Effects of Display Size and Resolution on User Behavior and Insight Acquisition in Visual Exploration

Khairi Reda\textsuperscript{1,2}, Andrew E. Johnson\textsuperscript{3}, Michael E. Papka\textsuperscript{1,4}, Jason Leigh\textsuperscript{2}
\textsuperscript{1}Argonne National Laboratory, \textsuperscript{2}University of Hawai’i at Mānoa, \textsuperscript{3}University of Illinois at Chicago, \textsuperscript{4}Northern Illinois University
\{kreda, papka\}@anl.gov, ajohnson@uic.edu, leighj@hawaii.edu

ABSTRACT
Large high-resolution displays are becoming increasingly common in research settings, providing data scientists with visual interfaces for the analysis of large datasets. Numerous studies have demonstrated unique perceptual and cognitive benefits afforded by these displays in visual analytics and information visualization tasks. However, the effects of these displays on knowledge discovery in exploratory visual analysis are still poorly understood. We present the results of a small-scale study to better understand how display size and resolution affect insight. Analyzing participants’ verbal statements, we find preliminary evidence that larger displays with more pixels can significantly increase the number of discoveries reported during visual exploration, while yielding broader, more integrative insights. Furthermore, we find important differences in how participants performed the same visual exploration task using displays of varying sizes. We tie these results to extant work and propose explanations by considering the cognitive and interaction costs associated with visual exploration.

Author Keywords
Large high-resolution displays; visualization; exploratory visual analysis; cognitive biases.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g., HCI): Miscellaneous

INTRODUCTION
Exploratory analysis represents one of the main use cases for visualizations. Often, the primary reason for employing visualization is to enable users to broadly explore their data, acquire new insights, and ask new meaningful questions [29]. Interactive visualization tools foster exploration by affording interactions that enable one to navigate large information spaces and perceive new relationships [6]. However, as datasets grow in size and complexity, visual exploration becomes a more difficult and straining process, requiring users to virtually navigate increasingly large information spaces (by panning and zooming or switching between windows, for instance). Such actions come at significant cognitive costs as they force users to switch between different projections of the data and integrate visual features across temporally-separated views. Conventional displays tend to perpetuate this behavior [3], for they typically provide limited resolution and can only show a small fraction of the data at a time. Faced with this overhead, a user’s natural response would be to reduce virtual navigation and potentially ignore parts of the information space [16]. Such accommodating behavior could negatively impact the quantity and breadth of discoveries.

Large high-resolution displays (LHDs) are increasingly being used as visual interfaces for the exploration of large datasets [21]. Compared to traditional displays, LHDs enable one to view a larger portion of the data space at once, reducing the need for virtual navigation [3]. Numerous studies have demonstrated the cognitive and perceptual benefits of LHDs [1, 2, 3, 4, 8, 13, 31]. However, the majority of these studies comprised relatively simple or abstract tasks under controlled conditions. The larger effects of adopting LHDs in exploratory visual analysis remain unclear.

In this paper, we report on a study examining the benefits of LHDs in multiple-view visualizations. We employ an insight-based methodology within the context of a realistic, open-ended visual exploration scenario. Analyzing participants’ verbal statements, we measure the quantity and breadth of insights they acquire under two experimental conditions (small vs. large display), and ask the following question: What are the effects of increasing the size and resolution of the display on knowledge discovery in exploratory visual analysis?

We review extant research and contribute a theoretical discussion of how interaction costs might affect user behavior in visual exploration, and how LHDs can potentially alleviate some of those costs. Following that, we elaborate on the methodology we used to quantify participants’ insight acquisition and exploratory behavior, and report on our results. These results provide preliminary evidence that larger displays with more pixels can significantly increase the number of discoveries reported by users during visual exploration, while yielding broader, more integrative insights. We propose explanations for these results and discuss the limitations of our study.

BACKGROUND AND RELATED RESEARCH
Visual exploration is a process of broadly surveying data from a flexible point of view to identify salient patterns and for-
mulate preliminary hypotheses and narratives to account for observations. In contrast to confirmatory analysis, which is typically aimed at answering specific questions, visual exploration is intended to lead to new meaningful questions [29] and broaden one’s understanding of their data. Hence, a crucial challenge for interactive visualizations lies in the design of usable interfaces that minimize the cognitive workload on users, and thus invite rather than hinder exploration.

Coordinated multiple views and virtual navigation
A useful paradigm for designing effective visualization interfaces is to distribute information along multiple coordinated views and juxtapose them. The main premise here is that “users understand their data better if they interact with the presented information and view it through different representations” [25]. Multiple-view visualizations have been proven effective in a variety of applications, ultimately becoming a standard technique in visual analytics. Yet, the scalability of this technique is arguably limited by display size and resolution, as adding additional views would require the sacrifice of precious pixels. Consequently, visualizations incorporate a limited number of views (typically 2–6). When the amount of information is too large to fit, views are presented in sequence using virtual navigation techniques, such as panning and zooming, overview+detail, and window switching [7].

Virtual navigation thus involves the temporal separation of views, requiring users to switch between different projections of the data. Conversely, juxtaposition entails the spatial separation of views, enabling more coherent access to information at the cost of complicating the interface. Visualization designers therefore need to balance a trade-off between separating views temporally and adding additional views to the interface. Wang Baldonado et al. suggest that extra views should be added parsimoniously to highlight diversity, correlations, and disparities within the data [30]. However, these design heuristics were primarily driven by the desire to minimize the pixel ‘footprint’ of the visualization at a time when display pixels where far more expensive. Here we examine an alternative perspective that considers the cognitive costs associated with visual exploration.

The costs of visual exploration
One useful model for understanding the cost structure of visual exploration is Lam’s framework of interaction costs [11]. Similar in spirit to the seven stages of actions [15], the framework articulates the cognitive costs incurred at each of the seven stages involved in using interactive visualization tools (see Figure 1). Taken collectively, these costs also imply a form of ‘cognitive resistance’ that users must overcome in order to form and set new exploratory goals for themselves. Since users tend to minimize the cost of acquiring information [18], a bias to the path of least resistance may cause them to (unknowingly) steer their strategy to exploit isolated subsets of data, which could negatively impact the discovery process.

Reducing visual context-switching by incorporating more views and using larger displays with more pixels could dramatically alter the cost structure of visual exploration. Specifically, we would expect a decrease in view-change costs (stage 6) as there would be fewer disruptions from virtual navigation [3], and users would need to rebuild their ‘mental map’ less frequently [20]. Additionally, we would expect a reduction in the costs of integrating information across multiple views (stage 7) as users would be able to physically switch between views with embodied actions [1] and access them in a non-sequential fashion. Experimental work also seems to support this conjecture. For instance, for visual comparison tasks, Plumlee and Ware found pan and zoom interfaces to be slower than interfaces with multiple views due to the increased workload on visual working memory imposed by the former [19]. Adding extra views, however, is likely to impose a different type of overhead, which manifests in rising physical motion costs (stage 4) from having to execute more strenuous physical actions (e.g., more frequent head turns). Furthermore, we would expect increasing visual clutter costs (stage 5) from having to operate on a more cluttered visualization and having to manage one’s attention in a larger spatial environment.

Evidence for benefits of large high-resolution displays
As display technology improves and its cost decline, the use of LHDs for data analysis is becoming more prevalent. There is a large body of HCI research on the benefits of using LHDs in general knowledge work [8]. In visualization, LHDs have been shown to improve user performance in tasks such as target search [4], spatial judgment [28], pattern perception [31], and object classification and manipulation [13]. These benefits are partly conferred by the reduction in the frequency of virtual navigation when users are given the opportunity to physically navigate the data space with eye movements and head turns [3]. However, most of these benefits pertain to low-level outcomes (i.e., performance time and accuracy) in basic tasks and static visualizations. In interactive visualization, results were largely dependent on the interaction method employed, with some standard techniques (e.g., context+focus lenses) resulting in poor performance [10]. In higher-order visual analytic tasks, LHDs provided considerable cognitive benefits to intelligence analysts working with large quantities of text documents [1]. Memory externalization and enabling the creation of semantically-rich spatial representations were two of the main affordances conferred [2].

Figure 1. The sequence of stages users take when interacting with a visualization (in bold) along with the associated costs (in italics) [11]. Users may experience ‘cognitive resistance’ that could dissuade them from forming new exploratory goals in response to high interaction costs.
Summary
Although several studies have considered the cognitive affordances of LHDs in complex sense-making activities, these studies adopted the analysis of text documents as a model task. Interestingly, users did not perform any better or worse in these studies when given a larger display with more pixels [1, 2]. Given the differences in how people process textual and visual representations [12], we would expect different outcomes in analytical tasks involving predominantly visual contents. Researchers have begun to study how interface characteristics (e.g., latency [14]) affect the exploratory behavior of users. However, the effects of display size on exploratory visual analysis remain poorly understood.

METHODOLOGY
Our first goal in this study is to understand the effects of increasing the physical size and resolution of the display on the quantity and breadth of insights acquired by users during visual exploration. Our second goal is to understand potential variations in user behavior as a result of adopting physically larger displays with more pixels. We investigate these effects in a realistic setting by employing an open-ended visual exploration task situated in a real-world scenario and dataset. To that end, we employ two experimental conditions: a small and a large display. While the two conditions provide an identical visualization tool, they afford distinct experiences. The large display allows for the juxtaposition of a larger number of views at the potential risk of overloading the visualization with too much information. The small display provides a simpler visualization with less information, requiring users to frequently switch between views.

Scenario and dataset
To induce a realistic exploratory behavior, we selected a scenario that revolved around the analysis of crime activity in the city of Chicago. The city has made available a dataset containing the majority of crimes that have occurred between 2001 and present. Crime records contain location (down to a block), date and time, and the type of crime. For our experiment, we used a subset comprising crimes committed between 2006 and 2012, inclusive. This resulted in approximately 2.8 million crime records, which we then categorized into 8 distinct crime types.

This dataset furnishes a suitable example to study the analytical behavior of users within the context of a visual exploration task. The scenario provides a semantically rich task for participants to try to find spatio-temporal patterns that govern criminal activity and to explore cause and effect narratives. We also consider the domain to be appropriate for a participant pool drawn from the general population.

Visualization
One of the challenges of conducting user studies on LHDs is that there are very few standard tools that can take advantage of the platform. For this study, we designed a custom visualization tool based on the coordinated multiple views model [25] with an overview+detail navigation technique [7]. We employed two types of visual representations: heatmaps to show aggregate crime density in a particular geographic area, and line charts to illustrate crime trends at three time scales: monthly, daily, and hourly. The two representations were integrated in compound views with the three line charts showing the aggregate number of crimes occurring within the extents of the geographic map (see Figure 2). Each compound view shows crime data according to crime type (e.g., narcotics-related, serious, or all) for a single year. The view can be panned using familiar mouse-drag operations, which also refreshes the line charts to reflect crime trends within the updated geographic extents.

We made a design decision to limit the compound views to a single scale (i.e., zooming was not supported). This eliminates a confounding factor introduced when participants are allowed to ‘read’ the visualization at multiple scales. However, we provided a slider to adjust the level of aggregation in the heatmaps.

View creation and coordination
Additional compound views can be created from the overview map, or by extending an existing view. Views created directly from the overview map are not coordinated, whereas views extended from existing ones are ‘chained’ together and coordinated to respond to interactions. Chained views share the same geographic extents and brushing-and-linking state, but can be set to show different crime types or different years. Additionally, panning the map in one view updates all chained views. The visualization also shows a visual link between the left/right-most view in a chain to the corresponding focal rectangle in the overview map, giving participants a visual cue to identify chained from non-chained views. To help users organize the views on the display, we provided two layout modes: a free form layout enabled users to position and resize views freely using a familiar window-based metaphor, and a structured layout which spreads the chained views around the center of the display and aligns them with the bezels.

The rationale behind the notion of view chaining was to give users some flexibility in specifying which views should be coordinated. In conventional visualization interfaces, all views are typically coordinated to collectively respond to all actions. On LHDs, however, there are advantages to employing

1The dataset is available from https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2
a loose coordination model in which a subset of the views are coordinated. For instance, non-coordinated views can function as independent magic lenses to view disparate parts of a large geographic map. Non-coordinated views can also serve to scaffold multiple exploratory branches, enabling a user to simultaneously pursue several lines of inquiry.

Brushing-and-linking

Rectangular bounding boxes can be created and moved within the heatmap, which are updated accordingly across all chained views. In addition to highlighting subregions of interest, bounding boxes cause additional trend lines to be visualized in the chart area reflecting crime patterns in the selected subregions within the context of the broader trend (see Figure 3). Moreover, line charts can be brushed to display the spatial distribution of crimes within a particular time window.

Apparatus

We used the CAVE2 environment as our LHD apparatus [21]. CAVE2 is a cylindrical system measuring 7.3 meters in diameter and 2.4 meters in height. The environment consists of 72 thin-bezel LCD panels arranged in 18 columns and 4 rows, creating an approximately 320-degree panoramic environment at a total resolution of 74 Megapixels. We choose CAVE2 because the cylindrical configuration of the environment allows users to view the displays at similar distances and visual angles form the center, which reduces geometric distortion in the periphery compared to flat wall displays [5].

Participants

Ten volunteer participants, mostly computer science graduate students, were recruited from our laboratory. The participants, 4 females and 6 males, ranged in age from 25 to 54 years (mean = 31.3, SD = 8.3). Participants reported their familiarity with interactive visualizations on a 5-points Likert scale, with 1 being “not at all familiar” and 5 being “very familiar” (mean = 4.2, SD = 0.7). Participants also reported how often they perform data analysis tasks as part of their work on a 5-points Likert scale, with 1 being “never” and 5 being “a great deal” (mean = 3.4, SD = 0.9). To reduce the chance of inducing an interface novelty effect, we recruited participants who had prior experience using LHDs (primarily for collaborative group meetings). Additionally, all participants resided in the Chicago metropolitan area, and thus were familiar with the city. While it is likely that participants brought their existing perceptions of crime prevalence throughout the city, we consider this to be an acceptable compromise as our goal was not to measure the accuracy of participants’ findings, but rather to understand their strategy and the nature of insights they acquire.

Procedure

Participants were seated in the center of the environment on a swivel, height-adjustable chair at a distance of approximately 4 meters from the displays, and interacted with the visualization using a standard mouse and keyboard setup. Participants were also provided with a notepad and a pen to use during the activity, if desired. We provided the participants with a description of the dataset and a list of questions as a starting point into the analysis. However, participants were told and encouraged to freely explore the dataset and to follow up on questions that interested them. Examples of the questions provided include:

- Identify crime ‘hotspots’ throughout the city. Are there distinct hotspots for specific types of crime?
- How does crime in the city vary with time of day, day of the week, and season?
- Are there correlations between different types of crimes?

The experiment began with a 15 minutes training session with the experimenter demonstrating the visualization environment and its various interactive features. Participants were then given 150 minutes (2.5 hours) to explore the Chicago crime dataset, and instructed to think aloud during the activity. Participants were informed that they could end the experiment when they felt that they have exhaustively explored the dataset. Otherwise, the experiment was terminated when the 150 minutes of allotted exploration time has elapsed. The session was video and audio recorded.

Study design

We used a between-subjects design with a single dependent variable: The display size (small, large). Half of the participants undertook the study using a small display while the other half experienced the visualization on a large display. Because large displays have been known to provide gender-specific benefits in some cases [27], we balanced the gender ratio in the two conditions. We also attempted to balance the average self-reported data analysis experience as much as possible across the two conditions. The mean reported data analysis experience for the small display condition was 3.6 compared to 3.2 for the large display.
Analysis
We focused our analysis on the video and audio data, which contained a record of the participant’s verbal protocol as well as the state of the visualization at the time. Due to the close correspondence between the verbal protocol and the state of the visualization, we decided to encode the verbalizations directly from the video.

Verbal protocol encoding
We developed a coding scheme inspired by insight-based evaluation methods [26, 14]. As our goal was to understand the effects of increasing the display size on insight acquisition and exploratory behavior, we focused on two themes: insights (observations and hypotheses) and exploratory goals. Coding was performed by a single coder using the following process: Data from two representative participants was coded in three passes to develop the final set of codes. In the first pass, the coder tagged discoveries, hypotheses, decision making, inquiries, deductions, and general comments. These were further aggregated in the second coding pass. In the third pass, the coder developed a criteria for measuring the breadth of insights. The final coding scheme comprised the following categories (see Table 1 for examples):

• Observation: A unit of knowledge acquired from looking at and interacting with the visualization.
• Hypothesis: A conjuncture made by the participant, usually as a result of making a series of observations.
• Question: A statement reflecting an inquiry about the data.
• Goal: A statement reflecting the formation of an exploratory objective. Questions and goals are functionally similar in the sense that they both serve to orient the exploration process. However, we found that objectives, when articulated, were more indicative of forthcoming exploratory actions, and thus were coded separately.
• Comment: Represents a generic verbal statement made by the participant, usually about some aspect of the visualization environment or the user interface.

Breadth of insights
In addition to the above codes, we assigned a score to each observation or hypothesis based on its “broadness” (or “narrowness”), on a scale of 1 to 5. In this scheme, higher scores reflect broader insights that join multiple dimensions of the information space, where as lower scores imply relatively isolated findings. The scoring criteria is outlined in Table 2.

Hypotheses
Our hypotheses for this experiment were:

H1—We hypothesize that the cumulative number of observations would be higher with the large display, driven by the ability to see more views simultaneously. We also expect the rate of observations (defined as the number of observations reported during a minute of analysis) to increase with the large display.

H2—We predict a higher number of hypotheses with the large display. Similarly, we also expect an increase in the rate of hypothesis formulation.

H3—Lastly, we expect participants to acquire more integrative insights with higher breadth score on the large display.

These predictions are based on a case study we undertook to evaluate the use of LHDs for the visual analysis of ensemble data [22, 24]. They also represent our extrapolation of results from earlier large display studies, which indicate improved performance time in basic information visualization tasks [3, 31]. We predict that these low-level productivity gains would encourage participants to dedicate more effort to the analytical task and ultimately improve the discovery process. A counter argument is that participants may not be able to incorporate the additional views afforded by the large display into their analysis, which would lead to comparable outcomes in both conditions.

Table 1. Examples of coded statements from the verbal protocol.

<table>
<thead>
<tr>
<th>Code</th>
<th>Example statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation</td>
<td>“In the Loop, the crime pattern is high on Monday, high on Sunday, low through the week... There’s a dip from Tuesday, Wednesday, Thursday, Friday, Saturday, but it’s high on Sunday and Monday.”</td>
</tr>
<tr>
<td>Hypothesis</td>
<td>“Maybe there was a drop in supply [in reference to a noticeable dip in narcotics-related crimes in 2012]. Because... there’s a lot of things that go into the narcotics trade; it’s not just a local phenomenon, it’s like a global thing.”</td>
</tr>
<tr>
<td>Question</td>
<td>“Why has the peak crime time at night changed [to an earlier hour] across all crimes?”</td>
</tr>
<tr>
<td>Goal</td>
<td>“So I’m gonna go and investigate why it is that, in the earlier years, there were more crimes at one point in time happening in the afternoon.”</td>
</tr>
</tbody>
</table>
Table 2. Criteria for scoring insights based on their breadth along with example statements.

<table>
<thead>
<tr>
<th>Score</th>
<th>Criterion</th>
<th>Example statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Isolated features in one view.</td>
<td>“I can see that there are a lot of non-serious crimes, which is kind of expected, for downtown Chicago.”</td>
</tr>
<tr>
<td>2</td>
<td>Comparative insights involving two level-1 features.</td>
<td>“In this time period [2012], [there are] more weapon violations near the lake [compared to inland areas].”</td>
</tr>
<tr>
<td>3</td>
<td>Temporal trends (over several years) or spatial patterns (involving multiple neighborhoods).</td>
<td>“Across all years it’s the same seasonal trends [in narcotics-related crimes] of peaking in the summer and then dropping off in the fall.”</td>
</tr>
<tr>
<td>4</td>
<td>Temporal or spatial trends combining an additional dimension (e.g., crime category).</td>
<td>“The spikes [in burglaries] seem to be consistent [across the years].... It looked like there’s a pattern... If you compare this [burglaries in the northside], with this [burglaries in the southside], same thing.”</td>
</tr>
<tr>
<td>5</td>
<td>Integrated spatio-temporal patterns with observed variations or correlations between multiple crime types.</td>
<td>“A lot of people in this region [near north-side] are buying drugs, but they’re not fighting. There’s no gang activity. There are much fewer deaths resulting from the narcotics trade. Whereas, disproportionately, in the southern regions there are more narcotics related homicides.”</td>
</tr>
</tbody>
</table>

RESULTS

We first look at the length of the exploratory activity, and report on the quantity and the rate of insight acquisition. We then examine the temporal distribution of insights and describe variations in the behavior of participants between the two conditions.

Exploration time

The length of the exploratory activity comprises the time period spent by the participants actively exploring the dataset. Recall that the maximum time for exploration was 150 minutes. However, participants were free to end the experiment when they felt they were finished, with most participants choosing to terminate the activity earlier. Figure 6 illustrates average exploration times under the two display conditions. We found a significant effect for display size on exploration time (two-sample $t(8) = 3.697, p < .01$). Participants with the large display spent an additional 40 minutes on average on the activity ($\text{mean} = 126.4, \text{SD} = 17.0$) compared to their counterparts with the small display ($\text{mean} = 86.4, \text{SD} = 17.2$).

![Figure 6. Average length of the exploration activity under the two display conditions. Error bars are 95% CIs.](image)

Reported observations

We analyzed the number of observations reported by participants during the exploratory activity, and also looked at the distribution of scores assigned under our breadth metric. Figure 7 (left chart) illustrates the average number of reported observations for each score-level in addition to the combined total. There was a significant effect for display size on the number of observations ($t(8) = 3.23, p < .05$), with the large display eliciting approximately 74% more observations on average. We performed chi-squared test with Yates’ correction to examine the relation between the display size and the breadth of observations. The display size had a significant effect on the distribution of scores ($\chi^2(4, N = 1327) = 263.37, p < .001$), with the large display inducing broader insights in general.

![Figure 7. Average number of observations reported by participants (left chart) and observation rates (right). We report the distribution of breadth scores (S1-S5) as well as the combined average/rate (right most column in either chart). Error bars are 95% CIs. Stars denote significant differences between the two display conditions ($p < .05$, Bonferroni corrected for 5 tests).](image)

We performed post-hoc analysis to further investigate the effects of display size on observations for each score level individually. Bonferroni correction was applied to maintain a significance level of $p < .05$ to each of the five comparison tests. There was a significant effect for display size on level-3 observations corresponding to temporal and/or spatial trends, with the large display eliciting approximately three times more observations on average ($t(8) = 8.161, p < .001$). Similarly, there were significantly more level 4 and 5 observations reported with the large display ($t(8) = 5.475, p < .001$ and $t(8) = 6.324, p < .001$ respectively). The numbers of observations scored at 1 and 2 were not significantly affected ($t(8) = .217, p = .83$ and $t(8) = 1.552, p = .15$ respectively). Interestingly, none of the small display participants reported observations that garnered a breadth score of 5.

Observation rates

While the number of observations reported by the large display participants is higher, this result does not factor in the longer exploration time exhibited by this group. Thus, we calculated a rate of observation which normalizes the number of observations for each participant by the time it took him/her to complete the activity. Figure 7 (right chart) illustrates these rates (in units of observation per minute of analysis). After normalization, we find no significant difference in the combined rates ($t(8) = 1.294, p = .23$). However, post-hoc analysis with Bonferroni correction finds the following significant differences: The rates of level 3 and 4 insights were significantly higher with the large display ($t(8) = 5.96, p < .001$ and $t(8) = 5.597, p < .001$, respectively). Level-5 observations were acquired at a rate of 0.3 compared to nil on the small display ($t(8) = 7.466, p < .0001$).
Hypothesis formulation

Figure 8 (left chart) illustrate the average number of hypotheses as well as their breadth distribution. In contrast to reported observations, we do not find a significant effect for display size on the combined number of hypotheses ($t(8) = 0.587, p = .57$). A chi-squared test with Yates’ correction, however, indicates a significant effect on the distribution of breadth scores ($\chi^2(4, N = 145) = 67.38, p < .001$). Again, the large display appeared to skew the distribution towards higher breadth scores with the distribution seemingly centered at level-4. Interestingly, none of the small display participants formulated hypotheses that were scored at 4 or 5. Post-hoc analysis with Bonferroni correction finds no further significant differences.

Hypothesis formulation rates

We calculated the rates of hypothesis formulation by normalizing the number hypotheses by the duration of exploration for each participant. Figure 8 (right) illustrates these rates (in units of hypothesis per minute of analysis). The combined rate was not affected by the display size ($t(8) = .215, p = .83$). Post-hoc analysis finds no further significant effects for the display size.

There was no significant effect for display size on the number of goals ($t(8) = 1.529, p = .16$) and questions ($t(8) = 0.985, p = .35$) articulated by participants.

Temporal distribution of insights

Lastly, due to the observed variation in the time that participants chose to allocate to the activity under the two display conditions, we looked at the temporal distribution of insights. Figure 9 illustrate the aggregate distribution of insights (including observations and hypotheses). Remarkably, the trend curves are quite similar in both conditions, although the peaks are approximately 25% higher and the curve is 33% wider in the large display. The first insight was acquired approximately 3 minutes into the activity with the small display compared to 6 minutes with the large display, which suggests a higher learning curve in the latter condition. Insight acquisition peaks halfway through the activity in both conditions (at 60 minutes in the small display vs. 75 minutes in the large display), after which the number of insights reported appear to decline steadily. Notably, we see an earlier peak in both conditions around the 30-minute mark, followed by a dip, then a second seemingly-higher peak. This suggests a strategy recalibration by participants after an initial exploration attempt. The distribution also indicates that insight acquisition was not temporally concentrated around a few points in time, but rather spread throughout the different phases of the activity.

Participants’ behavior and strategy

To understand potential effects for display size and resolution on user behavior, we looked for patterns in participants’ transitions between ‘mental’ and ‘interaction’ events [23]. That is, moments when the participant is performing mental computation in his/her head, and moments when he/she is interacting with the visualization, hence offloading some of the information processing onto the visualization tool. We calculate transition probabilities between these two classes of events, and compare them across the two display conditions. Since we are interested primarily in insight acquisition, we focused on three primary mental events: Making observations, Hypothesis formulation, and Goal formation. As for interactions, we were interested in two general categories of interactions:

- **Layout-preserving interaction**: Includes actions that do not result in major changes to the state and layout of the visualization. For our visualization, these actions comprised brushing-and-linking and map panning operations.
- **Layout-changing interaction**: Comprises action that result in major changes to the visualization state, potentially requiring the participant to rebuild his/her ‘mental map’ [20]. This category comprised the following interactions: Creating, closing, and positioning views as well as changing the year or the crime type in one or more views.

These two categories of interactions are likely to incur varying costs as the latter induce a temporal separation of views, whereas the former preserve the existing layout.

Event transition diagrams

For each participant, we combine the mental events from the coded verbal protocols with interaction events (which we manually coded from the video recording) into a single sequence sorted by timestamp. This sequence can be thought of as an ‘extended log file’ that captures both mental and computational processes within the combined human-computer cognitive system. We then visualize these transitions in a state transition diagram. To reveal systematic variations in user behavior, we averaged transition frequencies for participants under the same condition, giving us two ‘average diagrams’ corresponding to the small and the large displays. Figure 10 shows the two average diagrams side-by-side (top) as well as a difference diagram and adjacency matrix (bottom).

There are important differences between the two conditions; the large display diagram exhibits decreased transitions to the Modify layout state (column 1 of the adjacency matrix), indicating that participants were less likely to initiate layout-disruptive actions. Participants were also less likely
to perform brushing-and-linking and map panning operations. More importantly, however, we generally see an increased tendency for participants to transition from interaction states to insight-generating mental states with the large display (columns 4 and 5) in the large display. We also see increased likelihood for participants to form goals, suggesting that their activity was often driven by deliberate exploratory goals. Furthermore, we see an increased likelihood for participants to remain in insight-generating states (elements (4,4), goals. Furthermore, we see an increased likelihood for participants to transition from interaction states to insight-generating mental states with the large display. This result can be explained by taking into account changes to the cost structure of visual exploration wrought by the increase in the size and resolution of the display. The large display, by encouraging the spatial separation of views, affords the cognitive integration of features and patterns found in multiple views. We see evidence of this form of integration in our analysis of participants’ behavior. For example, Figure 10 shows a tendency for participants to remain in insight-generating mental states (e.g., making observations, where they are likely to integrate their findings and potentially develop more complex insights in this sense, a larger display is likely to keep participants in a more fruitful ‘cognitive zone’ by reducing interruptions due to virtual navigation [9]. Conversely, the small display participants, having had to frequently switch between temporally-separated views, where less likely to retain relevant information in their working memories, which would have decreased their chance of integrating individual observations.

Figure 10. Top: two transition diagrams comparing average transition probabilities under the small and large displays. Edge weights represent transition probability between two states (with log transformation applied); darker arrows represent more likely transitions. Bottom: a difference diagram illustrating the effect of increasing the display size on transition probabilities. Edges that have been strengthened are color coded with orange whereas weakened edges are color coded with purple. The adjacency matrix shows percentage changes for each edge, with outlined cells indicating significant differences between the two display conditions ($p < .005$, Bonferroni corrected for 25 tests).

**DISCUSSION**

Our results suggest important consequences for increasing the physical size and resolution of the display on the quantity and breadth of insights acquired, and on user engagement.

**Insight acquisition**

Our results show a significant increase in the quantity of observations reported by participants during visual exploration, given a larger display with more pixels. Participants generated 74% more observations on average with the large display, partially confirming our first hypothesis (H1). One possible explanation is that the larger display enabled participants to see more interesting features, by enabling them to perform the bulk of their analysis using mental map-preserving interactions (e.g. brushing-and-linking) and physical embodied actions (eye movements and head turns). These participants were also able to simultaneously see more views, representing crime patterns in different years, in multiple areas of the city, and/or distinct types of crime. Participants who used the smaller display, on the other hand, had to perform a greater amount of virtual navigation to cover the same breadth. The (hypothetically) lower cognitive cost of physical navigation may have induced participants to revisit information more frequently and ultimately acquire more insights, given a display that can accommodate more views.

However, and contrary to our prediction, we do not find significant differences in the rates of observation between the two conditions (H1). One at first may be inclined to conclude that the increase in the number of observations was simply due to participants spending more time on the activity with the large display (recall that participants spent an additional 40 minutes on average on the task with the large display). However, comparing the temporal distribution of insights in Figure 9, one can see that the large display generally induces more insights in the same time ranges compared to the small display. In retrospect, our limited sample set combined with individual variations between participants may have overwhelmed differences in insight acquisition rates, which are nevertheless visible at the aggregate level.

As for our second hypothesis, which predicts an increase in the number of hypotheses formulated by participants when given a larger display (H2), we do not find evidence to support it. Although there is a slight trend suggesting an increase in the cumulative number of hypotheses in the large display (see Figure 8), these differences are not significant. Furthermore, the aggregate rate of hypothesis formulation is quite similar for all participants, regardless of the display size.

**Breadth of insight**

A key difference between the large and small display lies in the distribution of breadth scores. On the large display, most observations garnered a score of 3. Conversely, the majority of observations on the small display were scored at 1 or 2. This is consistent with our prediction that the large display would afford the acquisition of more complex and integrative insights (H3). We see a similar trend in the distribution of hypotheses, though the differences were not significant.

This result can be explained by taking into account changes to the cost structure of visual exploration wrought by the increase in the size and resolution of the display. The large display, by encouraging the spatial separation of views, affords the cognitive integration of features and patterns found in multiple views. We see evidence of this form of integration in our analysis of participants’ behavior. For example, Figure 10 shows a tendency for participants to remain in insight-generating mental states (e.g., making observations, where they are likely to integrate their findings and potentially develop more complex insights. In this sense, a larger display is likely to keep participants in a more fruitful ‘cognitive zone’ by reducing interruptions due to virtual navigation [9]. Conversely, the small display participants, having had to frequently switch between temporally-separated views, where less likely to retain relevant information in their working memories, which would have decreased their chance of integrating individual observations.
From a top-down perspective, participants with the large display may have had more opportunity to form ambitious exploratory goals. Such goals could have been perceived as too costly on the small display. We find some evidence for this in Figure 10, which suggests a tendency for the large display participants to transition more often to Form goal. Anecdotal evidence from participants’ comments also supports this conjecture. For instance, when participant S3, who undertook the experiment with the small display, was asked why he had focused his analysis on crimes in 2012 (the most recent year), the participant responded: “I only bothered to look at the years when I knew something about an area– like Cabrini Green and the Taylor area”. This comment suggests that the cost of exploring temporal trends in this case was inhibitory that the participant generally avoided it unless he knew a priori what to focus on. In other words, in circumstances of weak information scent [18], a small display may skew the exploratory goals of users toward the acquisition of relatively isolated insights which represent a ‘low-hanging fruit’. This phenomenon is somewhat similar to the ‘narrowing in’ behavior observed by Patterson et al. [16], which caused intelligence analysts to inadvertently miss highly-profitable information that were beyond their self-set search scope. A larger display may afford a wider ‘spotlight’, enabling users to incorporate a richer variety of information in their analysis.

**Cognitive engagement**

Our study also raises the question of why participants chose to devote significantly more time to the task, given a larger display with more pixels. One possible explanation is that participants were simply more ‘engaged’ with the analytical task, with the large display serving to sustain their attention for longer durations. The notion of engagement is widely cited in the HCI literature as a positive quality that is often correlated with effective designs. In visual analytics, we can speak of cognitive engagement as the tendency of being invested in the analytical process beyond superficial interest [17].

A spatially larger visualization environment may provide users with more compelling and interesting visual stimuli by affording physical navigation as a data exploration modality. Alternatively, engagement may arises through a self-sustaining process where gaining insights is likely to drive users’ curiosity, prompting them to seek further discoveries. The usability of the visualization interface plays a crucial role here; by eliminating unnecessary cognitive costs, the visualization becomes more usable and feel more inviting to exploration. In this sense, LHD-based visualizations could engender more engagement by reducing the costs associated with visual context switching. Ultimately, increased cognitive engagement could improve the overall discovery process as suggested by our study.

**STUDY LIMITATIONS**

Our between-subjects study is based on a relatively modest number of participants. Therefore, we consider the results to be preliminary. Additional studies are thus needed to confirm and clarify some of the trends we observed.

Although participants in our study were occasional users of LHDs (primarily in collaborative group meeting settings), they did not have prior experience in the analytical scenario posed by the study. This may have impacted their level of motivation, selection of strategy, and their understanding of phenomena depicted in the visualization. A domain expert might be able to override some of the heuristic biases faced by participants, and potentially overcome some of the constraints experienced by participants in the small display condition.

Another limitation stems from the visualization environment we employed in the study. To make the experiment more manageable we opted for a minimalistic visualization, employing only two types of representations (heatmaps and line charts) in a small-multiples layout. Additionally, to reduce the learning curve, we utilized standard input devices (keyboard and mouse). One should therefore be careful when extrapolating our results to visualizations that employ more diverse representations and/or different interaction methods.

Lastly, owing to the lack of guidelines on designing visualizations for LHD environments, we had to make some unusual design decisions. For instance, we introduced a notion of loose coordination in which some views are coordinated and others are not. While we do believe these decisions are justified, we acknowledge the need for a more systematic investigation culminating in the articulation of a design space for LHD-based multiple-view visualizations.

**CONCLUSIONS**

Thanks to advances in display technology, large high-resolution displays (LHDs) are being increasingly seen as cost-effective instruments for the visual analysis of large datasets. Compared to conventional displays, LHDs enable greater quantities of information be visualized at once, potentially alleviating the ‘tunnel vision’ phenomenon experienced by data analysts. We have conducted a small-scale, between-subjects study to investigate the effects of adopting LHDs on exploratory visual analysis. We find that increasing the physical size and resolution of the display can significantly improve the quantity and breadth of insights acquired by users. Moreover, we find that a spatially larger visualization interface can improve user engagement and engender more effective exploratory behavior. These results provide preliminary evidence that the adoption of LHDs as a visualization instrument may positively impact the discovery process in data science applications. The aforementioned benefits may be attributed to a reduction in the cognitive costs of visual exploration wrought by LHDs, which promote the spatial separation of information and lessen the need for visual context switching.

**ACKNOWLEDGMENTS**

This paper has benefited tremendously from the feedback of anonymous CHI reviewers. We are very appreciative of their volunteer efforts. This work was supported in part by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357. Equipment used in this study was funded by National Science Foundation award CNS-0959053 and Department of Energy contract DE-SC005067.
REFERENCES