

Inking Your Insights: Investigating Digital Externalization Behaviors During Data Analysis

Yea-Seul Kim^{1,2} Nathalie Henry Riche² Bongshin Lee² Matthew Brehmer^{2,3}
 Michel Pahud² Ken Hinckley² Jessica Hullman⁴

¹University of Washington ²Microsoft Research ³Independent ⁴Northwestern University
 yeaseul1@uw.edu {nath, bongshin}@microsoft.com mb@mattbrehmer.ca
 {mpahud, kenh}@microsoft.com jhullman@northwestern.com

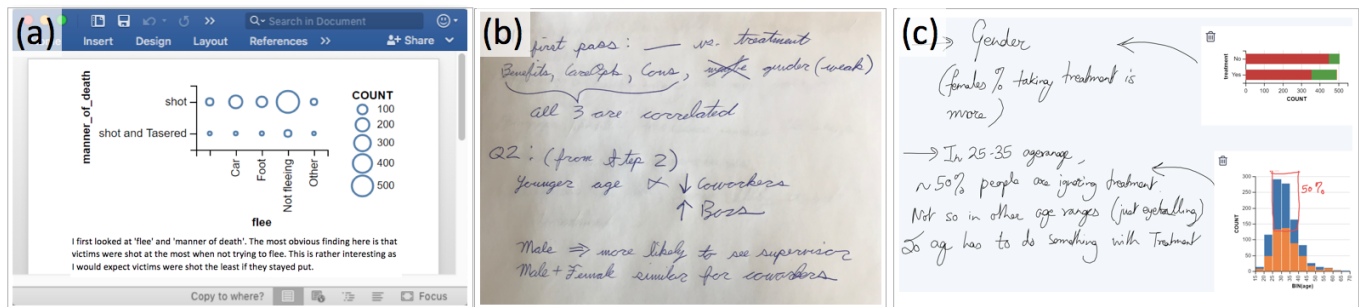


Figure 1. Examples of annotations and notes generated by participants in our two studies. Annotations and notes (a) typed in MS Word, (b) written with a pen on paper, and (c) written with a digital pen.

ABSTRACT

Externalizing one's thoughts can be helpful during data analysis, such as which one marks interesting data, notes hypotheses, and draws diagrams. In this paper, we present two exploratory studies conducted to investigate types and use of externalizations during the analysis process. We first studied how people take notes during different stages of data analysis using VoyagerNote, a visualization recommendation system augmented to support text annotations, and coupled with participants' favorite external note-taking tools (e.g., word processor, pen & paper). Externalizations manifested mostly as notes written on paper or in a word processor, with annotations atop views used almost exclusively in the initial phase of analysis. In the second study, we investigated two specific opportunities: (1) integrating digital pen input to facilitate the use of free-form externalizations and (2) providing a more explicit linking between visualizations and externalizations. We conducted the study with VoyagerInk, a visualization system that enabled free-form externalization with a digital pen as well as touch interactions to link externalizations to data. Participants created more graphical externalizations with VoyagerInk and revisited over half of their externalizations via the linking mechanism. Reflecting on the findings from these two studies, we discuss implications for the design of data analysis tools.

Author Keywords

Visualization; externalization; annotation; note-taking; data analysis

CCS Concepts

•Human-centered computing → Human computer interaction (HCI); Visualization; Empirical studies in visualization;

INTRODUCTION

People often externalize their thoughts as they interact with complex information [22]. For example, an analyst might jot down observations as she makes sense of unfamiliar data, or record interesting patterns on paper with the intent of revisiting them later. By offloading cognition to an external record such as a paper notebook or a digital annotation tool, people can more easily access, connect, and share insights [22]. Recognizing the utility of externalization for thinking about data, several visualization systems support capabilities such as note-taking (e.g., [5]), diagramming (e.g., [37]), and graphical annotation (e.g., [13]).

While such capabilities have become more common, there is a lack of empirical evidence to answer questions to inform their design in systems. For example, are there any differences in externalization behavior with respect to different data analysis stages? How do people refer to aspects of visualizations in an externalization? How does the affordance of different input modalities manifest in a data analysis context? To bridge this gap, we set out to understand when and how people externalize their thoughts and how these externalizations are used during exploratory data analysis. With an aim to inform the design of externalization support in data analysis tools, we conducted

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
 ISS '19, November 10–13, 2019, Daejeon, Republic of Korea.
 Copyright © 2019 Association of Computing Machinery.
 ACM ISBN 978-1-4503-6891-9/19/11 ...\$15.00.
<http://dx.doi.org/10.1145/3343055.3359714>

two studies to investigate the people's behavior when they are given different types of tools for externalization (Fig. 1).

In our first study, we observed how participants externalized their thoughts while using a visualization recommendation tool during exploratory data analysis. The tool allowed them to annotate visualizations using mouse and keyboard, in addition to their favorite note-taking method. Our results suggest that externalization most often takes the form of *notes* (i.e., externalization separate from the visualizations) made on paper or typed in a text editor. Typed *annotations* (i.e., externalizations on top of visualizations) within the data analysis tool were used only in the first stage of analysis while participants were getting familiar with the data. Annotations were often placed near the visual components that their insights referred to (e.g., near the tallest bar in a chart to emphasize the maximum value). As the analysis progressed, participants revisited their externalizations to determine which of their findings should be formulated into a hypothesis, and if their most recent observations were contradictory to previous findings. Since their notes were external to the data analysis tool, participants explicitly referenced data variables and values with the intent of revisiting them later. Yet, we found that over a third of these notes lacked clear identifiers and could not be definitively associated with the data, making this task difficult.

As the annotation behavior was observed only in the initial stage where participants explored the data without restrictions, similar to brainstorming, we speculate that free-form externalizations may support the feeling of being in “flow”, a state of heightened creativity [33]. Other research in HCI and educational psychology indicates further benefits of supporting free-form externalization. For example, it improves one's comprehension and retention of information in learning contexts relative to short-hand writing (e.g., typing) [31], and it may transfer the benefit of active reading to the reading of visualizations [39]. We thus hypothesize that people may benefit from free-form externalizations with a digital pen. Instead of placing typed text near the appropriate part of the visualization, people can freely draw or write their insights where they want. In addition, as participants actively attempted to review externalizations when they needed to formulate hypotheses in the initial study, we also hypothesize that it would be beneficial to provide a more explicit way to link between visualizations and externalizations to facilitate revisitation.

To investigate the potential of these opportunities, we conducted a second study with the same participants using VoyagerInk, a prototype that we developed to support free-form ink annotations, both atop visualizations and in a separate notes pane, where notes could be linked to visualizations via touch interaction. Our results reveal that participants produced more annotations atop visualizations than typed externalizations in the first study. They linked nearly two-thirds of their externalizations to visualizations. Unlike in the first study, participants externalized throughout the entire analysis process as opposed to solely during the initial phase. Their comments suggest that the cost of externalizing might be lower with VoyagerInk.

Our contributions are twofold. First, we report findings from two exploratory studies. Specifically, we report findings of the initial study with a conventional externalization mechanism that allowed us to identify potential opportunities of externalization mechanisms for visualization recommendation tools, and findings from a follow-up study that accommodated the opportunities. Second, we characterize externalization behaviors by different stages of analysis, the affordances of different modalities, and linking capabilities. Our findings provide a novel characterization of externalization during exploratory data analysis that can be used as a basis for approaching the design of the future data analysis systems.

RELATED WORK

This research sits at the intersection of visualization, pen and touch input, and active reading.

Externalization in Visualization

Offloading one's thoughts to an external medium can benefit data analysts in many ways, from making cognitive resources available for other tasks [4, 22] to improving information comprehension and retention [5, 7, 16, 44], or providing shared reasoning artifacts that facilitate communication and collaboration. Externalization is particularly helpful in the process of exploratory data analysis, given its iterative and complex nature. Once analysts have externalized their thoughts by note-taking, annotation, or other means, they can later revisit their insights to review, contrast, and relate them, potentially generating hypotheses and new insights leading to analyses of higher quality [6, 18, 19, 27, 28, 29, 47].

Many visualization tools provide note-taking functionality. Although the format and expressiveness of such functionality vary, existing tools tend to impose constraints on note-taking, particularly with respect to their input modality. Of the tools offering textual (e.g., [44]) and graphical annotation support, some provide a structured palette of annotation shapes (e.g., [10, 37]). Since the expressiveness of the input modality constrains what analysts can externalize, such limitations on annotation may impact the quality of analysis. However, other tools provide the ability to add free-form annotations by drawing with a mouse (e.g., [9, 13]) or with a digital pen (e.g., [25, 46, 34]). One tool [11] even supports the ability to record audio notes. While these annotation capabilities may offer input that is closer in spirit to pen and paper, the contents, intents, and types of annotation made via different input modalities are not understood well enough to guide a principled design of externalization support for visualization systems.

While mouse, keyboard, and even digital pens may degrade the experience of taking notes relative to a familiar, simple, and reliable pen and paper approach [33], digital note-taking can also provide significant advantages. For example, several tools (e.g., [13, 32, 35, 41]) provide a link between visualization and annotations referencing the visualization. This enables people to more easily surface annotations when viewing the visualization, and vice versa. For example, the Aruvi system [37] couples the visualization view with an annotation at the time when the annotation is added, so that analysts can retrieve the view by clicking the annotation.

Pen and Touch Interaction in Visualization

Recent evolution in high fidelity pen and touch technology and the popularization of touch screens on personal computers offer new opportunities for information visualization and data analysis [23]. As pen input supports creativity and expressivity, several recent systems have integrated pen input for creating visualizations and for storytelling (e.g., [24, 45]). Recent research has also indicated the potential use of pen and touch interaction for data exploration. For instance, SketchVis [3] leverages a set of hand-drawn marks and handwriting recognition to interact with charts. Since then, many other systems (e.g., [25, 46, 8, 17]) have leveraged pen and touch to tackle different aspects of visual exploration. Most of these systems support free-form annotation via digital pen input. These systems and their related studies focus on evaluating the interaction design for creating and interacting with visualizations, as opposed to a dedicated focus on externalization. Rather than proposing novel systems and annotation paradigms such as the recent ActiveInk [34], our work seeks to build knowledge on externalization behaviors, especially in the case of recommendation systems, generating a plethora of views for the user to browse and sort through.

Digital Externalizations and Active Reading

There is a substantial body of research on active reading of documents. Researchers have explored the possibility of incorporating free-form inking capabilities to annotate a document while reading (e.g., [1]) or to manipulate documents (e.g., [42]). Others have proposed applications (e.g., [30, 36]) that support digital pen inking that can automate search or content organization while maintaining the advantages of analog pen and paper. Other systems have incorporated free-form ink for externalizing and gathering information from the web (e.g., [15]), or for sharing ideas (e.g., [26]).

Riche et al. [33] conducted diary studies of participants using analog pens and their digital counterparts, attempting to understand activities and affordances that failed to transfer from analog to digital in popular commercial software. One salient finding is that participants often chose an analog pen over a digital one when externalizing their thoughts, such as by capturing information on scraps of paper or by annotating a printed document directly.

Several recent studies have explored the role of free-form sketching in visualization. Walny et al. [40] provided insights on inking marks and visual thinking on whiteboards, which motivated an investigation of the types of visualization that people sketched [38]. To the best of our knowledge, Walny et al. [39] provide the first study on free-form pen inking when reading visualizations presented in either an analog or a digital medium. Their results suggest that active reading behaviors transfer from documents to visualizations. Our work expands this knowledge by studying digital externalizations on visualizations during a more complex and realistic exploratory data analysis process.

STUDY 1: UNDERSTANDING EXTERNALIZATION

Currently, little formal knowledge addresses externalization behaviors during visualization-aided exploratory data analysis.

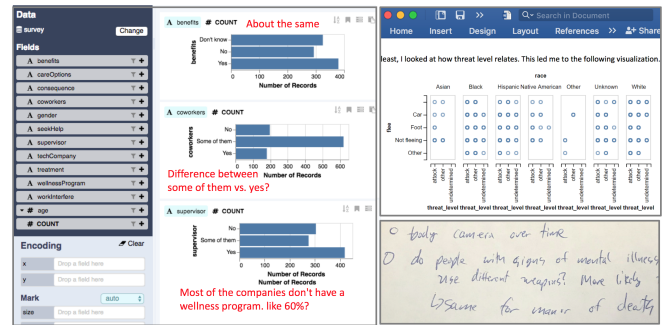


Figure 2. The interface of VoyagerNotes. Participants can annotate in VoyagerNotes or create externalizations using MS Word or on paper.

sis. Understanding the role of externalization is especially important for novel genres of visualization-based tools such as Voyager: recommending a series of visualization for the analyst to browse and derive insights from. The main goal of this study is to understand the types of externalizations and their occurrence in different phases of the exploration.

We opted to conduct our study in laboratory settings, to retain some control over the dataset explored. We divided exploration into three stages, in order to study externalization behavior in each of them.

Setup: VoyagerNotes + External Note-taking

Given our goal of observing participants' behaviors during free exploration of data, we adapted Voyager 2, a visualization recommender tool [43], which allows an analyst to select data fields of interest and receive suggestions for relevant views. Adapting Voyager 2 meant the tool minimized the burden of creating visualizations (e.g., considering which of multiple encodings to choose and specifying one visualization at a time). This is likely to have freed participants' attention for focusing on observations about the data, effectively speeding up what might otherwise be a more tedious process.

As Voyager 2 supports only the bookmarking of visualizations, we added basic annotation support in our study prototype, VoyagerNotes (Fig. 2), to allow an analyst to type a text annotation via keyboard at the desired location. To enable revisiting these annotated visualizations, we implemented an "Annotation Review" page that consolidates all typed annotations that the analyst created atop their corresponding visualizations.

We launched VoyagerNotes on a Microsoft Surface Studio, a desktop computer with a 28" display. We also asked participants their choice of digital note-taking application, such as Word or Notepad, and provided them with pen and paper.

Task and Procedure

First, we provided a brief introduction of the study and a tutorial for how to use VoyagerNotes. We asked participants to perform exploratory data analysis, breaking the process into three distinct phases and limiting the duration of each phase. This separation facilitated the analysis of the results according to analysis phase, as we hypothesized that externalizing behavior may differ across them.

- **Phase 1 - Familiarization:** Browsing the dataset and checking data quality.
- **Phase 2 - Formulating questions:** Generating interesting questions about the data.
- **Phase 3 - Answering questions:** Answering specific questions, which we provided for consistency.

To help participants concentrate on data analysis without worrying the added burden of time management across three phases, we instructed participants to spend 15 minutes in each phase. To allow for the iterative nature of exploratory analysis, however, we advised them that they should not let this suggested timing prevent them from returning to a previous phase or making observations relating to an upcoming phase. After each phase, we asked participants to briefly explain what they did and what they found.

To reduce the chances that participants would generate overly simple or out-of-scope questions in phase 2 (Formulating questions), we seeded their question list with two high-level questions. These questions required broad exploratory analysis while having definitive answers in the data, such as “What factors appear to have the most influence on whether a victim attempted to flee?” (for the police shooting data). We asked participants to think of annotations and notes as artifacts to rely on as they conducted their analysis and when explaining their findings to us, but not as polished artifacts intended to be presented to others. We took a screen capture recording of each session and collected any artifacts generated by participants, including digital and analog documents.

Once the analysis was over, we asked participants to complete a post-analysis questionnaire. The questions asked participants when, how, and why they annotated and revisited earlier phases in their analysis, and how these activities helped them at each phase. We also asked participants to describe their own ideal annotation system for visualization-based data analysis.

Results

We analyzed the results of Study 1 with respect to participants' preferred note-taking medium, the frequency and format of their externalizations, as well as the phase of data analysis. We also differentiated between *externalization formats*, including *annotations* (text added atop visualizations) and *notes* (text added in either a text editor or on paper).

Externalization Frequency & Format

Participants chose to take notes that were separate from the visualizations, using the medium that they felt was most natural. Six (P4-7, P9, P10) out of ten participants chose to use pen and paper to take notes. Three participants (P1, P2, P8) used only a Word document to take notes. One participant (P3) chose both pen and paper and a Word document.

To measure the frequency of annotations and notes created by each participant, we defined a *unit of externalization*. When a participant generated text, we counted each sentence (either complete or incomplete) as one externalization. When a participant generated non-text shapes, we counted each distinct shape (e.g., circle, line, arrow) as one externalization, since we observed that many participants tended to use single symbols to indicate independent units of thought. When more than

one shape was used, they were in different components of the visualization (e.g., one in the legend, one associated with a bar mark) or different regions in the visualization.

The participants created a total of 152 externalizations across their analysis session ($M = 15.2$ per participant, $\sigma = 3.9$). The majority of these externalizations (115 out of 152; 75.7%) were notes typed in a word processor or written on paper, while the remaining 37 externalizations were typed annotations added atop visualizations. Among the six participants who chose analog pen and paper as their primary note-taking method, four participants did not make any annotations.

Externalization by Data Analysis Phase

To understand how externalization manifests differently given the analysis phase, we compared the externalizations made during each of the three phases.

Phase 1 - Familiarization: We found that most of the annotations were made in the first phase (Phase 1: 42.1%, Phase 2: 22.4%, Phase 3: 35.5%). Participants often summarized the specific qualities of a data field with text. All of the annotations in this phase pertained to either the distribution of a data field (e.g., *roughly 50% of those shot are between the ages of 20 and 40*) or a correlation between two data fields (e.g., correlation between the data field “seek_help” and “communicate_with_supervisor”: *if they feel comfortable talking to their supervisor, they mostly don't seek help, and many gunshot victims did not have sign of mental illness*).

Among the 34 externalizations referencing a visualization, about a quarter of them were related to a univariate plot and about a half were related to a bivariate one. Given that none of the annotations created in phase 2 or 3 referenced a univariate plot, this may suggest that participants annotated more complex relationships within the dataset as the analysis proceeded.

Among all externalizations generated in this phase, most (53.1%) were annotations (cf. 0% in phase 2 and 5.6% in phase 3). Two participants (P4, P5) mentioned that creating annotations was advantageous because it kept them in a flow state, whereas note-taking in a separate medium could be distracting. P5 pointed out “[annotation] stays as part of the view so that I don't have to switch back and forth.” P4 mentioned that “[generating notes in separate media] takes viewers attention away from the main viz tool.”

We observed participants jotting down the name of data fields in the dataset, such as in Fig. 3. All occurrences were written with a pen (P6, P7). Some annotations contained the definition of the data field in the participant's own words.

Sometimes participants summarized the possible levels of nominal fields along with the name of the data field (e.g., Armed: weapon or unarmed). We found this use of annotation somewhat surprising as this information was present in each visualization (in axes or legends), so these annotations did not add new information. One possible explanation is that the act of note-taking was to “internalize” new information. P3 mentioned that “[In phase 1,] Most of my annotations were just mental notes to myself of things to look into.”

Phase 2 - Formulating Questions: In this phase, all externalizations were notes in a separate medium; most of these were questions that participants had formulated. Instead of creating more annotations about visualizations, participants revisited their annotations that they had created in phase 1. Among the five participants who indicated that they had revisited their prior annotations, four who had used annotation actively revisited their annotations to formulate questions related to a trend they had discovered earlier.

Revisiting prior annotations allowed some participants to gather salient findings that should be formulated into a question: *“The [revisitations] were useful when translating my general thoughts about visualizations into specific questions that I thought would be interesting to explore.”* (P1). Some revisitations were to check whether the question they had formulated was aligned with a new discovery: *“I revisited the notes to check if I am not contradicting any of my previous notes while making new questions.”* (P4).

While formulating questions by jotting them down or typing them, one participant (P6) felt that the act of note-taking itself helped with brainstorming, especially when using pen and paper: *[To generate a question, I annotated by] “paper. Generating ideas is easier to be written by hand.”*

Phase 3 - Answering Questions: Most externalizations (94.4%) generated in this phase were notes in a separate medium. We observed a few instances (8 notes) where participants referenced the visualizations in their notes. Participants who used the Word document to take notes (P2, P3, P8) copied and pasted the visualization to support their findings. Participants generated longer notes ($M = 13.6$ words per sentence) in phase 3 than in phase 1 (7.2). P8 mentioned that she preferred a medium that provided more space than in-place annotations, so she used MS Word in this phase: *“I ended up using Word. I wanted to type out longer explanations for what I was doing, so I preferred a larger tool.”*

Participants sought confirmation in the data for their insights more often in this phase. We observed many instances where participants revisited the visualizations alongside their corresponding annotations. Participants connected a particular region or trend in the visualization to the annotation.

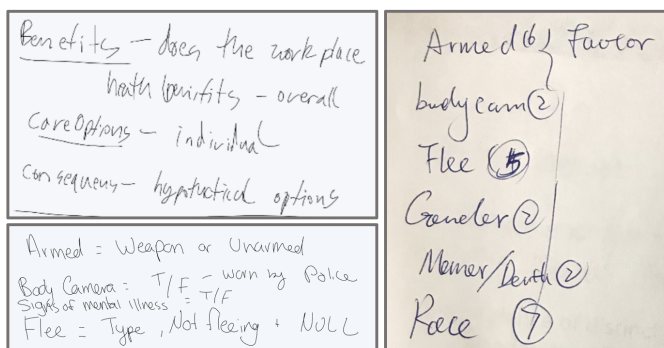


Figure 3. Examples of annotations contained the description of data fields, the level of the data field, or the number of levels in the data field.

Revisitation Behavior

Participants adopted strategies for linking their notes to relevant data, whether on paper or on a separate digital word processor. Most commonly they would refer explicitly to data fields in them. To estimate how *identifiable* notes were a posteriori (e.g., the degree to which the supporting visualizations that the participant viewed while writing the note could be identified), we further analyze the notes that we collected. We created a list of keywords consisting of (1) the name of each data field in the dataset (e.g., gender, armed) and (2) the levels of each data field (e.g., female, male, gun, knife). We then reviewed each note, counting when the externalization contained such keywords. Among the 152 annotations and notes that were created, 49 instances (35.6%) did not contain any keyword related to the dataset. This suggests that the participants may not have been able to locate the source of the externalization later in an analysis process.

Opportunities to Facilitate Externalization

Observations from our study align with several directions explored in the literature and suggest two opportunities—inking and linking—to facilitate externalization during the exploratory data analysis.

Inking

The literature suggests that externalizations using an analog pen has multiple unique benefits over conventional mouse and keyboard [33]. Observations and participants' feedback collected during our study suggests that the following three affordances of the pen are especially relevant in the data exploration process:

1. integrating information in context: participants commented on the benefit of annotating directly on top of reference material to keep annotations and visualizations together, especially in phase 1 when they had to browse and sort through numerous visualizations generated by the system;
2. interweaving symbolic and figurative content: the large majority of annotations summarized either a distribution or a correlation, both potentially easily depicted as graphical shapes, (e.g., a normal curve, a correlation line) overlaid directly on a chart;
3. immediacy of capture: several participants commented on the need to capture their fleeting thoughts during the exploration, mentioning the desire to stay “in the flow.” The analog pen being helpful for brainstorming in phase 2 also reflects the benefit of immediate inking for externalization.

Linking

We were surprised to observe that only two of the six participants who chose pen and paper ended up using this medium to write notes. This may be explained by the disconnect between analog notes and digital visualizations mentioned by three participants (P5, P6, P9). For example, P5 commented on the cost of maintaining such link: *“[with a pen, it is] hard to keep connections between specific notes and the graphs/data you refer to.”* This suggests supporting digital ink to provide support for linking ink to data.

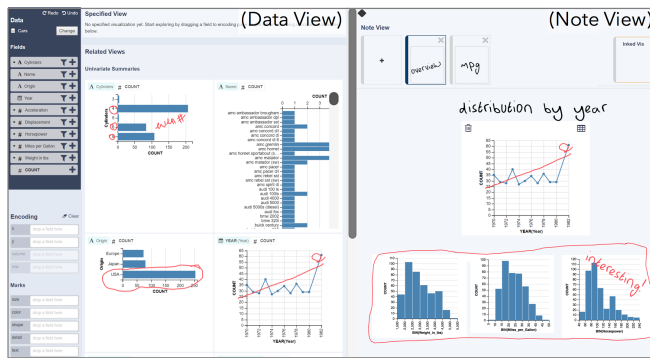


Figure 4. The interface of VoyagerInk. Participants can annotate on top of visualizations in (a) Data View or (b) Note View.

Many participants, whether using pen and paper, or the separate note-taking software, indicated that they desired linking notes to relevant visualizations. These observations align with prior research suggesting that a visualization that prompts an insight during exploratory data analysis should be considered as part of that insight [14]. Analysts also obtain deeper insights from aggregating findings from multiple visualizations [12]. Thus, the ability to link notes with visualizations supports analysts in revisiting insights and their corresponding visualizations in a later phase of analysis, helping them recall the context of the insight and develop further insights [37].

STUDY 2: ROLE OF INKING AND LINKING

In a second study, we investigated how participants would use inking and linking support during exploratory data analysis.

Setup: VoyagerInk

We built VoyagerInk (Fig. 4), a prototype extending Voyager 2 to provide a separate note panel for free-form digital pen-based note-taking as well as touch-based linking between notes and visualizations.

Inking

VoyagerInk allows for the free-form annotation of recommended visualizations (*Data View* in Fig. 4) using a digital pen. *Note View* (in Fig. 4) provides the analyst with space to add handwritten notes in a separated side panel. VoyagerInk uses the metaphor of multiple pages in a notebook, depicting them in a thumbnail ribbon at the top of the note view.

Linking

To provide linking between notes and visualizations, VoyagerInk uses three strategies (Fig. 5).

1- Linking by drag & drop: An analyst can simply drag and drop a visualization from the Data View to a note canvas in the Note View (Fig. 5c). Tapping the visualization in the note canvas brings up the corresponding view and shelf parameters in the Data View, allowing for further interaction.

2- Explicit linking while inking: An analyst can link a visualization with a note in the Note View by touching the visualization with their non-dominant hand while composing the note with the pen in their dominant hand *simultaneously* (Fig. 5d1).

This avoids multi-step sequential interactions, which may interrupt the flow of the analyst's exploratory analysis. After creating the note, the analyst can release their non-dominant hand from the visualization. At this point, the prototype creates a link between the visualization and the note.

3- Explicit linking after inking: To link an existing note in the Note View with a visualization, the participant can tap and hold the note with their dominant hand and touch a target visualization with their non-dominant hand *retrospectively* (Fig. 5d2). This allows for a single note to be associated with multiple visualizations, and vice versa.

For the last two strategies, VoyagerInk adds an icon adjacent to both the visualization and the note once a link is created. The icon indicates the number of other entities (visualizations or notes) that it links to. Touching the pencil icon next to a visualization will display thumbnail links to linked notes (Fig. 5b), and touching a chart icon next to an annotation will display thumbnail links to linked visualizations (Fig. 5e).

Study Procedure

To limit inter-subject variability, we recruited the same 10 participants who participated in the first study. The study procedure for Study 2 was identical to that of Study 1, with the following three differences.

First, we asked participants to use digital pen and touch input with the Surface Studio inclined like a drafting table. Second, we briefly trained participants on how to use VoyagerInk to create annotations, notes, and links, as well as how to revisit the linked annotations and notes. Lastly, after the post-session questionnaire, we asked participants to fill out an additional questionnaire asking them to describe externalization experience with both prototypes: when and how typing or writing on top of a visualization is useful, and to list the advantages and disadvantages of writing on paper with an analog pen, typing in a text editor, and annotating with the digital pen during data analysis. Participants also completed a demographic questionnaire which included questions about their data analysis experience and how often they use visualization tools.

Results

Our analysis approach in Study 2 is the same as Study 1.

Externalization Frequency & Format

Participants created a total of 202 externalizations in Study 2 ($M = 20.2$ per participant, $\sigma = 6.3$). Compared to the 152 total externalizations in Study 1, participants created roughly 33% more externalizations (mostly graphical annotations) using VoyagerInk (Fig. 6). While participants created a similar number of notes with both prototypes (VoyagerNotes: 115 vs. VoyagerInk: 110), they generated more annotations with VoyagerInk (VoyagerNotes: 37 vs. VoyagerInk: 92).

Participants annotated atop visualizations when the externalization was specific to a single visualization. If the insight was generated from multiple visualizations, they were more likely to use a dedicated note canvas. P1 explained this behavior: "I tended to use annotation on charts to jot down general trend lines, compare values between specific charts, and note

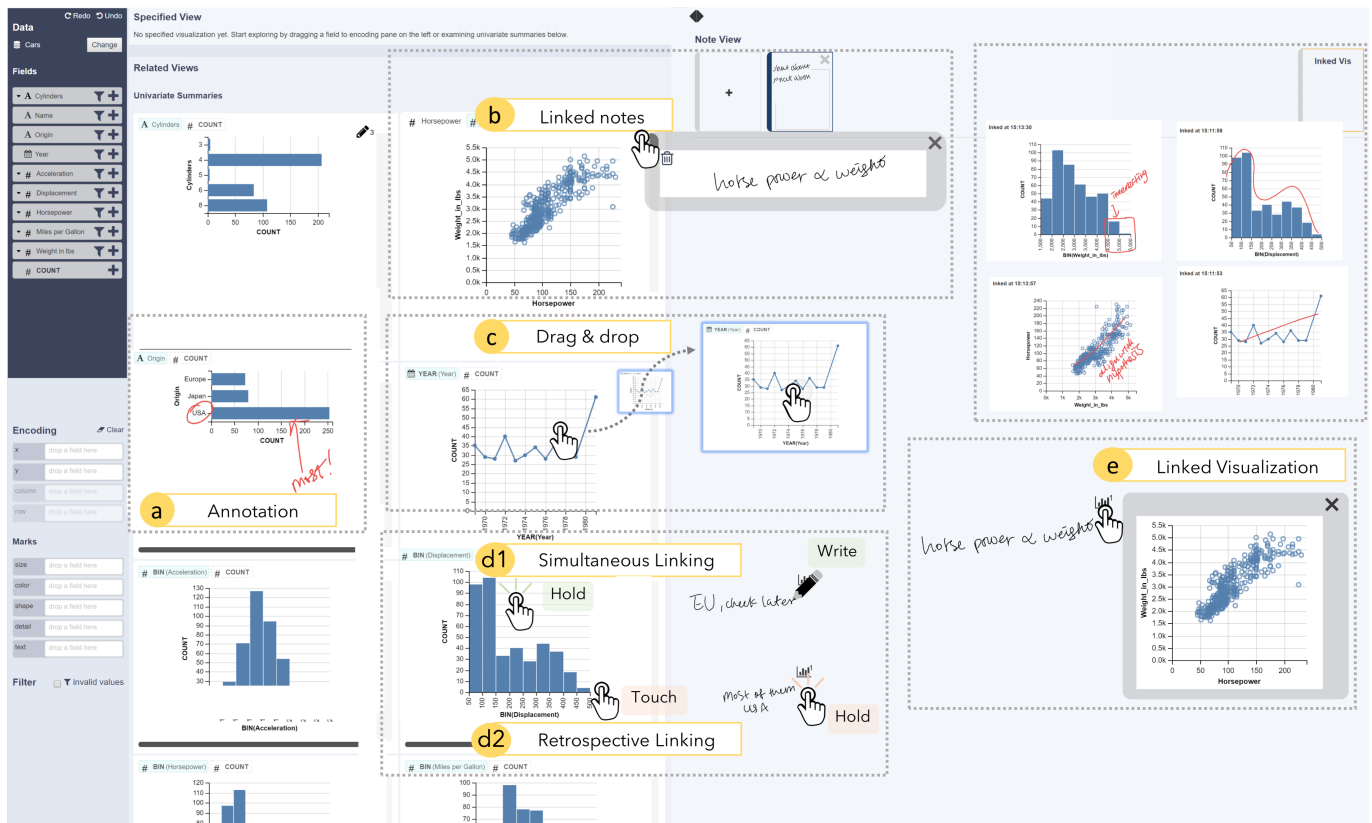


Figure 5. Key interactions with VoyagerInk: (a) draw or write with a digital pen to add a freeform annotation; (b) tap the pencil icon to display a corresponding note thumbnail; (c) drag a visualization from the Data View to a note canvas; (d1) write a note while holding a visualization; (d2) tap a visualization while holding on a note; and (e) tap the chart icon to display a corresponding visualization thumbnail. Actual interactions can be found in the supplementary video.

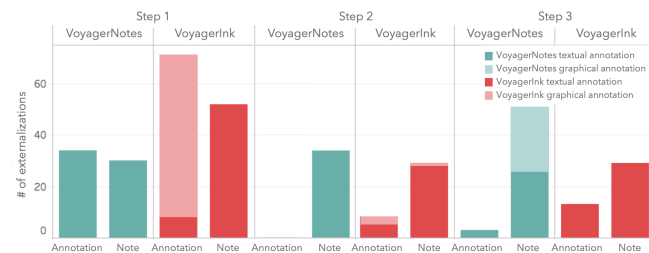


Figure 6. The total number of externalizations created by participants with two systems for annotation versus notes and data analysis phases. The lighter color indicates graphical annotation, and the darker color indicates text annotation.

any odd outliers. I preferred to use the notepad to aggregate multiple charts and make questions or notes based on them.”

Graphical Annotations: The majority (82.6%) of the annotations generated using VoyagerInk were graphical (e.g., Fig. 7) while only 17.4% were textual. We note that graphical annotation support is absent in many conventional visualization systems that support exploratory data analysis, including Tableau.

Participants’ self-reported intentions helped us characterize graphical annotation behaviors with VoyagerInk. The most

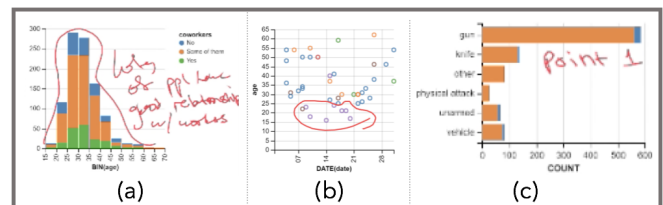


Figure 7. Examples of annotations created with VoyagerInk: (a) textual + graphical, (b) graphical, and (c) textual.

commonly specified usage of graphical annotation was to identify a trend (41.3%), while 26.7% served to identify an anomaly and 14.7% served to connect observations; Fig. 7 shows examples of each intention. In four instances, participants specified that their intentions were not captured by our options, so we recorded it as “other.” The cases contain a circle around an axis title to prompt further investigation of a data field, check marks to indicate the visualization was seen, and unconscious scribbling (Fig. 8, left).

Externalization through Linking: Many participants made use of linking to make notes—all notes made in VoyagerInk were linked to at least one view. Participants dragged an average of 5.8 visualizations from the Data View to a note

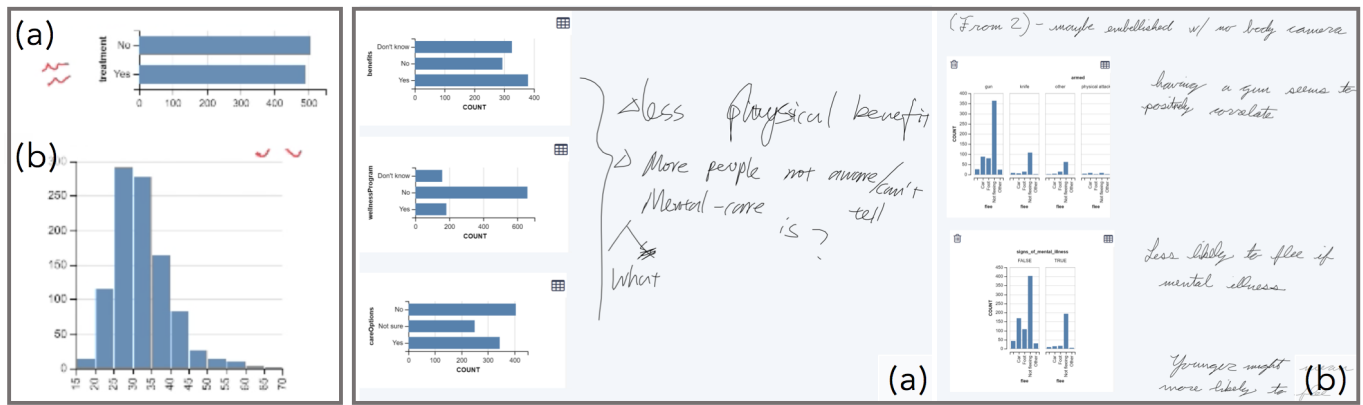


Figure 8. Left: Examples of graphical annotation where the specified intention was “Other.” (a) P1 made scribbles and (b) drew two check marks to indicate the visualization was examined by him. Right: Example usage of participants’ notes next to dragged visualizations. (a) P6 combined multiple visualizations with a curly line, (b) P5 presented the step of the analysis by presenting each visualization that led him to investigate the next visualization.

canvas in the Note View. We observed six instances where participants dragged multiple visualizations from the Data View and referenced them to support one note, such as in Fig. 8 (Right-a). One participant (P5) used it as if they were creating a storyboard, connecting multiple visualizations following his process, such as in Fig. 8 (Right-b).

Seven (P1–4, P7–9) of the ten participants created 4.8 links on average ($\sigma = 3.3$), all via the simultaneous linking interaction. Among all the notes, 26.4% of externalizations were linked to a visualization via a linking interaction. Combined with the implicit links afforded by annotations, 59.9% of externalizations created with VoyagerInk supported revisitation.

Externalization by Data Analysis Phase

We observed the same high-level trend seen in Study 1 with respect to externalization across the phases of analysis, albeit with differences in externalization format.

Similar to externalization behavior in phase 1 of Study 1, we observed annotation as a means of noting trends or data properties, possibly to help internalize the observations. In Study 2, many of the annotations generated in phase 1 were graphical marks on visualizations used to indicate salient trends, rather than text comments. More externalizations including many graphical annotations, were produced in phase 1 with VoyagerInk relative to phase 1 with VoyagerNotes. We observed that the distribution of phase 1 externalizations was more uneven for VoyagerInk, with 36.5% (19) more annotations generated than notes. When participants used VoyagerNotes, the difference between annotations and notes in phase 1 was only 13.3% (4).

Participants again displayed similar externalization behavior in phase 2 with VoyagerInk relative to phase 2 with VoyagerNotes. We saw continued use of annotation in phase 2 with VoyagerInk; in Study 1, we saw 8 instances in phase 1 and 0 in phase 2 (Fig. 6). These results suggest that graphical and textual annotations may support analysis beyond the phase of data diagnostics and familiarization.

In phase 3, participants frequently supported their annotations with references to visualizations; 42.7% of instances were

presented with one or more visualizations (cf. 21.9% in phase 1 and 26.8% in phase 2). All participants dragged the visualizations to a note canvas to support their annotations. P8 mentioned, “I wrote fuller explanations on the notes side. I liked dragging and dropping the graph so that I could see everything at once.”

Reflecting on the Use of Inking and Linking

In this section, we reflect on the use of inking and linking based on our observations and participants comments from both studies.

Inking

Immediacy of capture: One salient insight of study 2 is that participants using VoyagerInk created annotations throughout the session, as compared to using VoyagerNotes where participants added annotations predominantly in phase 1 (Fig. 6). Participants’ comments confirmed that they felt pen input more adequate for jotting down “quick thoughts” and that keyboard input was more convenient for entering “a long chain of thoughts.” This finding aligns with the immediacy of capture: “[inking is useful] when you are trying to make a quick note to yourself, or point out some kind of flaw” (P8); “Drawing with a pen was significantly easier when trying to jot down trends, look at outliers, and mark quick observations about the dataset.” (P1).

Integrating information in context: In line with previous findings [14, 13], we also found that participants appreciated the ability to reference a part of a visualization: “The ability to pinpoint exactly the issues through the use of a pen is nice” (P2). When we asked P5 why free-from in-place graphical annotation was useful, he pointed out that “Even just circling parts of a chart or axis that stand out or drawing a line to connect two parts can be really powerful.” Similarly, P1 indicated that graphical annotations helped him compare the data across visualizations since it called his attention to particular parts of a visualization without requiring that he re-analyzed the information.

Interweaving symbolic and figurative content: The digital pen was perceived to be informal and flexible: “I felt like I

had fewer constrictions. It felt easier to connect written notes to visualizations than the typed notes.” (P8). Our observations confirmed that participants used more dialectic symbols such as “=>” to indicate a correlation between two data fields or connect two observations when using the digital pen.

Further insights: Our observation also suggests that inking enables an informal mode of thinking. Six participants (P2, P3, P5–7, P9) described that typing text felt more structured relative to writing with a pen, with P5 stating that “[typing] feels very rigid, [I] feel pressure to make things nicely formatted.” This comment confirmed our observations of participants using complete sentences in textual annotations, such as in Fig. 9a whereas ink annotations, especially when atop visualizations were more informal, such as in Fig. 9b. Many externalizations created with the pen also contained abbreviations, such as “cat.” for “category”.

To get a better sense of the formality of externalization, we coded the completeness of each textual annotation: an annotation was coded as complete if it contained both a subject and a predicate. We found that 45.8% more incomplete sentences were generated with the digital pen than with the keyboard. This informality is aligned with participants’ testimonials that they considered the annotations generated with the digital pen as a record for themselves. P8 pointed out that “[externalization with the digital pen] feels more informal but also more comfortable. This would be for general note taking for self-reference.”

Participants such as P2 and P7 preferred to have a structured input like typing, with P7 stating “[typing is] great when creating a note that needs to be organized.” P3 also said that “By this I mean that I can type much faster than I can write with a pen.” Fast typing allowed participants to easily externalize their thoughts, especially longer sentences and questions, with P1 saying “faster typing allowed me to easily jot down interesting questions from trends.”

Linking

All seven participants who used linking interactions (P1–4, P7–9) used the simultaneous linking interaction. They all appreciated the capability to create a link between a visualization and a note, as it provided an easy way to revisit the note-visualization pair. P2 mentioned that “the ability to link [notes] to the visualizations was very useful as I could revisit the notes and click on the link directly to view the notes in the context of the visualization I made.”

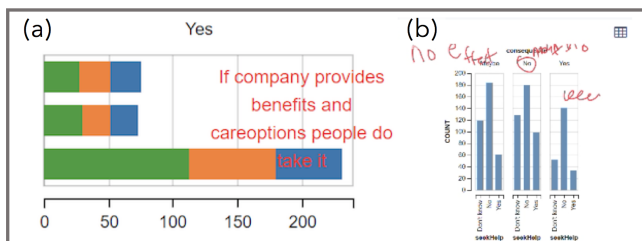


Figure 9. The annotations (a) typed by P9 and (b) drawn by P2.

Despite the ability to revisit annotations, some participants (P3, P5) perceived annotations as being ephemeral and disposable. For example, P3 stated that “Writing notes atop visualizations felt like I might not be able to find them later. I felt like moving visualizations to the note canvas was a more permanent way.” This comment suggests that the different mechanism for linking notes to visualizations may serve different purposes.

DISCUSSION

Toward a Framework of Externalization in Data Analysis

We observed participants externalizing their thoughts during data analysis in ways that suggested that they were trying to internalize unfamiliar information. As their analysis proceeded, externalizations became an external record that participants could revisit to check whether their hypotheses were supported by previous findings. These observations align with two theories in the education literature. First, note-taking is helpful because the act of externalization facilitates a deeper understanding of the content [20]. Second, by providing an external storage, notes benefit people by giving them the ability to review their thoughts [20]. We also observed that the perceived usefulness and the composition of sentences in notes did not vary according to whether the pen was analog or digital. This evidence matches the findings of Kim et al. [21], in which the quantity and quality of notes were similar when participants take notes with either type of pen. Some affordances of analog and digital pens for logging daily activities, such as externalizing thoughts and interweaving symbolic and figurative content [33], are in line with the affordances we observed in a data analysis context.

Participants generated far more annotations with VoyagerInk than with VoyagerNotes. Although they created annotations regardless of the phase of data analysis with VoyagerInk, participants generated no annotations in phase 2 and very few in phase 3 when using VoyagerNotes. We partially attribute this finding to two costs associated with a participant’s intent to add an annotation. The first is the cognitive cost that comes from the perceived pressure of generating a formal and complete annotation. The second is the physical cost incurred when participants had to switch their hand or move their body to generate an annotation, such as switching to the mouse to click to add an annotation and then returning to the keyboard to type. We suspect that both costs were high for generating annotations in VoyagerNotes, resulting in fewer annotations. The digital pen input might reduce these costs by offering a sense of informality, by including the ability to annotate graphically, and by providing fast switching between pen and touch input.

Design Implications

By synthesizing our study observations and the responses from the questionnaires, we now discuss how we could improve externalization support in future data analysis tools.

Supporting Graphical Annotation

Few existing exploratory data analysis tools support graphical annotation. Given our observation of how participants created graphical annotations, particularly during the initial phase of data analysis, graphical annotation capabilities can

provide great value to analysts. One takeaway from our study is that people may worry that graphical annotations made atop of a visualization may disappear as they interact with other visualizations. A visual cue can reassure the analyst that the annotation will be stored in the canvas, such as by animating the trace of the annotation to the note-canvas. Participants created many forms of graphical annotation, including vague “wiggly lines” and those that may not be especially meaningful in a long-term analysis process. Thus, prompts to remove or categorize such annotations will help analysts to organize their insights.

Supporting Linking of Visualizations and Notes

During the analysis process, participants revisited relevant visualizations to confirm insights that arose previously. Some participants accomplished this task easily by using visualizations that they copied and pasted into a separate medium in Study 1 or by using linked visualizations in Study 2. However, we did not observe the revisitations as often as we expected even in Study 2, compared to the number of links they created. This observation may be due to the limited session time in the lab studies. Another possibility is that participants may not know when the link will be important, and explicitly referring to attribute names all the time is costly for them.

Among the participants who attempted to re-visit the visualization, the ones who did not have links to revisit spent a considerable amount of time to find the source visualization by specifying the relevant configuration in the visualization tool. To make revisitation process easier and help analysts confirm their insights without interrupting their analytical flow, providing simple interactions along with appropriate feedback for linking will be important. For example, the interaction that can record and show how strong/important the link may help users later to determine which links they should investigate further as the analysis proceeds.

Supporting Data Familiarization Stage

One interesting finding from our studies is that participants' showed distinctive externalization behavior during the initial stage of exploratory data analysis from the two later phases, which suggests that generating and making use of annotations in a data analysis context can be further characterized by the phase of data analysis. We observed that in the earliest phase of analysis, externalization specifically helped participants make sense of unfamiliar data. By creating many annotations directly atop the visualizations, participants could quickly explore visualizations without switching back and forth between the visualization and an external medium where they would externalize their thoughts. Features that facilitate this familiarizing phase would enhance the overall quality of exploratory data analysis. For example, diagramming for reasoning about the relationship of the data fields, or automatically extracting the marked data points to promote further investigations would be a good addition.

Limitations and Future Work

Our study setting might not reflect various contexts in which exploratory analysis takes place. Our study datasets and procedure were artificially constrained in several aspects. Many analysts' work environments do not include a large touchscreen

device such as a Microsoft Surface Studio yet. Participants were also conscious that they were being observed and timed, thus our observation may not reflect their typical behavior.

We imposed three stages on participants' data analyses to differentiate externalization behaviors between stages. Though we informed participants not to let the protocol prevent them from returning to a previous phase, participants' behaviors might have deviated from their natural flow. We controlled the scope of the dataset given to participants to fit into a relatively short amount of study time, but participants may still have felt rushed to finish a stage. We also provided questions to answer in phase 3, questions they might not ask themselves in an unobserved analysis process.

In reality, an exploratory data analysis can take place over multiple sessions, whereas our studies captured a single session. It would be useful to study how externalization behaviors change over time across multiple sessions in a more natural setting.

Our choice of extending a visualization recommender system meant that participants could avoid manually specifying visualizations if they wanted to. This may have affected their externalization behaviors in some way. For example, those participants who did not manually specify their visualizations may have not thought about the visualizations as intentionally as they would have had they constructed them.

Participants used fewer words and more symbols in an externalization created with a digital pen than with a keyboard. This finding aligns with findings in education (e.g., [2, 31]), showing that handwriting is slower, so people tend to create shorter but more summative notes. Though we did not measure the quality of externalization and the quality of analysis, we expect that long-hand writing may promote a deeper understanding of the data based on participants' testimonials, and we believe that this is an important question for future work. In addition, further research could establish a more prescriptive account of how externalization can more effectively benefit data analysis by investigating how the length and character of externalization impact the quality of the analysis.

CONCLUSION

Despite the benefits of externalizing one's thoughts when interacting with complex information, we have a dearth of empirical knowledge in externalization behaviors in the context of exploratory data analysis. In this paper, we contribute the findings from two exploratory studies that investigated the types and use of externalizations with respect to the stage of exploratory data analysis. In the first study, we examined how people take notes during different stages of data analysis while using conventional visualization annotation capabilities, along with their favorite external note-taking tools. In the second study, we investigated how these externalization behaviors are affected by the ability to add free-form annotations with digital pen input and to link visualizations and externalizations with touch interaction. In addition, we provide a set of design implications and future research opportunities to improve externalization support in future data analysis systems.

REFERENCES

- [1] Annette Adler, Anuj Gujar, Beverly L Harrison, Kenton O'Hara, and Abigail Sellen. 1998. A diary study of work-related reading: design implications for digital reading devices. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM Press/Addison-Wesley Publishing Co., 241–248.
- [2] Nancy M Aguilar-Roca, Adrienne E Williams, and Diane K O'Dowd. 2012. The impact of laptop-free zones on student performance and attitudes in large lectures. *Computers & Education* 59, 4 (2012), 1300–1308.
- [3] Jeffrey Browne, Bongshin Lee, Sheelagh Carpendale, Nathalie Riche, and Timothy Sherwood. 2011. Data analysis on interactive whiteboards through sketch-based interaction. In *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces (ITS)*. ACM, 154–157.
- [4] Stuart K Card, Jock D Mackinlay, and Ben Shneiderman. 1999. *Readings in information visualization: using vision to think*. Morgan Kaufmann.
- [5] Yang Chen, Scott Barlowe, and Jing Yang. 2010. Click2annotate: Automated insight externalization with rich semantics. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST)*. IEEE, 155–162.
- [6] George Chin Jr, Olga A Kuchar, and Katherine E Wolf. 2009. Exploring the analytical processes of intelligence analysts. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 11–20.
- [7] Nelson Cowan. 1998. *Attention and memory: An integrated framework*. Oxford University Press.
- [8] Andrew Crotty, Alex Galakatos, Emanuel Zraggen, Carsten Binnig, and Tim Kraska. 2015. Vizdom: Interactive analytics through pen and touch. *Proceedings of the VLDB Endowment* 8, 12 (2015), 2024–2027.
- [9] Ivan Denisovich. 2005. Software support for annotation of visualized data using hand-drawn marks. In *Proceedings of the IEEE International Conference on Information Visualisation*. IEEE, 807–813.
- [10] Ryan Eccles, Thomas Kapler, Robert Harper, and William Wright. 2008. Stories in geotime. *Information Visualization* 7, 1 (2008), 3–17.
- [11] Sean E Ellis and Dennis P Groth. 2004. A collaborative annotation system for data visualization. In *Proceedings of the working conference on Advanced visual interfaces*. ACM, 411–414.
- [12] David Gotz, Michelle X Zhou, and Vikram Aggarwal. 2006. Interactive visual synthesis of analytic knowledge. In *Visual Analytics Science And Technology, 2006 IEEE Symposium On*. IEEE, 51–58.
- [13] Jeffrey Heer, Fernanda B Viégas, and Martin Wattenberg. 2007. Voyagers and voyeurs: supporting asynchronous collaborative information visualization. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 1029–1038.
- [14] William C Hill and James D Hollan. 1991. Deixis and the future of visualization excellence. In *Proceedings of the IEEE Conference on Visualization*. IEEE Computer Society Press, 314–320.
- [15] Ken Hinckley, Shengdong Zhao, Raman Sarin, Patrick Baudisch, Edward Cutrell, Michael Shilman, and Desney Tan. 2007. InkSeine: In situ search for active note taking. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 251–260.
- [16] Edwin Hutchins. 1995. *Cognition in the Wild*. MIT press.
- [17] Jaemin Jo, Sehi L'Yi, Bongshin Lee, and Jinwook Seo. 2017. TouchPivot: Blending WIMP & post-WIMP interfaces for data exploration on tablet devices. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 2660–2671.
- [18] Youn-ah Kang, Carsten Gorg, and John Stasko. 2009. Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST)*. IEEE, 139–146.
- [19] Youn-ah Kang and John Stasko. 2014. Characterizing the intelligence analysis process through a longitudinal field study: Implications for visual analytics. *Information Visualization* 13, 2 (2014), 134–158.
- [20] Kenneth A Kiewra. 1989. A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review* 1, 2 (1989), 147–172.
- [21] Kibum Kim, Scott A Turner, and Manuel A Pérez-Quirón. 2009. Requirements for electronic note taking systems: A field study of note taking in university classrooms. *Education and Information Technologies* 14, 3 (2009), 255–283.
- [22] David Kirsh. 2010. Thinking with external representations. *AI & Society* 25, 4 (2010), 441–454.
- [23] Bongshin Lee, Petra Isenberg, Nathalie Henry Riche, and Sheelagh Carpendale. 2012. Beyond mouse and keyboard: Expanding design considerations for information visualization interactions. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 18, 12 (2012), 2689–2698.
- [24] Bongshin Lee, Rubaiat Habib Kazi, and Greg Smith. 2013. SketchStory: Telling more engaging stories with data through freeform sketching. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 19, 12 (2013), 2416–2425.

- [25] Bongshin Lee, Greg Smith, Nathalie Henry Riche, Amy Karlson, and Sheelagh Carpendale. 2015. SketchInsight: Natural data exploration on interactive whiteboards leveraging pen and touch interaction. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, 199–206.
- [26] Guang Li, Xiang Cao, Sergio Paolantonio, and Feng Tian. 2012. SketchComm: a tool to support rich and flexible asynchronous communication of early design ideas. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 359–368.
- [27] Narges Mahyar, Ali Sarvghad, and Melanie Tory. 2010. A closer look at note taking in the co-located collaborative visual analytics process. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST)*. IEEE, 171–178.
- [28] Narges Mahyar, Ali Sarvghad, and Melanie Tory. 2012. Note-taking in co-located collaborative visual analytics: Analysis of an observational study. *Information Visualization* 11, 3 (2012), 190–204.
- [29] Narges Mahyar and Melanie Tory. 2014. Supporting communication and coordination in collaborative sensemaking. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of VAST)* 20, 12 (2014), 1633–1642.
- [30] Fabrice Matulic and Moira C Norrie. 2012. Supporting active reading on pen and touch-operated tabletops. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. ACM, 612–619.
- [31] Pam A Mueller and Daniel M Oppenheimer. 2014. The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological science* 25, 6 (2014), 1159–1168.
- [32] Donghao Ren, Matthew Brehmer, Bongshin Lee, Tobias Höllerer, and Eun Kyoung Choe. 2017. ChartAccent: Annotation for data-driven storytelling. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, 230–239.
- [33] Yann Riche, Nathalie Henry Riche, Ken Hinckley, Sheri Panabaker, Sarah Fuelling, and Sarah Williams. 2017. As we may ink?: Learning from everyday analog pen use to improve digital ink experiences.. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. 3241–3253.
- [34] Hugo Romat, Nathalie Riche, Ken Hinckley, Bongshin Lee, Caroline Appert, Emmanuel Pietriga, and Christopher Collins. 2019. ActiveInk:(Th) Inking with Data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM.
- [35] Kathy Ryall, Neal Lesh, Tom Lanning, Darren Leigh, Hiroaki Miyashita, and Shigeru Makino. 2005. Querylines: Approximate query for visual browsing. In *Extended Abstract Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 1765–1768.
- [36] Bill N Schilit, Gene Golovchinsky, and Morgan N Price. 1998. Beyond paper: Supporting active reading with free form digital ink annotations. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 249–256.
- [37] Yedendra Babu Shrinivasan and Jarke J van Wijk. 2008. Supporting the analytical reasoning process in information visualization. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 1237–1246.
- [38] Jagoda Walny, Samuel Huron, and Sheelagh Carpendale. 2015. An exploratory study of data sketching for visual representation. *Computer Graphics Forum (Proceedings of EuroVis)* 34 (2015), 231–240.
- [39] Jagoda Walny, Samuel Huron, Charles Perin, Tiffany Wun, Richard Pusch, and Sheelagh Carpendale. 2018. Active reading of visualizations. *IEEE transactions on visualization and computer graphics* 24, 1 (2018), 770–780.
- [40] Jagoda Walny, Bongshin Lee, Paul Johns, Nathalie Henry Riche, and Sheelagh Carpendale. 2012. Understanding pen and touch interaction for data exploration on interactive whiteboards. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 18 (2012), 2779–2788.
- [41] Martin Wattenberg. 2001. Sketching a graph to query a time-series database. In *Extended Abstract Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 381–382.
- [42] Pierre Wellner. 1993. Interacting with paper on the DigitalDesk. *Commun. ACM* 36, 7 (1993), 87–96.
- [43] Kanit Wongsuphasawat, Dominik Moritz, Anushka Anand, Jock Mackinlay, Bill Howe, and Jeffrey Heer. 2016. Voyager: Exploratory analysis via faceted browsing of visualization recommendations. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 22, 1 (2016), 649–658.
- [44] William Wright, David Schroh, Pascale Proulx, Alex Skaburskis, and Brian Cort. 2006. The Sandbox for analysis: concepts and methods. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 801–810.
- [45] Haijun Xia, Nathalie Riche, Fanny Chevalier Bruno De Araujo, and Daniel Wigdor. 2018. DataInk: Enabling direct and creative data-oriented drawing. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*.
- [46] Emanuel Zraggen, Robert Zeleznik, and Steven M Drucker. 2014. PanoramicData: Data analysis through pen & touch. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 20, 12 (2014), 2112–2121.

- [47] Jian Zhao, Michael Glueck, Simon Breslav, Fanny Chevalier, and Azam Khan. 2017. Annotation graphs: a graph-based visualization for meta-analysis of data

based on user-authored annotations. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of VAST)* 23, 1 (2017), 261–270.