the design space. We focus on two dimensions that expose a hole in the current design space – how history is directly tied to an object (*degree of permeation*) and whether history represents personal or group activity (*personal vs. social*).

As a direct result of this analysis, we propose **HindSight – a representation of personal interaction history that directly encodes interaction history as a visual variable on the data.** At its most basic level, HindSight modifies the saliency of data after a user engages with it, leaving visual markers of interaction history. Given an indication of what they have visited, users can quickly segment what parts of the data they have explored as well as what remains unexplored– using their perceptual system rather than their memory. The technical barrier of integrating HindSight into visualizations is low, requiring only simple modification to existing visualization infrastructure.

Direct encoding of interaction history on data has potential benefits that align with aspects of Shneiderman's arguments for direct manipulation [27]: increased visibility of object and actions, for example, or rapid and incremental actions with immediate feedback. Direct encoding puts interaction history right in front of the user, supporting visual recognition of previous interactions rather than relying on recall, short-cutting the mental translation of history information. Compared to indirect history encoding techniques common in visualization research [4, 14], direct encoding doesn't require users to process spatially separate regions to relate history information back to the data.

Given these observations, we hypothesized that the combination of direct encoding and personalized histories in HindSight would positively impact user behavior during exploratory analysis. To test our hypotheses, we applied HindSight to three visualizations, analyzing exploration behavior during interaction, as well as user-reported insights after exploring the visualization. Our cases include:

- *"The Rise and Decline of Ask MetaFilter" by Jim Vallandingham* ($N = 92$): 16 line charts of topic trends over time at MetaFilter that can be reordered by Count or Name.
- *"How the Recession Reshaped the Economy, in 255 Charts" by the NYTimes* ($N = 116$): a scatterplot of 255 line charts showing how jobs have changed across industries over the past 10 years.
- *"Where are the Big Polluters since 1971" by Jeremy Boy* ($N = 206$): a coordinated view map and line graph showing CO2 emissions that can be filtered by year or country [5].

In controlled experiments of over 400 participants, we found that HindSight designs encouraged people to visit more data and recall different insights after interaction. These results illustrate that the long-standing design principles developed by visualization research– principles that allow us to effectively map data to visual variables– can also be used to encode interaction, allowing us to leverage our perceptual system in interactive exploration and sensemaking.

2 Background

Interaction becomes a key mechanism in exploratory data analysis when the size or complexity of the data eclipse what the visual display can handle [28]. To this end, research has historically focused on interaction techniques that empower users to effectively reveal and reconfigure data in visualization systems. More recent work addresses the challenges of supporting user exploration and their awareness in the information foraging process. We describe several seminal results and research threads in this area, focusing on how they shape our contributions.

2.1 Wexelblat and Maes' Interaction History Framework

Objects are *history-rich* if they contain "historical traces that can be used by people in the current time" (p. 270, [36]). In the physical world, we note the wear on a tool to help us understand how it has been gripped in the past, or observe footprints in the snow to help us see areas that have previously been already explored. Embedding history rich objects into the digital realm enables people to either leverage their own experience that they have accumulated over time, or leverage the combined experience of people who have interacted in the same space. Citing results from Pirolli and Card, Wexelblat and Maes argue that without interaction history we are "forced to become information foragers over and over again" [36, 25].

Wexelblat and Maes describe six properties to articulate a design framework for interaction history: the extent to which people find a space to be transparent and easily understood vs. needing background or training to engage with it (*proxemic vs. distemic*), the degree of effort needed to record history (*active vs. passive*), the degree to which an object is changed by history (*rate/form of change*), the extent to which history is directly tied to an object or recorded separately (*degree of permeation*), whether history is tied to an individual or a group (*personal vs. social*), and finally the information we choose to represent history (*kind of information*). Each dimension of these six properties will nudge user behavior as they engage or use their own histories.

Consider the interaction when we click a link on a webpage – an example of an information-rich environment. The link that I click (*high degree of permeation*) automatically (*passive*) turns purple (*form: color as history*), and indicates whether I (*personal*) visited the site or not (*kind of information, binary rate of change*). Contrast this interaction with how our browser represents visit history. Our browsing history is also automatically collected (*passive*), but contains more detailed information than the purple links (*kind of information: time, url, etc.*). However, seeing our visit history requires us to navigate to a history page that is spatially separated from the original data (*low degree of permeation*). This shift from a high to low degree of permeation enables focused views of our browsing history, but sacrifices the availability of that information by relegating it to a secondary display.

These design tradeoffs are critical to weigh when designing history-rich tools and have implications for guiding exploration or engagement in any information foraging task. In particular, the change in permeation from the previous example shifts the notion of history from "How did I get here?" to "Where have I been before?" and "What is left to explore?". In the next sections, we highlight the benefits of reframing history in this manner, and explore whether these same benefits can be translated to data visualization contexts.

2.2 Interaction History from HCI to Visualization

The direct encoding of interaction history has been studied in HCI since the early 90s, when Hill *et al.* proposed the notion of computational wear ('read wear' and 'edit wear') to display authorship history [15]. Alexander *et al.* later analyzed principles of wear mechanisms – in this case marks on the scrollbar – to return to previously edited regions of a document. They found that marking the scrollbar with interaction history decreased visitation time, was highly preferred by participants, and was scalable to a large number of marks [2].

Following these foundational papers, researchers in HCI have applied interaction history to support users in novel ways. Gutwin, for example, visualized the traces of multiple mouse-pointers in a collaborative system to make users aware of where other people were focusing

[10]. They found that a direct representation (or *high degree of permeation*) of interaction history (the pointer trail) was easy to understand, and helped users understand the context of their collaborators current actions. Bridging the gap from HCI into data visualization, Skopik and Gutwin, introduced the notion of "visit wear" in the context of fish-eye pointers [31]. Using visual indications of history, they show that users were more readily able to trace their previous steps. Building on this work, Gutwin and Anton examined the extent to which users could remember their path after information history was removed [11]. Gutwin also carried some of these findings back to HCI, by integrating a "recency cache" in a list-interface to improve revisitation [12].

Beyond this, however, we also hypothesize that directly encoding interaction history is useful beyond revisitation. As we will demonstrate, even the most simple indications of history not only benefit revisitation, even more so, they impact the exploration patterns and insights of users.

2.3 Interaction History in Visualization

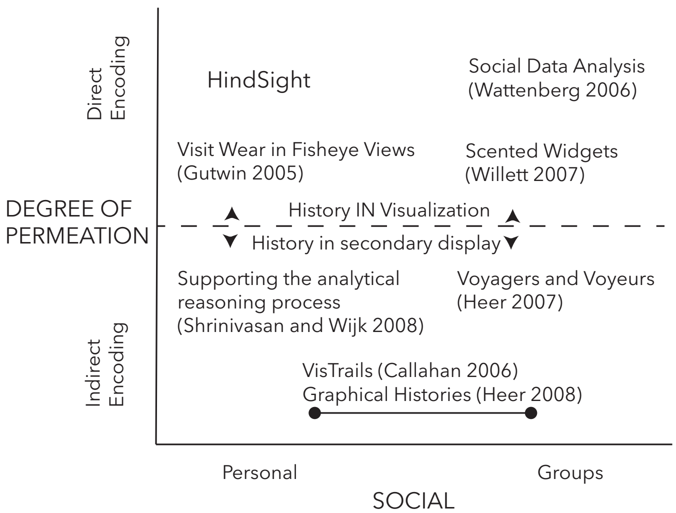

Fig. 2: With the exception of Gutwin's implementation of visit wear in fisheye views, research in data visualization has typically focused on three quadrants defined by Wexlblet and Maes. HindSight lies in the fourth– a direct encoding of personal interaction history.

Broadly, several threads of visualization research have focused on interaction history. In formal terms, Jankun-Kelly *et al.* propose a model for capturing the exploration process [16]. This work enabled several extensions, including VisTrails from Bavoil *et al.*, which used formal models of exploration to support scientific provenance in visualization systems [4], and Shrinivasan and van Wijk, who propose methods of transferring these provenance techniques to visual analytics [30]. However, the combination of direct, personal representations of history in the HCI community has not be suitably transferred and explored in the context of data visualization (see Figure 2).

2.3.1 Direct vs. Indirect encoding

In visualization, interaction history widgets typically use **indirect encoding** to represent history in secondary displays. This spatial separation from the data allows history to be expressed using a diverse palette of design characteristics that will not interfere with existing visual encodings. For example, textual or graphic representations of history may be spatially organized as a linear sequence of items, on continuous timelines, using branching metaphors, or in network diagrams [8, 14]. In addition, these views support a broad set of operations on historical information such as navigation, editing, annotation, searching and filtering, and exporting [30]. For a more thorough examination of these displays, see [14].

Outside of Gutwin's "visit wear" study, examples of visually encoding interaction history directly onto the data are more difficult to come by. Since interaction is represented in the same space as the data, the design space is constrained to visual features that are separable from the visual encoding. However, direct encoding of interaction history on data has clear usability benefits because it situates history signifiers directly onto the data. For example, Willet's *Scented Widgets*, which places small data visualizations next to interface widgets to guide exploration, found that users exploring unfamilar data make up to twice as many unique discoveries [33]. Instead of relegating interaction history to a secondary display that requires a mental translation, direct encoding leverages preattentive processes to spatially put interaction history next to or on top of the data itself.

2.3.2 Personal History vs. Social History

A second distinction we make is the use of history to communicate personal interactions with the data or group-driven interactions with the data. While most work in this space has focused on facilitating collaboration, we believe that directly encoding interaction history can improve *personal data exploration* with a fraction of the overhead.

History-focused interface widgets in data visualization typically appear in the context of asynchronous collaboration [1], or are shown indirectly through secondary displays [14]. A relevant example similar to our proposed work is Wattenberg and Kriss [19] who, when describing the visual encodings used in NameVoyager, briefly mention directly encoding personal interaction histories (p. 556):

> *"color by history" ...causes any visited series to appear in gray... We refer to this as "road-less-traveled navigation": Instead of using previous visits as a cue to importance, as in traditional social navigation interfaces, we treat it as a cue to staleness and hope to draw a user's eye to new territory, thus suggesting a unique perspective to each user.*

We propose that this concept can be broadened into a general design principle for interactive data visualizations: directly encoding personal interaction histories, or HindSight. In the context of exploratory data visualizations and in contrast to indirect displays of history which capture a "moment in time", encoding history directly on the data frees users to explore new spatial organizations without losing context. We hypothesize that HindSight-inspired techniques will encourage personal exploration of data and yield benefits such as higher levels of engagement, more systematic exploration, and as a result, more diverse insights about a particular dataset. While we have included an experiment that targets these measures, we first discuss the design process of building interaction history directly into existing visualizations.

3 HINDSIGHT DESIGN PROCESS

The core idea of HindSight is that designers can architect visualizations not only by visually encoding data, but also by encoding their users' interactions in the visualization itself. In this section, we pose questions for designers when they are considering to apply HindSight – *how do we define history, how do we represent history, and is it worth it?* – and share the principles we have developed while applying HindSight to a range of existing visualizations.

3.1 What type of history is important to this visualization?

As we mentioned in the previous section, HindSight shifts our perspective of history from "How did I get here?" to "Where have I been before?" and "What is left to explore?". As a result, HindSight may be most beneficial for visualizations in which exploration is a design goal. For example, when interactive news visualizations reveal important context only after users hover over data, encouraging exploration may lead to more nuanced insights that complement the story.

On the other hand, HindSight is less suitable when it is important for users to retrace their steps. Since spatial encodings are likely already in use by a visualization, it is not able to represent sequence data without interfering with the existing design. While we see this as the primary limitation of direct encoding, designers must generally make informed decisions about framing the user's mental model of history.

What data entities best represent a 'unit' of history?: Since we can refer to data at various levels of abstraction in a graph (*e.g.* chart-level vs. data-level), it is important to carefully weigh the entities we choose when applying HindSight. For example, in the small multiples visualization in Figure 1.a, we could consider interaction with each chart as meaningful (encoding history at the chart level) or we could consider interaction *within* each chart to be meaningful (for example, highlighting explored regions of the area graph). In this case, because chart reordering was a core interaction mechanism in the visualization, we encoded HindSight at the chart level, enabling visited charts to remain salient even as the data is reorganized. Additionally, encoding HindSight at the chart level encourages exploration of different topics in the MetaFilter visualization rather than secondary trends within a single topic. Choosing an appropriate level of coding for HindSight has the potential to unify exploratory goals with the capabilities of our perceptual system, making user history immediately available for further exploration and discovery.

What duration of user interaction represents *meaningful* interaction?: Interaction history is dynamic. Users may visit charts multiple times, or accidentally visit a chart when en route to another. In our initial pilots, we found that triggering a "visit" immediately was not ideal, whereas a short delay (*i.e.* 500ms) led to more predictable results. While definitive guidance on timing is beyond the scope of this paper, a general principle is to delay for long enough that the visit is considered "intentional".

3.2 Which visual channels should be used?

One broadly applicable way of encoding interaction history is changing the opacity of the element after interaction. Opacity is just one of many visual channels that may be used, however. Designers should be aware of the relative efficacy of visual channels such as position and color, as well as concepts such as integral and separable channels [34]. A poor choice of encoding– significantly increasing line size, for example– may severely interfere with the other data in the visualization, especially as the user spends more time interacting. Here we give high-level guidelines for selecting visual channels based on the current design of the visualization and the goals of the designer. We categorize three use-cases for applying HindSight encodings:

- *augmentation:* when unused visual channels are available, augment existing data with additional visual encodings to the target visualization to show interaction history. For example, we identified opacity as an unused visual channel that could be used to encode interaction history in the area charts shown in Figure 1.a.
- *addition:* There is often empty space available in a visualization that can be repurposed for interaction history. When history can be represented in unused regions of a chart, modify unoccupied visual layers with interaction data. Transforming the background of a scatterplot into a heat map, for example, could clearly communicate regions of the plot that were already explored.
- *adaptation:* when no visual channels are available but displaying history is deemed important, adapt the target visualization to show interaction history by modifying visual channels that are already occupied by data. If there are no available visual channels, existing encodings can be manipulated to represent interaction history. Note that this approach runs the risk of undermining the perceptual benefits of some visual encodings.

How important is interaction history to the goals of the visualization?: One helpful way of assessing design tradeoffs is to consider interaction history as an additional data attribute. Weighing interaction history's impact on understanding in relation to other data attributes enables designers to use the *principle of importance ordering* to map both data *and* interaction history onto visual variables. For example, encouraging exploration in a complex news visualization may be critical enough to the success of a graph that representing interaction using color will yield stronger results than using that same channel to encode an additional data dimension.

Similarly, in The New York Times "255 Charts" visualization, there are many visual variables which could be used to encode history (see Figure 1.b). Line charts are the primary encoding in this visualization, representing the most important information – the financial growth of the particular industry. Color is also used on each line chart to show whether a particular industry has grown (green) or fallen (red). Since color is a redundant encoding, we may decide that the benefits of representing interaction history outweigh the benefits of aligning multiple visual channels with a single dimension of data.

However, assessing the importance of encoding interaction raises the inevitable question: what are the benefits? While prior work such as Gutwin *et al.* suggest that showing users where they've been can help when revisiting previously visited elements [10, 12], it is not clear from existing research whether making users aware of their interaction history impacts any other aspects of the exploration process. The duration of this paper, in particular our three experiments, are dedicated to examining this question.

4 Evaluation

The goal of our study was to determine the effect of directly encoding personal interaction history on the following factors:

- **exploration behavior:** how does HindSight impact exploration behavior such as number of charts visited, total time spent exploring the data, and patterns of exploration?
- **post-interaction insight:** how does HindSight impact the insights that people recall immediately *after* interacting with a visualization?

To this end, we used a between-subjects design to test HindSight principles in three different interactive data visualizations. Two were selected to vary in complexity and design, and the third was chosen to draw comparisons with recent work by Boy *et al.* [5] that evaluates exploration and engagement in visualization. In each visualization, we tested conditions with and without HindSight:

- **control:** we present an interactive visualization in its original form, removing only extraneous information
- **hindsight:** we apply a straightforward encoding of user's interaction history.

4.1 Procedure and Tasks

Participants were recruited through Amazon's Mechanical Turk (AMT) to participate in a maximum of one of our three studies. AMT is a crowdwork platform where "Workers" select from a range of available tasks, including research experiments [13, 20]. Each participant was randomly assigned to either the *control* (original-visualization) or *hindsight* (original with HindSight techniques) condition. Based on time data in pilot experiments, participants were paid $1.00 in order to exceed US Minimum Wage. All participants were shown a standard consent form before continuing.

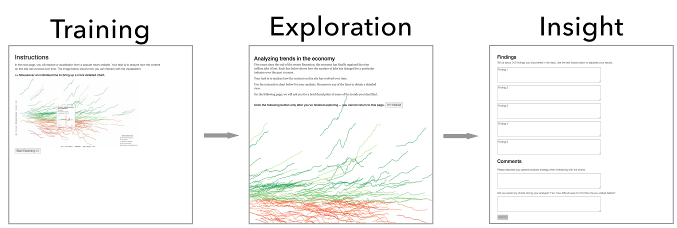

Fig. 3: In each of our three experiments, participants completed a training phase before heading to the exploration section. When they were finished exploring the interactive (no time limit), they moved to the insight section and were asked to describe their findings.

Our procedure consisted of three phases: *Training*, *Exploration*, and *Insight*. In the *Training* phase, we provided participants with an instruction page that briefly described their task and the interaction mechanisms in the visualization. For example, for the *metafilter* experiment participants were told:

> *In the next page, you will explore a visualization from a popular social media site. Your task is to analyze how the content on this site has evolved over time. The image below shows how you can interact with the visualization.*

In the *hindsight* condition, an extra sentence explained that visited charts would be made visually distinct, and an image showed HindSight being triggered.

Following training, the *Exploration* phase began with a paragraph that introduces participants to the visualization and their task. Participants were instructed that they may interact with the visualization without any minimum or maximum time limit. They were also reminded that after they finish, they would be asked to describe several of their findings. When participants finished exploring the visualization, they advanced to the *Insight* phase through a button press.

As a final step, participants entered the *Insight* phase. After the visualization was hidden, participants were instructed to describe 3-5 of their findings in individual text boxes. Additional text boxes were included to allow for more freeform comments about their experience.

4.2 Measures

Given the *Exploration* and *Insight* phases of the experiment, we draw on both quantitative and qualitative measures for evaluation. For quantitative exploration metrics, we build on work from Boy *et al.* [5], recording *visited* items and *exploration-time*. We also include the *revisit* metric from Gutwin *et al.* [10].

- *visited* : the number of unique charts that a person directly interacts with during exploration.
- *revisited* : the number of instances when a user interacts with a previously visited chart.
- *exploration time* : the total amount of time spent interacting with charts. We use this metric to try and capture *active* use of the visualization, mitigating when external distractions artificially inflate the time spent in the exploration phase.

For qualitative metrics, we referred to work by Saraiya *et al.* on analyzing insights from interactive data visualization [26]. We used faceted coding, where independent coders mark what elements of the visualization (*e.g.* a particular topic or year) appear in the comments.

- *mentions* : the number of times a chart is directly referenced in findings during the *Insight* phase of our experiment.

Finally, we asked participants to describe their general analysis strategy and to reflect on the difficulty of revisiting charts. We will draw on these open-ended comments to contextualize our findings.

4.3 Pilots, Analyses, and Experiment Planning

We conducted several pilot experiments using the *metafilter* visualization (fully described in Section 5) to help establish our measures and procedure. In response to concerns about the limitations of null hypothesis significance testing [7, 35], we model our analyses on recent visualization research that seeks to move beyond these limitations [6], primarily focusing on confidence intervals and effect sizes. Following Cumming [7], we compute 95% confidence intervals using the bootstrap method, and effect sizes using Cohen's d– which is the difference in means of the conditions divided by the pooled (*i.e.* both conditions') standard deviation. While we include significance testing and related statistics, it is with the intention of supplementing these analyses.

The results of our pilots showed some measures from the *Exploration* phase were non-normally distributed, according to a Shapiro-Wilk test. These measures include $exploration-time$, $visits$, and $revisits$, all of which were right-skewed with long tails. Because common transforms (*i.e.* log, square-root) did not cause a significant change in the Shapiro-Wilk result, we use the non-parametric Mann-Whitney test to compare the *control* and *hindsight* conditions.

Analyzing the findings left during the *Insight* phase, we turned to three independent coders and inter-coder reliability metrics. The coders were undergraduate students who had little-to-no visualization experience, and were not involved in this project. The coders annotated each comment by assigning tags to indicate the entities mentioned (*e.g.*, the social media topic mentioned). Fleiss' Kappa was calculated to measure the agreement among the three coders [9]. We took the majority agreement when 2 out of 3 coders agreed on all entities mentioned in a given comment. If all coders disagreed, the comment was discarded from analysis.

	metafilter	**255charts**	**storytelling**
control	44	57	99
hindsight	48	59	107
Total	*92*	*116*	*206*

Table 1: We tested HindSight using a between-subjects design on three visualizations. The table above shows participant numbers for each visualization, which were determined by running effect size and power analyses on pilot studies.

In order to ensure our experiments included enough participants to reliably detect meaningful differences between the *hindsight* and *control* conditions, we conducted effect size and statistical power analyses. Specifically, we used pilot experiments to estimate the variance in our quantitative measures, and combined these with the observed means to approximate how many participants were needed. This procedure was repeated for each of our three experiments (see Table 1).

5 Visualization 1: Metafilter

We first chose to apply HindSight to a relatively simple interactive visualization. Many interactive visualizations people encounter on a day-to-day basis consist of a few views and simple interactions such as clicks and hovers to uncover more information. From an experiment control perspective: a simple visualization should lead to less variance between participants, making it more likely to detect reliable effects.

After evaluating several alternatives, we selected an interactive small-multiples area chart - *The Rise and Decline of Ask MetaFilter*. Obtained from a popular data visualization blog [32], it depicts posting trends across topic categories in a community weblog. There were twenty area charts in total. Mousing-over any chart brought up a cursor at the corresponding x-axis (time) location on all other charts, and a toggle button allowed the charts to be reorded either by alphabetical order or post count. We used a between-subjects design with the following conditions:

- **control:** the original design of the visualization
- **hindsight:** interaction history was encoded through a small change in opacity. If a chart was visited for more than 500 milliseconds, it received a slight increase in opacity and became more salient in the visualization.

The original visualization and HindSight encoding can be seen in Figure 1.a. Pilot experiments with $metafilter$ coupled with a power analysis (see 4.3) indicated that at least 76 participants would be needed to detect a large effect (*e.g.* a difference of 3+ charts visited).

5.1 Results

We recruited 92 participants through AMT for this experiment. Through random assignment, we gathered 48 responses for the *hindsight* condition and 44 responses for the *control* condition.

5.1.1 Behavior/Interaction Analysis

Shown in Figure 4.d, the average participant in the *hindsight* condition visited more area charts ($M = 9.4$ visits 95% CI $[7.5, 11.3]$) than those in the *control* condition ($M = 5.4$ $[4.4, 6.5]$). Given the upper and lower limits of the confidence intervals, the average participant visits at least 1 additional chart with *hindsight*, and up to 7 more ($d = 0.75$ $[0.34, 1.11]$). There was little difference in participants' time spent interacting with charts in the *hindsight* condition ($M = 43.4$ seconds $[32.6, 65.6]$) compared to the *control* condition ($M = 36.1$ $[25.7, 51.6]$, $d = 0.15$ $[-0.27, 0.53]$).

Qualitative analysis of visits indicate that participants in the *control* condition tend to focus on the top region of the chart (*i.e.* the top two rows). While this trend held for HindSight, additional visits were more evenly spread across the entire chart (see Figure 4.a).

5.1.2 Insight Analysis

92 participants left a total of 363 findings in the Metafilter experiment. Following the methodology in Section 4.3, three people independently coded each finding to determine whether a specific posting topic was referenced. A statistical analyses of the 363 comments indicate strong

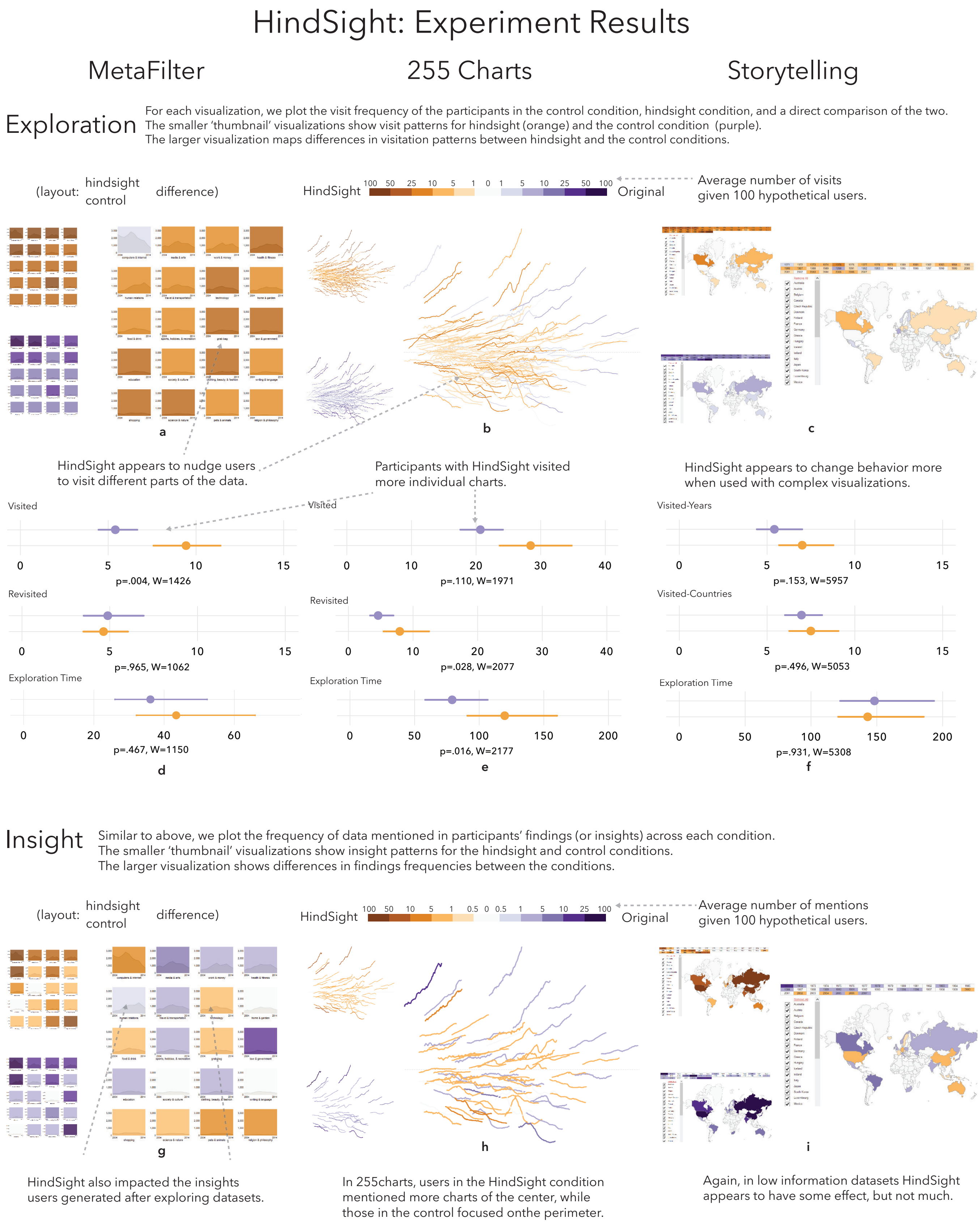

Fig. 4: Experimental results comparing basic HindSight encodings with three visualizations. Exploration metrics suggest that HindSight generally encourages more exploration and nudges users towards investigating different parts of the data. While insight metrics indicate that quantity of comments users produced are similar, HindSight impacts the diversity of insights generated. Note that in the map legends, the average number of visits for a given chart are low (*i.e.* 0.6); we scale to natural frequencies using hypothetical participants to aid comparison.

agreement for the posting topics mentioned ($\kappa = 0.89$). For posting topic, there were 362 comments with majority agreement (*i.e.* at least two out of three coders agreed), and one comment with complete disagreement (this was discarded).

Qualitative analysis of posting topics mentioned in findings for the *metafilter* experiment indicate that participants in the *hindsight* condition overwhelmingly referenced the bottom region of the chart-grid more often than in the *control* condition (see Figure 4.g). In contrast, findings from participants in the *control* condition appear to tend are more evenly distributed across the chart-grid.

We also analyzed the number of unique charts referenced by a minimum number of participants (analyzed by 1, 2, ..., up to 5). From this data, we see very little difference between the *hindsight* and *control* conditions (see Figure 4.j). This suggests that in the *metafilter* visualization, HindSight did not encourage more findings overall, but different findings. Turning to the open-ended comments, this change in behavior may be the result of HindSight enabling more systematic exploration strategies. When asked about their approach, participants in the *hindsight* condition often responded with a clearly defined strategy similar to the following: *"I looked at every chart one by one, sorted by 'Count'"*.

6 VISUALIZATION 2: 255 CHARTS

The datasets people encounter on a daily basis are often larger and more complex than the *metafilter* visualization discussed in the last section. For this reason, we examined whether HindSight impacted behavior with more advanced interactive visualizations.

Towards this goal, we adapted a popular interactive visualization from The New York Times titled, "How the Recession Shaped the Economy, in 255 Charts" [3]. Shown in Figure 1.b, the 255*charts* condition includes 255 line charts distributed across the viewport in a scatterplot-like fashion. Each line in 255*charts* represents how a particular industry of the US Economy – Home Health Care Services or Air Transportation, for instance – grew or declined from 2004 to 2014. Mousing-over an industry's chart brought up a detailed view showing specific values, years, and industry information.

The original article included multiple stages with animations, transformations, and annotations, which the user controlled through scrolling. To better control our experiment, we isolated the part of the visualization where users are given the opportunity to freely explore the charts. We also repositioned the introductory explanation to avoid obscuring any part of the data.

The open-ended nature of 255*charts* coupled with its large data size makes it an ideal candidate for examining how HindSight impacts exploration with more complex data. Again, we used a between-subjects design with the following conditions:

- **control:** the design of the visualization as described above.
- **hindsight:** if a line chart was visited for more than 0.5 seconds, it received a slight increase in width and opacity to represent interaction history.

Running a power analysis on pilot experiments of 255*charts* (see Section 4.3) indicated that at least 102 participants would be needed to reliably detect a large effect (*e.g.* a difference of 5 or more charts visited).

6.1 Results

We recruited 116 participants through AMT for this experiment. Through random assignment, we gathered 59 responses for the *hindsight* condition and 57 responses for the *control* condition.

6.1.1 Behavior/Interaction Analysis

Shown in Figure 4.d, the average participant in the *hindsight* condition visited more charts ($M = 28.4$ visits $[23.3, 34.2]$) than those in the *control* condition ($M = 20.7$ $[17.9, 24.2]$). Given the limits of the confidence intervals, the average participant will at least visit the same number of charts with HindSight, and up to 16 more ($d = 0.44$ $[0.12, 0.75]$). In addition, the average participant in the *hindsight* condition appears to revisit more charts ($M = 7.9$ *visits* $[5.6, 13]$) than the *control* condition ($M = 4.6$ $[3.2, 6.7]$, $d = 0.32$ $[-0.06, 0.57]$). Similarly, we see that the average participant in the *hindsight* condition may spend more time interacting with charts ($M = 119.7$ *seconds* $[93.1, 166.9]$) compared to the *control* condition ($M = 79.1$ $[59.7, 110.6]$, $d = 0.36$ $[-0.02, 0.68]$).

Qualitative analysis of visits in the 255*charts* experiment indicate that participants in the *hindsight* condition tend to focus more attention than the *control* condition on industries in the center of the visualization, where the data density is at its highest (Figure 4.b). In contrast, participants in the *control* condition appear to focus on charts in periphery, particularly the top left and bottom right.

6.1.2 Insight Analysis

116 participants recorded a total of 492 findings in the 255*charts* experiment. Three people independently coded each finding to determine whether a specific industry was referenced. The statistical analyses of the 492 comments indicate moderate agreement for the industry mentioned ($\kappa = 0.59$). For the industry mentioned, there were 444 comments with majority agreement (*i.e.* at least two out of three coders agreed), and 48 comments with complete disagreement (these were discarded).

Qualitatively, the maps showing referenced findings (Figure 4.h) indicate trends that mirror behavioral patterns. When compared to the *control* condition, participants in the *hindsight* condition were more likely to reference industries in their findings that were spatially in regions of high data density. This is also reflected in the map of industries participants visited (Figure 4.b). We also analyzed the number of unique charts mentioned by at least 1 participant, 2 participants, etc (see Figure 4.j). While most findings still gravitated towards a handful of charts, in contrast to *metafilter*, the trends in Figure 4.h and Figure 4.j suggest that HindSight not only encouraged a different set of findings, but more diverse set of findings. These benefits were reflected in open-ended comments: *"... it was relatively easy to find the chart that I wanted to see again because it had been changed to a bolder and darker line which is a great feature seeing as how there are a whole bunch of lines mixed up together."*

It's possible that the increased data in 255*charts* amplified the effect of HindSight in comparison to *metafilter*, however, more experiments would need to confirm this hypothesis.

7 VISUALIZATION 3: STORYTELLING

We turn to existing research in exploratory data analysis to choose our third visualization. In a recent study, Boy *et al.* examined the impact of storytelling techniques across several quantitative measures of user engagement [5]. While we adapt several of the measures they use throughout our experiments (see 4.2), we also replicate one of the conditions of their experiment, thanks to their releasing the study's experiment materials online.

The *CO2 Pollution Explorer* was one of the primary interactive visualizations in Boy *et al.* [5]. Consisting of a world map, a year selector, and a line chart showing a country's pollution over time, this interactive visualization allows users to compare pollution from a particular country across several decades (see Figure 1.c). User interactions included the ability to hover on a country to highlight the corresponding trend on the line chart, and click on a year to update the map and year marker in the line chart. As in previous visualizations, we used a between-subjects design with the following conditions:

- **control:** the design of the visualization as described above.
- **hindsight:** if a country shown in either the map, list, or line chart view was visited for more than 0.5 seconds, its opacity increased slightly in each view. Similarly, the color of a visited year button changed from gray to light red, and a light red border also appears in the line chart indicating the year's range.

Pilot experiments with *storytelling* coupled with a power analysis (see Section 4.3) indicated that at least 177 participants would be needed to reliably detect a medium effect (*e.g.* a difference of 5 or more years or countries visited). In contrast to the *metafilter* and 255*charts* conditions, the variance in behavioral metrics in the *storytelling* pilots was higher, leading to a larger number of participants needed.

	HindSight, 2016		Boy *et al.*, 2015	
Metrics	**Exp**	**Control**	**Control**	**Exp**
meaningful interaction	54.8	48.6	44	33
meaningful hover	22.8	19.8	35	26
meaningful click	32.2	28.8	8	6
semantic - inspect	14.6	13.6	26	17
semantic - connect	8.1	6.2	10	8
semantic - select	21.1	19.8	5	3
semantic - explore	8.7	6.6	3	2
semantic - filter	2.44	2.39	0.2	0.1
exploration time	140.3	148.2	108.8	54

Table 2: Meta-analysis of HindSight applied to one of the primary visualizations from Boy *et al.*, 2015. While the control condition in the present experiment led to generally higher results, HindSight appears to reliably outperform the other conditions– past and present.

7.1 Results

We recruited 206 participants through AMT for this experiment. It took approximately one day to gather all responses. Through random assignment, we gathered 107 responses for the *hindsight* condition and 99 responses for the *control* condition.

7.1.1 Behavior/Interaction Analysis

The behavioral metrics for the *storytelling* visualization differ slightly from the previous graphics. Specifically, instead of reporting "visited" items, the original work from Boy *et al.* distinguishes between years visited and countries visited. We adopt their approach here.

The visit quantities for years and countries were largely the same. The average participant in the *hindsight* condition visited a similar number of countries ($M = 7.2$ $[5.9, 8.7]$) as those in the *control* condition ($M = 7$ $[6, 8.1]$, $d = 0.04$ $[-0.24, 0.3]$). Participants in the *hindsight* condition also appeared to visit a similar number of years ($M = 6.7$ $[5.6, 8.5]$) as the *control* condition ($M = 5.4$ $[4.4, 6.9]$, $d = 0.19$ $[-0.09, 0.47]$). Qualitatively, the maps showing which years and countries participants visited were largely similar (see Figure 4.c), particularly when compared to the differences in the *metafilter* and 255 maps (Figure 4, a and b). In terms of timing, participants in the *hindsight* condition spent roughly the same amount of time in the exploration phase of the experiment ($M = 140.3$ $[117.4, 180.7]$) compared to the *control* condition ($M = 148.2$ $[123.3, 188.6]$, $d = -0.05$ $[-0.32, 0.22]$).

Meta Analysis. In their study, Boy *et al.* analyzed additional metrics such as hover and click interactions. We also tracked these metrics in our experiment to facilitate a meta-analysis with the results of [5]. While the raw data from Boy *et al.* was not available, we carefully inferred means from the confidence interval plots in [5]. Their experiment hypothesized that the addition of storytelling prompts would increase several of these measures. However, they found the opposite occurred – users in the experiment condition generally interacted less with the visualization. In contrast, we found that HindSight produced small gains across the board in identical behavioral metrics when compared to our control (Table 2).

7.1.2 Insight Analysis

206 participants left a total of 831 findings in the *storytelling* visualization. Coders labeled two dimensions– whether a specific country or year was referenced in the comment. A statistical analyses of the 831 comments indicate strong agreement for the country mentioned ($\kappa = 0.87$), and substantial agreement for the year mentioned ($\kappa = 0.76$). For the country and year mentioned respectively, there were 821 and 826 comments with majority agreement (*i.e.* at least two out of three coders agreed), and 10 and 5 comments with complete disagreement. The latter were discarded.

Qualitatively, behavioral visitation trends did not transfer to year or country references in the findings (Figure 4.i). While countries of increased interest in *hindsight* appear to reflect the most significant stories in the data, the effect is not strong enough to make more generalizable claims. We also found that participants in the *control* condition referenced a more diverse set of years from the visualization while participants in the *hindsight* condition focused their findings on major trends in the data. We will contextualize these findings in the discussion section.

8 Discussion

HindSight's simple encoding of interaction history generally changed users' behavior as well as the details that they remembered. In both *metafilter* and 255*charts*, we saw significantly increased interaction with data. As indicated in the insight maps (Figure 4), users also reflected on a more diverse set of findings with HindSight, although they identified dominant outliers and trends less often.

In the *storytelling* condition, we noticed slightly different results. There were few differences in the amount of data explored (*e.g.* visited countries, years, and exploration time). We did see, however, a small improvement in most behavioral exploration metrics recorded in the original study (Table 2). This change raises the question: when should we expect techniques like HindSight to cause a noticable change in user performance?

The results of these experiments generally confirm our hypothesis that subtle indications of interaction history impact user behavior in data visualizations, while the degree of impact may vary across different visualizations, e.g., 255*charts* versus *storytelling*. Our goal now is to discuss the implications of these findings more broadly and make recommendations for the use and development of HindSight.

8.1 Benefits on Exploration, Engagement and Insights

We found that HindSight generally encourages people to interact with more data. We also observed that HindSight impacts the findings that users report after viewing a visualization – nudging users towards areas that are typically unexplored in a visualization (for example, areas of high data density). While it is difficult to make value judgements about exploration patterns, our findings suggest that at the very least, HindSight redirects attention to *different* data. Whether more interaction is a good thing – for instance leading to a deeper understanding of the dataset as a whole – remains an open question for future research.

The quantitative results suggest that the effects of HindSight may be amplified by larger, more complex data visualizations. This observation is supported by the comparison of results between *metafilter* and 255*charts* visualizations. As the amount of data between *metafilter* and 255*charts* increased (20 to 255), the effect of HindSight on exploration time also increased (see Figure 4.e).

We also believe that HindSight improves levels of the sustained attention on a visualization, which is one marker of engagement. This raises the question: Why does HindSight nudge exploration behavior?

One plausible explanation is that HindSight helps negate attentional biases related to the spatial placement of data on a page by making people more aware of their own navigational patterns. As an example, the visit spatial pattern of the control condition in *metafilter* appears to mirror the typical F-shaped gaze patterns observed in eye-tracking studies of product websites [22]. In these website studies, users typically explored the top rows and down the left side of a webpage, avoiding the center. While some form of top-to-bottom bias still holds for HindSight in the *metafilter* visualization, visit patterns and findings suggest users with HindSight engaged with the bottom row of charts much more frequently than in the control condition. Another possible explanation is that HindSight gamifies interaction by providing immediate visual feedback and anchor points from which users can systematically navigate complex data.

The *storytelling* condition is of note because we did not observe the same changes in behavior and insight. There were several factors that made the *storytelling* visualization unique, however– countries were not available to interact with due to limitations of the underlying dataset, and several participants commented that the animated pollution clouds interefered with their ability to select European countries. In the insight maps, country references were largely focused on just a handful of nations, suggesting that *storytelling* contained fewer significant insights that could be gleaned from the data. The regions of the map in which HindSight provoked the most findings tended to align with the major pollution contributors (Figure 4.i). These factors suggest that HindSight may help users more systematically navigate datasets where fewer insights are to be found. In other words, when considering techniques like HindSight, designers should ensure that their data contains many possible stories that may benefit from exploration (*i.e.* not just a few outliers).

Overall these results confirm that HindSight impacts user engagement and exploration patterns. As visualization research continues to add language and metrics that capture user interaction strategies (*e.g.* Ottley *et al.* [24]), techniques such as HindSight should be developed in parallel to help support the cognitive task of exploration in interactive visualizations.

8.2 Low Technical Barrier

The cost of implementation effort versus the added value to users is a tradeoff rarely discussed in visualization design. We see this dynamic as one of the core advantages of HindSight. HindSight can be applied to existing visualizations by adding just a few lines of code and without changing any technical infrastructure. For example, modifying the visual encoding of data in response to mouse behavior is a trivial change in dominant visualization libraries such as d3js. This enables designers to leverage the benefits of interaction history we have established without having to dramatically alter existing code bases (necessary for indirect coding approaches) or by adding server-side storage mechanisms (necessary for social applications). We envision future research targeting the long-term support of visualization navigation (*i.e.* beyond a single-session), similar to the topic of analytic provenance from the visual analytics community [23].

8.3 Design Tradeoffs

HindSight's direct encoding of interaction history, much like Gutwin's "visit wear", can be compared to the concept of direct manipulation as defined by Shneiderman [27] and following research. While changes in visual encoding occur passively, they are triggered by explicit actions. This encoding creates a continuous and dynamic indication of data of interest, allowing users to rapidly and incrementally tweak their interaction strategy.

As a result, some of the same advantages of direct manipulation outlined by Shneiderman and Plaisant can also be considered within the context of HindSight [29]. Immediate visibility of user actions a) results in reduced error rates, b) promotes usage by novices with minimal knowledge or instruction [27], and c) encourages exploration [29]. While we did not investigate error rates or visualization expertise, exploration benefits are reflected in our results. Looking forward, the concepts explored in direct manipulation (*e.g.* reversible actions) may serve as inspiration for future research related to HindSight.

We must also consider the constrained design space of directly encoding interaction history onto visualizations. HindSight's definition of history to this point has shifted from the traditional notion of "How did I get here?" to instead focus on "Where have I been before?" and "What is left to explore?'. In designs that already map several data variables to visual variables, identifying additional separable channels is difficult [34]. Over-representing history information, for example, may interfere with existing spatial encodings of data. While there is no silver bullet for design, the examples and principles we lay out in the design space (Section 3) are intended to help architects of interactive visualizations maximize benefit and minimize tradeoffs.

8.4 Expanding the HindSight Design Space

Returning to Wexelblat and Maes' interaction history framework, we can consider other unexplored facets of design within HindSight's core properties (direct encoding of personal interaction history):

- *rate/form of change:* rather than a binary encoding of visitation, we can map the accumulation of history to visual variables, both in terms of quantity as well as frequency.
- *proxemic vs. distemic:* encoding more complex interaction data is possible, however, it will increase the training needed to understand it. While our experiments targeted a broad audience, this tradeoff may be worth considering in analytical contexts.
- *active vs. passive:* a combination of explicit bookmarking operations with the implicit recording of visit data could yield benefits that merge users' goal-oriented thought processes with their natural exploration patterns.
- *kind of history:* encoding visit history benefits from simplicity both in terms of intuition and implementation. However, rich representations of interaction history will likely require a more diverse set of input, both in terms of operations (*e.g.* transformations on the data) and input sources (*e.g.* eye-tracking data).
- *personal vs. social:* similar to the spectrum of active vs. passive, there may be unknown benefits in considering a more nuanced continuum between personal encodings and encodings meant to enhance asynchronous collaboration.

8.5 Limitations

Although the validity of AMT workers has been analyzed and validated for visualization studies [13, 20], the population is incentivized differently than the average user visiting The New York Times and similar sites. It is not clear how these motivations may change exploration strategies or willingness to engage. In addition, our study captured a user's first exposure to HindSight and the results may not be generalizable to longterm use. Further research and deployment will clarify the impact of HindSight "in the wild".

We currently lack a precise language to discuss the strategic approaches that emerged as a result of HindSight usage. Current measures do not take into account the actual spatial layout of the visualization or the encodings used. While HindSight impacts behavior, more research in the strategies that people use when interacting with visualizations, *e.g.* Lam's interaction cost framework [17] or Ottley *et al.* [24], will allow us to better reason about the impact of interaction support tools, and enable designers more accurately assess the tradeoffs.

9 Conclusion

As visualization becomes more widely used by everyday people, research should focus on low-barrier interaction support techniques that can benefit people without expertise or training. We believe that HindSight offers an opportunity to do exactly that.

Building on preliminary evidence from Gutwin *et al.*, we used Maes and Wexelblat's interaction history framework to identify gaps in existing interaction history encoding approaches used in visualization. A direct encoding of personal interaction history not only is trivial to apply to many web-based visualizations, but as we discovered, can yield high benefits for the low cost. In three experiments, we found simple applications of HindSight techniques changed exploration behavior – increasing the amount of data covered and the range of insights articulated after encountering a visualization.

HindSight provides cognitive support for interaction through visual encodings, and yields benefits beyond enabling users to "retrace" previous steps. Our results suggest that HindSight may hold immediate benefits for practitioners. News organizations who are building expository visualizations similar to the designs we tested in our experiment may use HindSight to help encourage their users engage more deeply with the data presented. As visualization research continues to define and understand the interaction process, techniques like HindSight should be further developed and evaluated to ensure users have as much cognitive support for exploratory data analysis as possible [1].

Acknowledgments

We thank Dave Brown, Steve Franconeri, Martin Wattenberg, and others who gave excellent feedback on early versions of this work.

References

[1] M. T. A Sarvghad. Exploiting analysis history to support collaborative data analysis. *Graphics Interface*, pages 123–130, 2015.

[1]**To facilitate future work, all experiment materials, participant data, and analyses scripts are available online:** `https://github.com/wpivis/hindsight`.

[2] J. Alexander, A. Cockburn, S. Fitchett, C. Gutwin, and S. Greenberg. Revisiting read wear: analysis, design, and evaluation of a footprints scrollbar. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1665–1674. ACM, 2009.

[3] J. Ashkenas and A. Parlapiano. How the recession shaped the economy, in 255 charts. http://www.nytimes.com/interactive/2014/06/05/upshot/how-the-recession-reshaped-the-economy-in-255-charts.html, 2014 (accessed 2016-03-01).

[4] L. Bavoil, S. P. Callahan, P. J. Crossno, J. Freire, C. E. Scheidegger, C. T. Silva, and H. T. Vo. Vistrails: Enabling interactive multiple-view visualizations. In *IEEE Visualization, 2005.*, pages 135–142, 2005.

[5] J. Boy, F. Detienne, and J.-D. Fekete. Storytelling in information visualizations: Does it engage users to explore data? In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1449–1458, 2015.

[6] F. Chevalier, P. Dragicevic, and S. Franconeri. The not-so-staggering effect of staggered animated transitions on visual tracking. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):2241–2250, 2014.

[7] G. Cumming. *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge, 2013.

[8] 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*, pages 1663–1672, 2012.

[9] J. L. Fleiss. Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76(5):378–382, 1971.

[10] C. Gutwin. Traces: Visualizing the immediate past to support group interaction. In *Graphics interface*, pages 43–50, 2002.

[11] C. Gutwin and D. Anton. Does location come for free? the effects of navigation aids on location learning. *Technical Report HCI-TR-06-03: Computer Science Department, University of Saskatchewan*, 2006.

[12] C. Gutwin and A. Cockburn. Improving list revisitation with listmaps. In *Proceedings of the working conference on Advanced visual interfaces*, pages 396–403, 2006.

[13] J. Heer and M. Bostock. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 203–212, 2010.

[14] J. Heer, J. D. Mackinlay, C. Stolte, and M. Agrawala. Graphical histories for visualization: Supporting analysis, communication, and evaluation. *Visualization and Computer Graphics, IEEE Transactions on*, 14(6):1189–1196, 2008.

[15] W. C. Hill, J. D. Hollan, D. Wroblewski, and T. McCandless. Edit wear and read wear. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 3–9. ACM, 1992.

[16] T. Jankun-Kelly, K.-L. Ma, and M. Gertz. A model for the visualization exploration process. In *IEEE Visualization, 2002.*, pages 323–330, 2002.

[17] H. Lam. A framework of interaction costs in information visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 14(6):1149–1156, 2008.

[18] Z. Liu and J. Heer. The effects of interactive latency on exploratory visual analysis. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):2122–2131, 2014.

[19] J. K. M Wattenberg. Designing for social data analysis. *Visualization and Computer Graphics, IEEE Transactions on*, 2006.

[20] W. Mason and S. Suri. Conducting behavioral research on amazon's mechanical turk. *Behavior Research Methods*, 44(1):1–23, 2012.

[21] G. A. Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[22] J. Nielsen and K. Pernice. *Eyetracking web usability*. New Riders, 2010.

[23] C. North, R. Chang, A. Endert, W. Dou, R. May, B. Pike, and G. Fink. Analytic provenance: process+ interaction+ insight. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 33–36, 2011.

[24] A. Ottley, H. Yang, and R. Chang. Personality as a predictor of user strategy: How locus of control affects search strategies on tree visualizations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3251–3254, 2015.

[25] P. Pirolli and S. Card. Information foraging in information access environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 51–58, 1995.

[26] P. Saraiya, C. North, and K. Duca. An insight-based methodology for evaluating bioinformatics visualizations. *Visualization and Computer Graphics, IEEE Transactions on*, 11(4):443–456, 2005.

[27] B. Shneiderman. Direct manipulation: A step beyond programming languages. In *ACM SIGSOC Bulletin*, volume 17, page 1993, 1981.

[28] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pages 336–343, 1996.

[29] S. B. Shneiderman and C. Plaisant. Designing the user interface 4 th edition. *ed: Pearson Addison Wesley, USA*, 2005.

[30] Y. B. Shrinivasan and J. J. van Wijk. Supporting the analytical reasoning process in information visualization. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1237–1246, 2008.

[31] A. Skopik and C. Gutwin. Improving revisitation in fisheye views with visit wear. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 771–780, 2005.

[32] J. Vallandingham. The rise and decline of ask metafilter. http://projects.flowingdata.com/tut/linked_small_multiples_demo/. Accessed: 2016-03-01.

[33] M. A. W Willett, J Heer. Scented widgets: Improving navigation cues with embedded visualizations. *Visualization and Computer Graphics, IEEE Transactions on*, 2007.

[34] C. Ware. *Information visualization: perception for design*. Elsevier, 2012.

[35] R. L. Wasserstein and N. A. Lazar. The asa's statement on p-values: context, process, and purpose. *The American Statistician*, 2016.

[36] A. Wexelblat and P. Maes. Footprints: history-rich tools for information foraging. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 270–277, 1999.