Evaluating the Role of Time in Investigative Analysis of Document Collections

Bum Chul Kwon, *Student member, IEEE*, Waqas Javed, *Student member, IEEE*, Sohaib Ghani, *Student member, IEEE*, Niklas Elmqvist, *Member, IEEE*, Ji Soo Yi, *Member, IEEE*, and David Ebert, *Fellow, IEEE*

Abstract—Time is a universal and essential aspect of data in any investigative analysis. It helps analysts establish causality, build storylines from evidence, and reject infeasible hypotheses. For this reason, many investigative analysis tools provide visual representations designed for making sense of temporal data. However, the field of visual analytics still needs more evidence explaining how temporal visualization actually aids the analysis process, as well as design recommendations for how to build these visualizations. To fill this gap, we conducted an insight-based qualitative study to investigate the influence of temporal visualization on investigative analysis. We found that visualizing temporal information helped participants externalize chains of events. Another contribution of our work is the lightweight evaluation approach used to collect, visualize, and analyze insight.

Index Terms—Qualitative evaluation, investigative analysis, temporal visualization, insight-based evaluation.

1 Introduction

Time has always received special treatment in the visualization literature [1]. It is used for a wide variety of tasks such as understanding causality, discovering trends, and predicting future events. A wide array of techniques has been developed over the years for visualizing temporal data, such as for time-varying quantitative data, event sequences, and storytelling.

Despite much of this prior work including results from empirical user studies, there exists very little knowledge on the actual role of temporal data and temporal visualization for **investigative analysis** [2]. In the field of visual analytics the concept of time is central to the analytical process [1], [3], [4] and widely utilized in many visual analytics tools. However, there has not been sufficient work on the role and impact of temporal information on the thinking process of an investigative analyst. In particular, recent progress in empirical evaluation of visual analytics systems [4], [5], [6], [7] have failed to clearly deal with this topic.

In this paper, we attempt to address this issue by presenting and discussing results from a qualitative evaluation comparing the performance of participants conducting a investigative analysis task using a visual analytics tool with and without access to temporal visualization. It should be noted that the purpose of this work is **not** to answer the question whether temporal information and temporal visualization is useful or not—the answer to this question is a clear "yes"—but rather to study differences in how users utilize temporal information when it is explicitly presented in a temporal visualization, as opposed to when no such visualization is available. Our ambition is that these findings will in turn allow us to derive practical and workable results that have general application across a wide array of visual analytics tools.

Having said that, it is important to realize that evaluation of visualization and visual analytics is difficult and still in its infancy [8], [9], [10]. While the field of human-computer interaction has a long tradition of performing aptitude tests on low-level cognitive and perceptual tasks, it is not clear that extrapolating such tradition to higher-level sensemaking and decision making tasks is possible [11]. The overarching investigative analysis task is generally too individual, volatile, and amorphous to afford quantitative evaluation and comparison. Our evaluation is, therefore, qualitative in nature, and we make no efforts to derive quantitative measures on time and error, which is often meaningless in the context of investigative analysis. Instead, our findings revolve around observation, semi-structured interviews, and informal performance analysis. We thus follow in the footsteps of Kang et al. [4] but focus on a hitherto neglected aspect of visual analytics. We have also quite deliberately taken a lightweight approach to this qualitative comparison that we think may be of general use for evaluating visual analytics tools.

Thus, we see the main contributions of this paper as the following: (i) results and observations from a qualitative comparison of investigative analysis with and without access to temporal visualization; (ii) design implications on how to best design and utilize temporal visualization in visual analytics tools; and

- *B.C. Kwon and J. S. Yi are with the School of Industrial Engineering, Purdue University, West Lafayette, IN, USA.*
  E-mail: {kwonb, yij}@purdue.edu
- *W. Javed, S. Ghani, N. Elmqvist, and D. Ebert are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA.*
  E-mail: {wjaved, sghani, elm, ebertd}@purdue.edu

(iii) a novel evaluation approach for *lightweight qualitative comparison* that strikes a balance between time and cost versus depth and explanatory power.

2 Background

This work explores the role of time and temporal visualization in *investigative analysis*. In this background section, we first discuss how temporal information has been used in intelligence analysis and existing work that supports temporal data and visualization. We then motivate our evaluation by reviewing the state of the art in visual analytics evaluation.

2.1 Investigative Analysis and Time

Investigative analysis is defined as making discoveries and finding hidden truths in large collections of data [2]—in a way, detecting the expected and discovering the unexpected. It is a cognitively taxing task performed by a wide variety of user groups including business analysts, journalists, scientists, intelligence analysts, and law enforcement officers.

More specifically, investigative analysis involves understanding the connections, causality, and relationships between different entities, scattered in multiple documents, collected from multiple sources, and represented in multiple different formats [2]. Due to the limited working memory of the human mind, the analytical process becomes increasingly difficult as the number of entities involved in the analysis grows [12].

In addition to the complexity and large scale of the data, identifying potential explanations and hypotheses, as well as testing these hypotheses by finding evidence from collected data, is an onerous task. Because of this problem, the dominant approach of investigative analysts is similar to that of historians rather than that of scientists [13]. Instead of deriving all possible cases and scientifically evaluating them, which is often difficult if not impossible, they tend to find a coherent narrative to explain the interesting phenomena. Thus, collecting evidence to confirm or reject hypothetical stories is an important step for intelligence analysts [14]. Because of this tendency of creating stories out of evidence, time becomes very important in investigative analysis. Time is essential for suggesting and sometimes determining sequential orders, thereby clarifying cause and effect relationships. Temporal information can also be used to rule out unlikely hypotheses (e.g., if Bob's visit to a particular place happened much earlier than a suspicious event, his visit may not be related to the event) and identify impossible hypotheses (e.g., Bob was killed in 1998, so he cannot have bombed a building in 2000).

However, investigating the temporal aspects of evidence on top of already complex data is challenging. Visualization and visual analytics can aid the analysis by providing effective graphical representations of the available information—what cognitive scientists call *external cognition* [15], [16]—and by allowing for interactive exploration of these representations.

2.2 Temporal Visualization and Analysis

As discussed in Section 2.1, time is an inherent dimension in data analysis because of its unique semantic meaning; for this reason, it has always received special attention in visualization [1]. Therefore, it is not surprising that many visual analytics tools include some functionality for temporal analysis or temporal visualization. The most common mechanism is the timeline view, where events and time-varying data are visualized on a chart where one dimension (often the horizontal) is time. Another example is representing temporal dynamics in geo-spatial visualizations [17]. For investigative analysis, the data usually consist of discrete events in time, so we largely ignore the considerable body of work on visualization of time-varying or time-series data. Examples of discrete event timeline views can be found in Jigsaw [2], LifeLines [18], [19], and Similan [20]. Certain tools take temporal aspects a step further: GeoTime [3] has a story building mechanism for constructing narratives from event sequences, and CzSaw [21] maintains analysis provenance to facilitate reflection and replay.

Despite the prevalence of temporal visualization, we have not been able to find any studies that particularly investigate the role of temporal visualization in the analytical process. In fact, there exists very little work that empirically studies the analytical process in general (e.g., [4]), let alone its temporal aspects.

2.3 Evaluating Visual Analytics

Most empirical evaluations in visualization and visual analytics study low-level analytical tasks like search, navigation, and queries, and are therefore more of a physical aptitude than a cognitive nature [9]. Only a few studies investigate higher-order analytical activities like sensemaking, decision-making, or even comparison, correlation, and organization.

In a sense, this dearth of empirical knowledge is an effect of the difficulty of evaluating visualization in general [9], [22], and investigative analysis in particular [4]. This is mostly due to the open-ended nature of many visual analytics tasks, which makes drawing clear conclusions from quantitative data difficult [23]. In fact, it is sometimes difficult to even *collect* quantitative data in the first place: what should really be measured? This is also the reason for the heavy emphasis on more qualitative and exploratory user studies of visual analytics tools in the literature.

Many such existing studies are relevant to our purposes. In separate work, Bier et al. [5] and Jeong et al. [6] studied quantitative performance for professional analysts solving sensemaking tasks in intelligence and financial analysis, respectively. Similarly, Isenberg et al. [24] and Robinson [25] independently performed exploratory studies of collaborative sensemaking in paper-based settings; both of these papers are particularly interesting due to their use of timeline visualizations to present results. Gotz and Wen [26] conducted an empirical study of user interaction behavior during visual analysis to propose general guidelines for user-driven visual analytics tools. Mark and Kobsa [14] compared group and individual performance with collaborative information visualization environments through a quantitative experiment and derived a stage model that explains the users' collaboration process. Park et al. [27] qualitatively reviewed how collaborators in virtual environments work together to perform several tasks on visualized oceanographic data. Also of interest is the description of Plaisant et al. [10] of how the VAST contest judged the utility of the submitted visual analytic tools.

The recent qualitative study [4] on the Jigsaw [2] system is particularly instructive. The authors conducted a between-participants study, which divided participants into groups of four, each group having only partial access to a subset of functions of Jigsaw, and made all of them work on the same intelligence reports to identify a fictional terrorist plot. Two external raters graded the score of findings in the experiment based on correctness of answers and also provided subjective grade on narrative debriefings. They also measured the elapsed time, amount of notes, documents viewed, etc. Rather than providing statistical measures, they tried to deduce particular strategies being used in each group to understand the role of the visual analytics system in the analysis.

Insight-based evaluation is another option. Saraiya and her colleagues [28] used insight reports to collect findings for microarray data analysis. The reports were then evaluated and proved helpful for understanding what kinds of insights that participants generated while using different tools. This kind intrusive methodology rather than measuring time and accuracy of users' performance could be useful in order to capture the cognitive analysis process [29].

3 Investigative Analysis Software

As our literature review shows, there already exists a number of investigative analysis tools such as Jigsaw [2], CzSaw [21], and Analyst's Notebook [30] that we could use in our evaluation of temporal visualization. However, these tools require significant training and they often include many different views and methods for solving a particular task. In addition, it is difficult for outsiders to instrument these tools to collect user interaction data (e.g., click stream). Therefore, we felt that a better approach would be to identify the canonical tasks in investigative analysis and develop a minimalistic tool that supports them.

3.1 Canonical Tasks

Based on our focus on time and on the interaction categories proposed by Yi et al. [31], we derive the below canonical tasks for investigative analysis:

- **Reading documents:** Reading is a central activity in investigative analysis [2]. (*Elaborate* [31])
- **Viewing relationships:** Relationships between entities suggest association, information exchange, and causality. (*Connect* [31])
- **Selecting:** Marking entities allows for structuring work and correlating relationships. (*Select* [31])
- **Filtering:** Entities that are irrelevant to the analysis should be possible to discard. (*Filter* [31])
- **Viewing temporal relationships:** Causality is an important relationship [1]. (*Reconfigure* [31])

3.2 TimeInvestigator

Guided by these canonical tasks, we developed an investigative analysis tool, called TimeInvestigator, consisting of five cross-linked views where an operation in one view (e.g., selecting and filtering) would affect all other views accordingly (Figure 1):

- The **Entity-Relationship view** shows entities and their co-occurrences using a graph (Figure 1(a));
- The **Timeline view** shows entity occurrences on a timeline (Figure 1(b));
- The **Document view** shows reports with the entities highlighted and color-coded (Figure 1(c));
- The **Document list** shows names and dates of currently matched documents (Figure 1(d)); and
- The **Recycle bin** contains entities that have been removed from other views; e.g., filtering out irrelevant entities from the Timeline (Figure 1(e)).

Using these views, users were able to dynamically add and remove entities from the application—this essentially meant moving entities to and from the recycle bin. On starting up the application, no entities were shown in the main views. The analyst could then add whole ranges of entities, or just select a few.

3.3 Entity-Relationship View

The Entity-Relationship (ER) view (Figure 1(a)) is the main view of TimeInvestigator and is designed to partly mimic the Graph view of Jigsaw. The view displays the entities in the document collection as nodes and their co-occurrence in documents as link relations between the nodes. Nodes are labeled with their entity names and are color-coded depending on their type (i.e., Places in blue, Organizations in green, and Persons in red). Nodes can be moved so that the user can partition the space during the analysis.

Beyond browsing, the view also supports free text search using a query box (top right in Figure 1(a)). Matched nodes are highlighted in yellow. Finally, the ER view also incorporates an entity legend (just above the query box in Figure 1(a)) that supports toggling visibility of nodes by entity type simply by clicking on the label. Finally, entities can be filtered out using a double-click (sending them to the recycle bin).

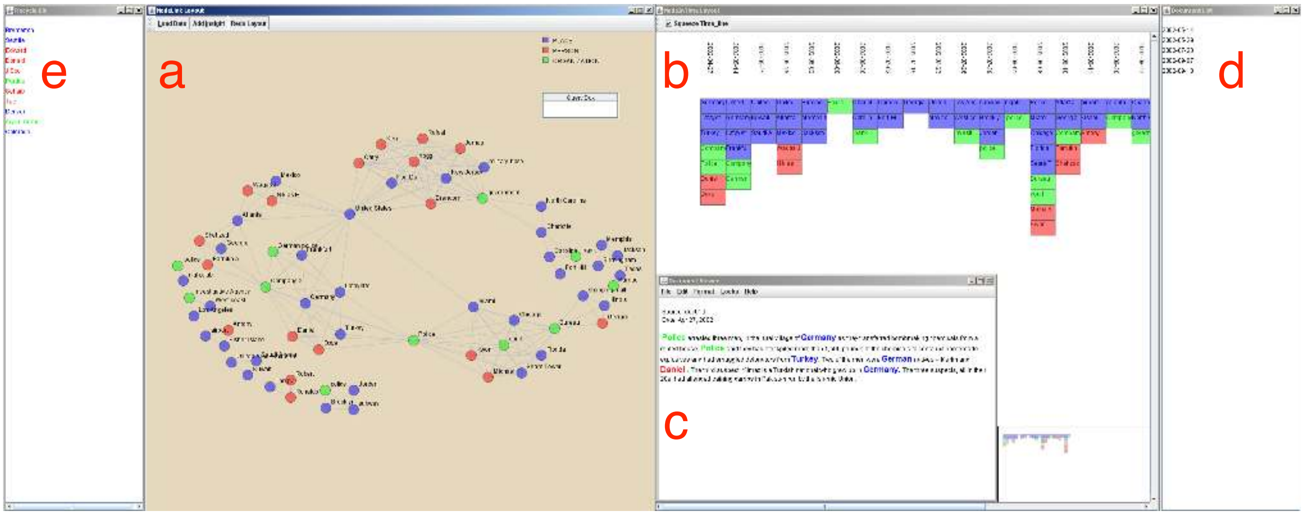

Fig. 1. Overview of the TIMEINVESTIGATOR tool. (a) Entity-Relationship view for color-coded entities and their co-occurrence in reports. (b) Timeline view showing entity distribution in time. (c) Document view with color-coded entities highlighted. (d) Document list for currently matched documents. (e) Recycle bin showing entities that have been discarded as irrelevant. Blue entities are places, red are persons, and green are organizations.

3.4 Timeline View

The Timeline view is a temporal visualization that displays entity occurrences organized along a temporal axis (Figure 1(b)). This is done by aggregating all reports in the document collection by their dates, and then showing all of the entities for each particular date. Each entity in the Timeline view is represented by a labeled box that is color-coded according to the entity type. Since a single entity may appear in more than one report at different dates, entity boxes may be duplicated for several dates along the timeline.

For each date, entity boxes are grouped according to their type to make the display consistent; for our example in Figure 1(b), we group entities in the order of "Place", "Organization", and "Person" from top to bottom. Furthermore, the order within each entity type group depends on the number of occurrences of a particular entity in the whole document collection, organized in descending order (i.e., the entity with most occurrences is placed at the top). The view can be scrolled horizontally to support long event sequences contained in large document collections. A small viewport in the right bottom corner shows an overview of the whole timeline to aid overview and navigation. To further ease temporal navigation, the user can toggle timeline compression, where all dates containing no currently selected entity are removed. Finally, a user can remove an entity using a double-click (sending it to the recycle bin).

3.5 Reading Documents

The Document list enumerates documents where matched or selected entities occur (Figure 1(d)), and updates as the user selects, queries, and filters entities. The Document view allows for reading the actual text of reports in the document collection (Figure 1(c)), a vital part of investigative analysis [2]. Just like in Jigsaw, the view highlights entities using color-coding based on type, and also draws them using a bigger font. Any number of reports can be open at a time.

3.6 Recycle Bin

The Recycle bin is a list of entities that have been removed (i.e., filtered out) in order of removal (Figure 1(e)). Entities can be returned into the dataset by double-clicking on its entry. The ER and Timeline views also support undo and redo for the delete operation, moving entities to and from the recycle bin.

While the recycle bin is not typically a canonical component found in investigative analysis tools, its existence in TIMEINVESTIGATOR is a side effect of the decision to allow participants to discard entities from views and recover those if necessary. Discarding entities corresponds to filtering in many existing tools. Furthermore, no participant reported difficulties in understanding this function, probably because it resembles the trash can in major operating systems.

3.7 Recording Insights

For the purposes of performing insight-based evaluation [28], we created a view called the Insight Report view. A new insight report can be generated at any time at the click of the "Add Insight" button; doing so will take a screen capture of the whole TIMEINVESTIGATOR desktop, and will open a text field where the user can type in free text related to the insight. The insight report also asks which view (Document, Entity-Relationship, or Timeline view) helped inspire the insight. The text, screenshot, and time stamp are saved when submitting the report.

The insight report view was a pure byproduct of the evaluation methodology, and we tried to minimize its impact on the analysis process. In particular, existing reports created at an earlier stage were **not** available for consultation at later stages. In the study (see below), we stressed the need to report findings using this mechanism, but we did not explicitly remind participants to do this during sessions. Despite these steps to minimize its impact, it is entirely possible that the inclusion of insight report generation changed the structure of the analysis process; this was also noted by Saraiya et al. [28] in their original work.

4 EVALUATION

Our ambition with this work, as noted above, was to study the influence of temporal visualization on investigative analysis of document collections. Below we discuss the general method we employed, as well as specifics on participants, equipment, and task.

4.1 Method

Out of the evaluation methods reviewed in the literature, we found the controlled study approach by Kang et al. [4] that involved single non-expert users in contrasting conditions working on an extensive constructed scenario with ground truth to be the most appropriate for our work. We therefore decided to adopt this methodology, but to reduce time investments by not video recording sessions, and instead to use a combination of observations, screen captures, click streams, and insight-based evaluation [28] to collect deeper insights about the analytical process.

We call this *lightweight qualitative comparison*, and submit that it may be a useful evaluation method that strikes a balance between in-depth qualitative (or even ethnographic) evaluation performed using domain experts, and low-overhead quantitative evaluation involving non-expert participants.

4.2 Participants and Apparatus

We recruited 12 paid participants ($10 per hour)—7 males and 5 females recruited from the engineering student population at our university—randomly divided into two groups: 6 participants with **no** access to the Timeline view in the TIMEINVESTIGATOR tool (Group N), and 6 participants with full access to the Timeline (Group T). The reason for choosing students as participants as opposed to professional analysts is that we were unable to get access to such analysts in our traditional university setting. We discuss the implications of this limitation further in Section 6.6.

The study was performed on a desktop computer equipped with two 19″ monitors (1280 $\times$ 1024 pixels) to accommodate the multiple views of TIMEINVESTIGATOR. Participants were not told the name of the tool, nor the special emphasis on temporal analysis to minimize any unexpected biases. Prior to starting the experiment, participants underwent a training session of approximately 20 minutes using a dummy dataset.

4.3 Task

The task consisted of identifying a hidden terrorist plot in a collection of 50 fictional intelligence reports. This dataset was the same that was used in the recent evaluation by Kang et al. [4]. Participants were allowed to take up to one hour to complete the task and were encouraged to make use of the full time.

Participants were instructed to create insight reports whenever they learned something significant about the document collection. They were told that these reports would be the main evaluation instrument in the study, and thus that creating reports was important.

Upon finishing the experiments, participants were told to write a short narrative on the suspected terrorist plot. They were then issued a questionnaire on their experiences of the method, strategy, and view primarily used to perform the task. Finally, we also conducted exit interviews with all participants.

4.4 Measures

We collected several measures to understand the experiences of the two groups of participants (Groups N and T), including interviews and insight reports (which view a participants got the insight from, text, and screen capture). We also instrumented the TIMEINVESTIGATOR tool to collect participant usage patterns (i.e., the uses of the different views and clicks of entities with timestamps). The purpose was to use these quantitative measures to aid our understanding.

We coded insight reports systematically using two coders, who are also authors of this paper, working independently and using a shared coding rubric (Cohen's kappa coefficient = 0.49, which is considered "good" or "moderate"). The independent code streams were then merged, discussed, and unified. Score was based on five main plot points (and a number of subplots per plot point) that we had extracted from the ground truth of our dataset; taken together, these five plot points explained the full story. Every plot point that was discussed in an insight report was scored from 0.0 to 1.0 depending on the accuracy of the insights and the coverage of subplots, so each insight report was scored between 0.0 and 5.0. The final score per participant is based on the accuracy and comprehensiveness of cumulative insight reports.

Some participants falsified—i.e., disproved certain elements of the dataset as not being relevant or correct—story components that were not included in the main plot and/or reported incorrect speculation and confirmation. Though these are notable aspects of insight reports, including them into scores is problematic (e.g., PT3 successfully falsified story components in 10 different insight reports, but how much is each successful falsification worth?). Thus, such elements are separately codified and not included in the scores.

5 Results

We collected results using a combination of interaction logs, observations, interviews, and insight reports.

5.1 Visualizing Evaluation Results

To aid our understanding of participant analysis processes, we decided to visualize our study results. Inspired by timeline visualizations created by Isenberg et al. [24] and Robinson [25], we created the visualization in Figure 3 to show the temporal event sequences we collected during the study: which view a participants interacted with, when an insight report was submitted, the individual scores for insight reports, and the number of entities visible in the tool. Figure 2 gives a legend to aid in understanding Figure 3. Some notable observations follow:

First, the lack of patterns in the visualization suggests the great variation in analysis method between individuals. Participants demonstrated wide variation in their final scores, how frequently insight reports were submitted (e.g., PN4, PT2, PT3, and PT6 submitted insight reports more frequently than the average across all participants, but PN1 and PN5 did less frequently), and which views they frequently used (e.g., PT4 used the Timeline view heavily, but PN4, who did not have access to the Timeline, mainly used the ER and the Document list). These large differences made us doubt that simply recruiting more participants would yield statistically significant results.

Second, as discussed previously, Group N viewed fewer entities in the TimeInvestigator tool than Group T when their first insight reports were submitted. The black solid lines in the colored band in Figure 3 indicate that the numbers of entities are generally increasing in Group N as time progresses, but decreasing in Group T. In other words, Group N seemed to base their analysis on progressively adding supporting evidence, whereas Group T instead iteratively removed circumstantial or unrelated evidence. This result is in line with the interview results reported in Section 5.2. Participants in Group N seemed to have difficulties in dealing with many entities, so they tended to carefully add entities. In contrast, ones in Group T, except for PT1, started with almost all entities on the screen and progressively filtered out ones that were irrelevant.

Third, some participants (PT3 and PT6) in Group T successfully falsified irrelevant plot points, indicated in green boxes in Figure 3, but Group N participants collectively demonstrated only four successful falsifications. Actually, insight reports submitted by PN3,

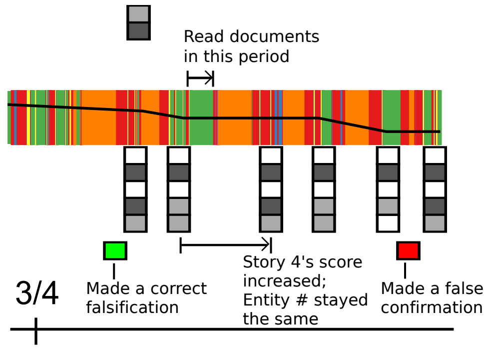

Fig. 2. Key for the activity timeline in Figure 3.

PN4, and PN6 included quite a few incorrect speculations and confirmations, indicated by red boxes.

These findings suggest that Groups T and N employed significantly different analysis methods, and the likely cause for this is the absence or presence of the Timeline view. The benefits of the Timeline seems to be in organizing and externalizing the events and then helping in discarding irrelevant entities. Below we study these aspects in more depth.

5.2 Interview

The results of interviews revealed three notable benefits of the Timeline view: for (1) making sense of the order of event chains; (2) identifying important nodes; and (3) discovering relevant documents easier.

The first and rather obvious benefit of the Timeline view is that it helps making sense of the sequential and/or logical order of event chains, pointing to the prevalence of stories in temporal reasoning. When asked about how the Timeline view helped, PT1 said, “[It helps to] figure out how relationships change over time.” PT2 also added, “The timeline helped me understand the order [of events].”

Second, the Timeline view also seemed helpful in identifying important nodes. In a question asking “Please describe how you knew you had found the main plot,” PT4 remarked, “The frequency of a node appears to be significant. If a node appears multiple times in the [Timeline] view, it is more important.” PT5 also said, “[In the Timeline view,] strong connections can be shown if entities show up multiple times.” PT6 described his or her strategy as “See a person first. Follow the timeline. See if they are linked to the plot. Pay attention to areas with lots of blocks [i.e. events] in the Timeline view.” In contrast, Group N participants seemed to have difficulties in discerning which is important or not. For example, PN3 reported, “[I] keep them all till the end because there is no way to decide if any of them are important.”

Third, Group T seemed to find it easier to discover relevant documents than Group N. The Timeline view not only shows the temporal information but also serves as an easy access to documents ordered

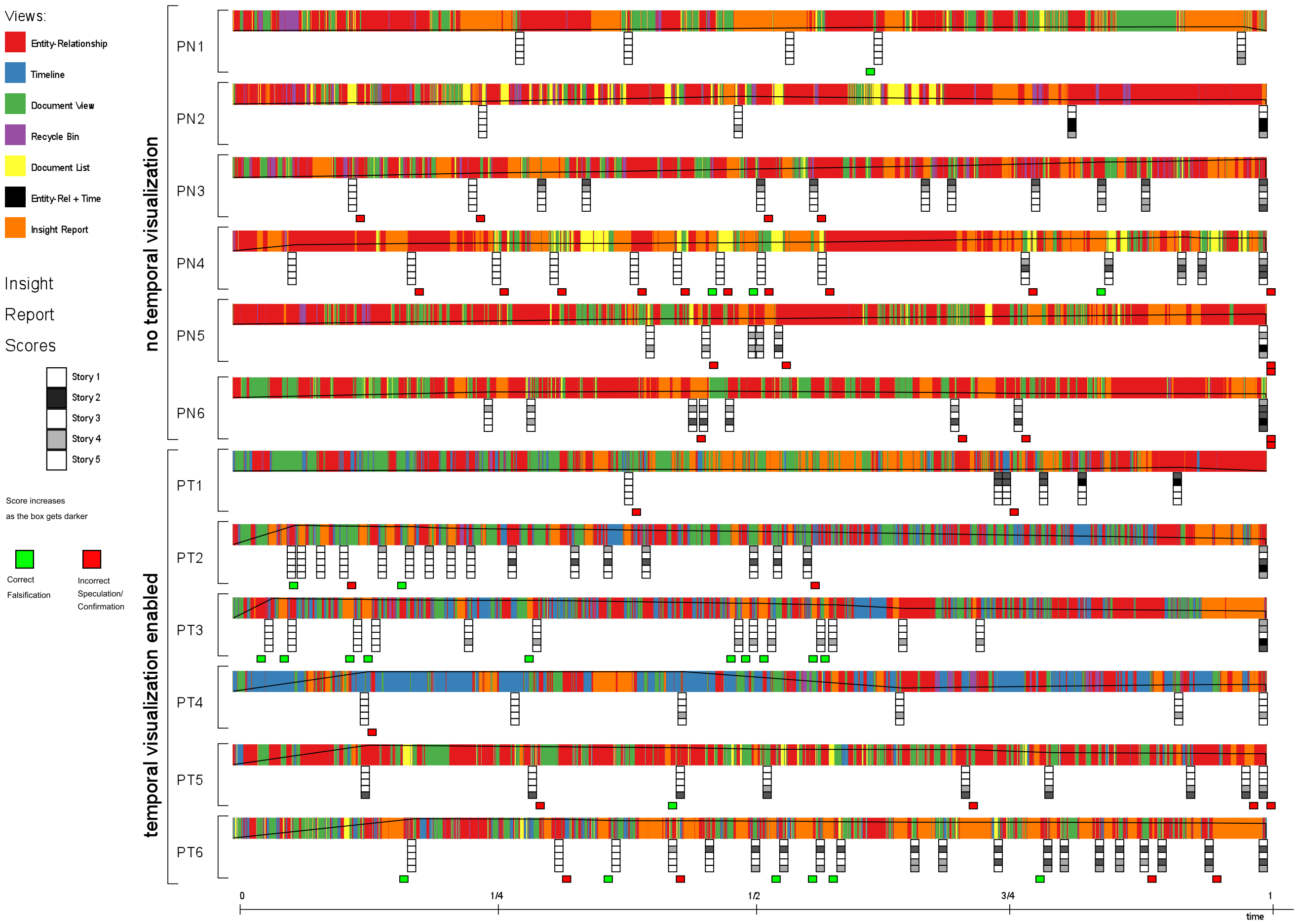

Fig. 3. Activity timeline for study sessions for all 12 participants (inspired by Isenberg et al. [24] and Robinson [25]). The color-coded band indicates which view the participants were interacting with. Black lines across the color-coded band show the number of entities shown in the ER view. A stack of boxes under the color-coded band shows the score of an insight report. Each box in a stack corresponds to one of the five major storylines; higher scores results in darker color of each box for that particular storyline.

chronologically, which helped users to easily follow suspicious entities over time. (PT2: "Read through time. [Find] what they did in the past and the future." and PT3: "With the remaining nodes, [I] used the Timeline view to make sense of stories."). In contrast, interviews with Group N hinted at the impact of having no access to temporal visualization: they were forced to read a lot of reports. When asked "How did you go about confirming that the plot is actually threatening?," Group N participants stated that they had to search multiple documents without any order. (PN1: "I could confirm the story only after reading the whole document" and PN5: "[I] verified it by reading reports associated with it. I would find the suspected node and verify it by using the [Document] view.").

5.3 Summary Statistics

First, we reviewed summary statistics to see if there were any significant differences between Groups N and T. However, because the number of participants is small, it was difficult to calculate statistical analyses with reasonable confidence. As expected, we did not see many notable differences between the two groups, as evidenced by Tables 1 and 2.

The only statistically significant difference was the number of entities that they placed on the ER view[1] (not in the Recycle bin) when they submitted the first insight report. This is in line with our observations from Section 5.1. Except for PT1, all Group T participants had over 100 entities on the views when the first insight reports were submitted. In contrast, Group T participants had on average less than 40 entities. This difference is statistically significant (Mann-Whitney $U = 4, p = 0.0292$). However, the difference in the number of entities diminished when they submitted the final insight reports (no statistical difference).

We also investigated how much time was spent for each view and from which views participants gained their insights (see Table 3). For both groups, participants tended to spend most of their time on

1. The number of entities on the Timeline view is identical to those on the ER view because the views are synchronized.

TABLE 1
Summary statistics of all participants with no access to temporal visualization (Group N).

	Group N							Statistics
Measures	**PN1**	**PN2**	**PN3**	**PN4**	**PN5**	**PN6**	**Avg**	**Mann-Whitney U**
Final score	0.37	2.17	1.45	0.98	1.23	2.42	1.44	$U = 22, p = 0.5887$
Earliest insight report (min:sec)	16:04	20:23	7:08	3:40	29:15	15:14	15:00	$U = 25, p = 0.3095$
# of insight reports	5	4	12	14	6	8	8.18	$U = 9.5, p = 0.1962$
# of entities in initial insight report	6	19	11	40	36	40	25.33	*$U = 4, p = 0.0292$
# of entities in final insight report	29	30	120	85	69	27	60	$U = 17, p = 0.9361$

TABLE 2
Summary statistics of all participants with access to temporal visualization (Group T).

	Group T							Statistics
Measures	**PT1**	**PT2**	**PT3**	**PT4**	**PT5**	**PT6**	**Avg**	**Mann-Whitney U**
Final score	1.17	1.73	1.67	0.20	0.63	1.27	1.11	$U = 22, p = 0.5887$
Earliest insight report (min:sec)	25:35	4:02	2:11	6:53	7:24	10:35	9:00	$U = 25, p = 0.3095$
# of insight reports	6	17	14	6	9	20	12	$U = 9.5, p = 0.1962$
# of entities in initial insight report	12	123	127	127	127	126	107	*$U = 4, p = 0.0292$
# of entities in final insight report	28	36	44	44	71	95	53	$U = 17, p = 0.9361$

TABLE 3
Time usage and insights generated from each view.

	Group N		Group T	
TimeInvestigator view	**Avg.**	**S.D.**	**Avg.**	**S.D.**
Usage times (min:sec):				
Entity-Relationship	36:01	10:01	22:41	6:34
Timeline view	–	–	11:30	7:40
Document list	8:43	4:38	1:18	1:52
Recycle bin	3:58	3:28	0:50	0:43
Document view	11:29	3:30	14:26	5:03
Insight report	10:38	4:42	9:48	5:31
Number of insights:				
Entity-Relationship	2.83	2.79	4.17	3.87
Timeline view	–	–	2.5	2.59
Document view	5.33	3.83	5.67	6.09

the ER view, which is somewhat surprising to us because we had anticipated that the Document view would consume significant amount of time because they should read reports anyway to know the details. It is also interesting to see that the majority of reports were based on insights gained from the Document view (see Table 3), a result consistent with Kang et al. [2]. However, there is no statistically significant difference between the two groups except for those due to the presence and absence of the Timeline view.

5.4 Insight Reports

The insight reports proved to be a rich source of qualitative information on the investigative strategies employed by our participants. One main finding is that the timeline helped Group T participants in finding the correct results more quickly. Below we pull out the main such trends and discuss them in depth.

5.4.1 *Falsification*

We first studied insight reports submitted by PN4, whose reports had ten instances of incorrect speculations/confirmation. What we found is that PN4 often started with a suspicious activity based on the layout of entities in the ER view (PN4-0[2]: "There is a group of 6 people that communicate with each other a lot. They could be plotting something.") The suspicion continued in PN4-1 and ended at PN4-7, when PN4 opportunistically found evidence showing that PN4's initial suspicion was incorrect. While proceeding with the investigation, PN4 basically found initial cues, uncovered additional entities related to the clue, and expanded the network around these initial suspects. The procedure was then repeated. This seems to be a fairly natural investigative analysis process that mixes intuition and guesswork with evidence and reasoning, and is consistent with earlier results [4], [24], [25].

We found that the Timeline view for Group T had a significant impact on the investigative process. Some participants made falsifications based on the duration of entities co-occurring throughout the timeline. For example, PT3 simply removed three entities based on their occurrence (PT3-0: "Since **Robert D'Onfrio**, **Tampa**, and **Jesus Vazquez** were only mentioned once and mentioned in the same report, I am removing them as suspects"). If PT3 would not have had access to the Timeline view, PT3 may have followed more wrong leads similar to PN4 because the three entities look closely connected in the ER view. Furthermore, in one instance, PT6 stopped tracing a person because the person had no appearance after a specific time (PT6-2: "**Julio** and **David** were removed since their act doesn't connect with the terrorist attack at this moment.").

In summary, Group N participants lacked the additional cues that Group T had from the Timeline view. The lack of these additional cues made Group N par-

2. PN4-0 stands for the first insight report from participant PN4. Note that the insight report number starts from 0.

ticipants (particularly, PN3, PN4, and PN6) consider irrelevant information as a part of main plots.

5.4.2 Alias Detection

The 50 intelligence reports contain three aliases that are crucial for understanding the terrorist network because seemingly disconnected networks suddenly become connected when two separate names turn out to denote the same person. Although we cannot show statistical significance, we found that Group N noticed fewer instances of such aliases and made more mistakes in dealing with aliases than Group T. More specifically, Group N correctly identified a total of five aliases in PN1-4, PN2-2, PN5-0, PN5-2, and PN6-0 while Group T identified nine in PT1-1, PT1-3, PT2-16, PT3-12, PT3-13, PT5-3, PT6-9, PT6-11, and PT6-17. In addition, PN4 and PN6 treated one person with two aliases as two separate people (PN4-9 and PN6-2), and PN5 found a wrong alias for an entity (PN5-1), while no participants in Group T made such mistakes.

Note that identifying such aliases does not require a global understanding of the terrorist network. Instead, it requires simply reading a specific document containing evidence like "**Abu H.**, who was released from custody after the September 11 incidents and whose fingerprints were found in the U-Haul truck rented by **Arnold C.** (report on 2002-10-22)" Thus, the performance of identifying aliases largely depends on the ability to identify such a document.

Even after investigating all of the insight reports and associated screenshots relevant with these aliases, we failed to find a single and universal explanation why there is a difference in identifying aliases between the two groups. One speculation is that it is a mere positive side-effect of the Timeline view as an additional overview that helped participants find more relevant information efficiently as discussed in Section 5.2. Another speculation is that the Timeline view may make a document containing evidence for an alias more salient than the ER view. In the ER view, such a document appears as a single link between the two identities, which could be easily overlooked in a complex network. However, in the Timeline view, the two names would separately appear in different documents except for the one or two documents containing the evidence showing the connection between the two. This visualization would be more visible than a single link in the ER view as shown in Figure 4.

5.4.3 Screen Captures

We also analyzed the screen captures that were taken at the moment that participants submitted reports.

Interestingly, the layouts of entities in the ER view generated by Groups N and T are drastically different, as exemplified by Figures 5 and 6. It was clear that several participants (PN2, PN5 and PN6) in Group N tried to place entities in temporal orders on the ER view to reflect the identified storyline (PN1 and PN4 reorganized the ER view, but there was no temporal order visible). Figure 5 shows the ER view with dotted lines indicating these storylines. There is also an exception to this pattern: PN3 did not change the layout of the ER view at all for any spatial organization.

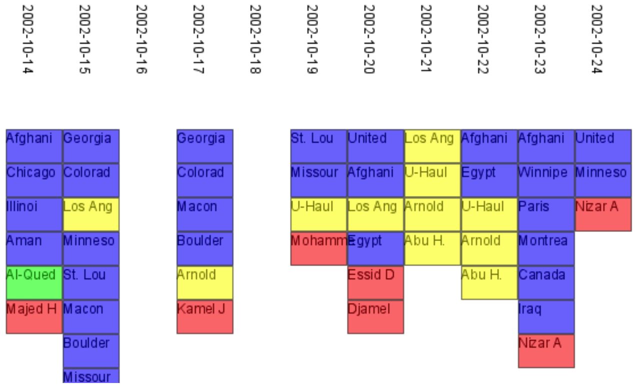

Fig. 4. A portion of the screenshot submitted with PT3-12. Note that Arnold and Abu H. appeared in two adjacent documents (2002-10-21 and 2002-10-22), which reveal fingerprint matches between two identities.

That Group N used the ER view for recovering temporal order is not unexpected. Group N did not have any external media (not even paper and pencil) to record or build storylines. Thus, the ER view is a natural medium for them to externalize their storylines while conducting investigative analysis. This might also explain why the initial numbers of entities on the screen were lower in Group N. We speculate that the participants probably did not want their storylines to be polluted with irrelevant entities.

In contrast with Group N, we did not find a particular layout pattern on the ER view for Group T. Instead, we found that entities on the ER view were more or less randomly spread or held their initial positions without much changes. Figure 6 exemplifies how PT2 and PT4 organized entities on the ER view.

The random entity layouts generated by Group T were unexpected. Because Group T participants also did not have any other external media for recording storylines, they likely used the Timeline instead of the ER view for externalizing analysis. This is interesting because the Timeline does not allow users to change the layout of entities on the screen, which we initially thought would make it unsuitable as a story mechanism. This is discussed in more detail next.

6 Discussion

Our emphasis with this work was on understanding the role of temporal visualization using the Timeline view in TimeInvestigator as an instantiation. As acknowledged earlier, our intention is not to provide statistical comparisons between the two groups. We also acknowledge that the addition of the Timeline view may improve the analysis by simply providing

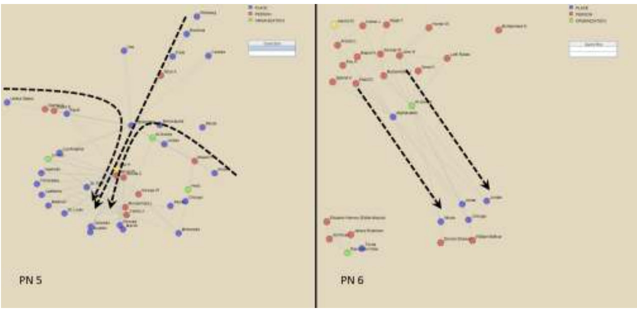

Fig. 5. The layouts in the Entity-Relationship view for PN5-4 (left) and PN6-7 (right).

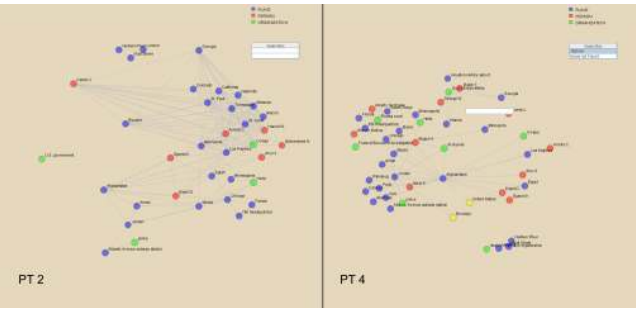

Fig. 6. The layouts of the Entity-Relationship view for PT2-16 (left) and PT4-5 (right).

another external visual representation. However, the results shown above clearly indicate that participants were indeed influenced by the presence and absence of this view. Therefore, we focused our attention on comparing behavior as opposed to comparing low-level performance between the two groups.

6.1 Benefits of the Timeline View

In summary, we noted three benefits of the Timeline view from our study: the Timeline view (1) is vital for uncovering important entity relations; (2) allows for filtering out unimportant entities; and (3) helps identify patterns that are invisible in the ER view.

First, the most obvious benefit of having access to the Timeline view is that it aids in uncovering important relationships. The ER view only shows binary relations, i.e., whether there was a relationship between certain entities. On the other hand, the Timeline view shows the development of relationships over time. This, in turn, seemed to provide Group T participants with (1) chronological and logical order of events; (2) the importance of events/entities; and (3) changes in relationships over time (based on feedback from PT1, PT4, and PT6). The absence of temporal visualization may be a roadblock for Group N because the chronological order of events were not visually available, so they needed to be remembered or recorded by some other method. Group N's heavier use of the Document List (average 8:43 in Group N vs. average 1:18 in Group T in Table 3), which listed the dates of reports in chronological order, indirectly shows that participants needed external cues to organize events in time. The interesting layouts of nodes and links, built by PN3 and PN4, in Figure 5 also showcase a tendency to want to make sense of stories and time.

Second, the Timeline view also provides additional cues to identify unimportant and irrelevant information. Two groups had unique patterns in the use of falsifications. Group N tended to be hesitant to falsify and thus discard entities, presumably due to the difficulty of overseeing long-term implications of such an action. Interview quotes from Group N clearly showed this aspect: for example, PN3 reported that he or she felt that he/she could not remove any entities because they may somehow be important. Because of the absence of temporal visualization, participants in Group N seemed to struggle in following entities of interest. On the other hand, Group T could follow entities through time and discard if they turned out not to be important. For example, when some entities appeared only in a certain time period, which was clearly visible in the Timeline view, these entities were easily disregarded (PT1 and PT6). This seemed to help Group T make falsifications with confidence.

Third, the Timeline view appears to be effective in highlighting a specific pattern, such as aliases, which could be easily obscured in the ER view. Identifying aliases is a particularly interesting activity in the context of investigative analysis because evidence showing aliases is often subtle but may drastically influence the analysis outcome. More specifically, it makes the aliased entity very suspicious and often helps to better understand its local neighborhood of entities. Although we failed to collect evidence showing that the Timeline view directly helped identify such aliases, we speculate that the Timeline view generally makes such subtle evidence more visible.

These benefits of the Timeline view were slightly different from our initial speculations derived from the literature (Section 2.1). We thought that time would be important because it can definitely rule out some hypotheses because they are impossible or unlikely due to causality (i.e., effects cannot come before causes). However, we found that causal relations cannot be easily visualized because these subtle causalities are often buried in documents. Entity co-occurrence in a single document is far from enough for causal relations distributed in time (and thus generally over several documents). Therefore, an analyst must often read several documents carefully to extract these casual relationships. We found that while temporal visualizations cannot directly provide causality information, they can certainly support the task of deriving causality. Entity frequencies, which are relatively simple to visualize, served as important cues for our participants to discern which nodes might be more important than other nodes.

6.2 The ER View vs. the Timeline View

Another interesting aspect of the Timeline view is that it not only provided additional benefits on its own but also influenced how the ER view was used. As shown repeatedly, Group N tended to start with a small number of entities and added more entities as the analysis progressed, while Group T did the opposite. In other words, Group N seemed to use the ER view in the additive fashion of "drawing on a canvas" to record the progress of investigation. Thus, they did not want to overload the ER view with irrelevant entities, which may distract them. We also noticed that, when too many entities were loaded into the ER view, participants tried to divide and allocate particular regional space in the ER view for specific use (e.g., Figure 5). This intelligent use of space for simplifying choice, perception, and computation in the real world has also been observed in general cognitive science studies, e.g., by Kirsh [32].

Group T, on the other hand, seemed to use the ER view in a subtractive fashion of "carving a sculpture." This is a radically different approach. Rather than adding important entities to the ER view, participants in Group T chose to eliminate unrelated entities from the view. The ER view did not seem to be used for story building as discussed in Section 5.4.3.

This difference is most likely caused by the fact that Group N participants used the ER view as a story building mechanism, whereas Group T participants did not. The unorganized layouts on the ER view generated by Group T participants are particularly intriguing. Assuming that Group T also built storylines for their investigative analysis, Group T participants appeared to use the Timeline view as a story building mechanism. However, it should be noted that the Timeline view does not provide much degree of freedom in organizing visual elements because the horizontal and vertical locations of entities are predetermined. The only additional interaction that participants could do is either highlighting or eliminating entities. Thus, the Timeline view appears to be more limited compared to other story building mechanisms, such as the Shoebox feature in Jigsaw [2], which provide various evidence marshaling capabilities.

But if this is true, how were Group T participants able to build stories using the Timeline view? One speculation is that people may not need sophisticated interface support to build a story (as evidenced by oral storytelling tradition). Perhaps the single key feature of the Timeline is that it presents events in chronological order, the precise order necessary to conduct investigative analysis. The temporal entity order and highlight features may be sufficient to help people construct stories out of various reports. Perhaps, Group T participants built stories in their minds using only the external media to help organize the story more efficiently. This speculation is in line with the notion of distributed cognition [33] and its application to visualization [34].

However, we acknowledge that the complexity of the main storyline in our study may not be complex enough, allowing the Timeline view to serve as a story building mechanism in only this particular case. For a more complex storyline (e.g., multiple branching and merging of storylines along the main plots), its lack of spatial interaction may cause it to be insufficient as a story building mechanism because it does not support the expressive power needed by the analyst. In addition, the role of the ER view seem to be still important even though it was not used as a story building mechanism for Group T participants. The time spent on the ER view for Group T is still substantial, and highlighting (i.e., brushing) entities appeared to help Group T participants connect the two views and see relational and temporal information simultaneously. The speculations made on top of our findings in this study should be tested in future research.

6.3 Externalizing Temporal Data

In general, the above benefits of the Timeline view all seem to stem from the fact that this view externalizes temporal relationships in a form more amenable to human comprehension than many other representations. This is an instantiation of the concept of *external cognition* [16], which has been quoted as one of the primary mechanisms of general visualization. Adapting Scaife and Rogers' terminology, we think that timeline representations aid the user in the following ways:

- **Computational offloading:** Timeline representations make the story of the data explicit in the world (i.e., on the computer screen), reducing the need for users to mentally formulate and store this information in their minds.
- **Re-representation:** Temporal order is key to uncovering causal information, which in turn is central to identifying an overarching plot in investigative analysis. Unlike other visual representations, timeline representations make the temporal order between events clearly visible.
- **Graphical constraining:** Mapping time onto a screen dimension provides an explicit graphical constraint on the temporal order of events that is not present in the ER view, where the nodes are placed according to different criteria. This constraining allows users to quickly rule out impossible hypotheses (e.g., effects before causes).

Naturally, the above mechanisms are true for all visual representations in general. However, this treatment makes some progress towards understanding the actual mechanics of why timeline visualizations are useful. Time is a fundamental aspect in our world and for visual analytics [1], so these findings will be useful in acquiring a better understanding of how visualization helps the user understand the data.

6.4 Design Guidelines

Based on our findings and discussions above, we provide below a set of design guidelines for temporal visualization when used in investigative analysis:

Supporting temporal analytic tasks: We found that temporal visualization could help discern which entities and relationships are important or not by presenting the following patterns: (1) entities appearing multiple times over the timeline; (2) several entities co-occurring multiple times over the timelines; (3) entities appearing before and after a certain time; (4) two entities appearing only once together but separately appearing multiple times over the timeline. When designing future temporal visualizations, the designer should confirm that these patterns are indeed visible.

Combining temporal and relational information: We also found that the ER and Timeline views were often used together by Group T. Thus, our recommendation is to provide both relational and temporal information, or, better yet, to create visualizations that combine relational and temporal information in the same view. In addition, interaction techniques such as brushing and linking, which combine data from temporal and relational views, should be used.

Supporting story building: Although the importance of supporting story building is well understood and accepted in the visual analytics community, our findings on the use of the Timeline view are intriguing and suggest a need to study this topic in much more detail. In our particular study, the single key feature of the temporal visualization seemed to be that it showed selected events in chronological order: this was enough to off-load the participants sufficiently so that they felt no need to externalize the storyline in the ER view. While this may be an effect of the relatively small dataset we used in the study, this in turn suggests that visual representations do have a significant impact on the cognitive effort of the user. A story building mechanism need not be overly complex and full-featured if the temporal visualization provides sufficient information for the user to be able to reconstruct the storyline in their head.

6.5 Evaluating Visual Analytics: Lessons Learned

We did not record video due to its high cost in codifying and analysis, which went against our lightweight evaluation approach. We also found in pilot testing that video was of limited use since our study was based on single-user analysis restricted only to the TIMEINVESTIGATOR tool and with no external aids. In other words, the optimal use of video in our study would be to record the contents of the screen. In general, we feel that this gives rise to a recommendation on how to capture user behavior during investigative analysis: *select user behavior capture mechanisms by carefully considering analysis costs versus potential gains.*

Based on this reasoning, our TIMEINVESTIGATOR system was heavily instrumented to capture large amounts of interaction data and screenshots during each experimental session. However, this left us with tens of thousands of lines of log events. Our solution was to turn the analysis of visual analytics evaluation into a visual analytics problem of its own. In this work, we have explored ways of applying visualization techniques to both analyze our data as well as to expose it to our audience. We are surprised by the scarcity of such approaches in the literature (notable exceptions include Isenberg et al. [24], Robinson [25], and Tang et al. [35]), and thus we feel our recommendation on this is both novel and useful: *use visualization to analyze complex evaluation results.*

Furthermore, inspired by the insight-based evaluation employed by Saraiya et al. [28], we introduced a replacement of the think-aloud protocol in the form of our impromptu insight reports and screenshots that we use for collecting the participants' thought process throughout analytical sessions. We think that this is a useful technique for visual analytics evaluation because of its smooth integration into the analytical process (noting down intermediate thoughts and ideas is not uncommon when studying complex problems), so we recommend to *allow participants to record insights and results throughout a session, and not just at the end.*

6.6 Limitations

Below we discuss the most important limitations of our study and their potential impact on our findings.

6.6.1 Custom Tool

We developed a custom tool—TIMEINVESTIGATOR—for this evaluation rather than using an existing, established tool like Jigsaw [2], CzSaw [21], or the Analyst's Notebook [30]. This means that our results may be more difficult to generalize to other tools for investigative analysis. Furthermore, it could also be argued that our implementations for different views are suboptimal compared to those of established tools.

However, we made this decision because of the need to be able to fully instrument the tool with our testing environment, and to constrain the tool to have a minimal subset of operations. Existing tools often have several ways of accomplishing the same task, whereas our approach allowed us to make the different views of the tool as orthogonal as possible.

6.6.2 Participant Expertise

Just like Kang et al. [4], our evaluation included only novice analysts from the student population at our university. As a result, our results may have been different if the study participants had been professional intelligence analysts. However, as noted by Kang et al., intelligence analysts are a small and highly inaccessible population, so including them in exploratory studies of this nature is difficult, at least in a university setting. Given the lightweight evaluation methodology used in this paper, we wanted to investigate the depth and breadth of findings possible even with non-professional analysts as study participants.

What impact this choice had on our results is difficult to establish. The fact that we constrained participants to using an unfamiliar tool would eliminate effects of practice and presumably uncover emerging strategies that would be same across both populations. All participants also received 20 minutes of training using a small dataset before each session.

Furthermore, our participants were all engineering students, whereas many analysts may come from a broader social or political science background. Again, we are unable to predict what impact this difference would have on the results: one hypothesis may be that social and political science majors are more accustomed to reading and summarizing large amounts of text, whereas the visual representations used in our tool would benefit engineering students better. Additional evaluation is needed to answer such questions.

6.6.3 Solution Grading

Kang et al. [4] used external reviewers (graduate students in the larger research group) to grade the solutions derived by each participant to avoid bias. We used two of the authors of the paper to grade and code the insight reports independently.

To maintain objectivity in spite of this fact, we established a strict coding rubric, performed the two coding sessions independently, and then merged and discussed the results into a single coding metric. This approach is common practice in much qualitative evaluation in social science. Therefore, we do not think it affected the results significantly.

7 Conclusion and Future Work

We have presented a qualitative evaluation of temporal visualization for investigative analysis. Our evaluation clearly showed that having a temporal visualization (the Timeline view) provides participants with additional aids to find important clues and falsify irrelevant information, so that they more easily can find the correct solution. These positive outcomes are a result of the temporal view not only serving as a passive view showing temporal information, but also serving as an external memory aid for viewing complex event sequences and for building storylines.

It is clear that visual analytics evaluation is still a wide-open research topic. Our future work will focus on studying the analytical process in more detail. In particular, we think that the low-cost evaluation approach used in this paper will be helpful in extending our studies of investigative analysis to other settings, scenarios, and tasks beyond the intelligence domain.

Acknowledgment

Thanks to Youn-ah Kang and John Stasko for sharing the dataset that was used in their paper [4]. We are indebted to the anonymous reviewers for their extensive feedback on early versions of this work.

References

[1] W. Aigner, S. Miksch, W. Müller, H. Schumann, and C. Tominski, "Visual methods for analyzing time-oriented data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 1, pp. 47–60, Jan./Feb. 2008.

[2] J. T. Stasko, C. Görg, and Z. Liu, "Jigsaw: supporting investigative analysis through interactive visualization," *Information Visualization*, vol. 7, no. 2, pp. 118–132, 2008.

[3] R. Eccles, T. Kapler, R. Harper, and W. Wright, "Stories in GeoTime," *Information Visualization*, vol. 7, no. 1, pp. 3–17, 2008.

[4] Y. Kang, C. Görg, and J. Stasko, "Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study," in *Proceedings of the IEEE Symposium on Visual Analytics Science & Technology*, 2009, pp. 139–146.

[5] E. A. Bier, S. K. Card, and J. W. Bodnar, "Entity-based collaboration tools for intelligence analysis," in *Proceedings of the IEEE Symposium on Visual Analytics Science & Technology*, 2008, pp. 99–106.

[6] D. H. Jeong, W. Dou, F. Stukes, W. Ribarsky, H. R. Lipford, and R. Chang, "Evaluating the relationship between user interaction and financial visual analysis," in *Proceedings of the IEEE Symposium on Visual Analytics Science & Technology*, 2008, pp. 83–90.

[7] W. Wright, D. Schroh, P. Proulx, A. Skaburskis, and B. Cort, "The sandbox for analysis: Concepts and methods," in *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 2006, pp. 801–810.

[8] S. Carpendale, "Evaluating information visualizations," in *Information Visualization: Human-Centered Issues and Perspectives*, ser. Lecture Notes in Computer Science, 2008, pp. 19–45.

[9] C. Plaisant, "The challenge of information visualization evaluation," in *Proceedings of the ACM Conference on Advanced Visual Interfaces*, 2004, pp. 109–116.

[10] C. Plaisant, G. G. Grinstein, J. Scholtz, M. Whiting, T. O'Connell, S. J. Laskowski, L. Chien, A. Tat, W. Wright, C. Görg, Z. Liu, N. Parekh, K. Singhal, and J. T. Stasko, "Evaluating visual analytics at the 2007 VAST symposium contest," *IEEE Computer Graphics and Applications*, vol. 28, no. 2, pp. 12–21, 2008.

[11] C. North, "Toward measuring visualization insight," *IEEE Computer Graphics and Applications*, vol. 26, no. 3, pp. 6–9, 2006.

[12] G. Klein, B. M. Moon, and R. R. Hoffman, "Making sense of sensemaking 1: Alternative perspectives," *IEEE Intelligent Systems*, vol. 21, no. 4, pp. 70–73, 2006.

[13] R. J. Heuer, *Psychology of intelligence analysis*. United States Government Printing Office, 1999.

[14] G. Chin, Jr., O. A. Kuchar, and K. E. Wolf, "Exploring the analytical processes of intelligence analysts," in *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 2009, pp. 11–20.

[15] D. A. Norman, "Cognition in the head and in the world," *Cognitive Science*, vol. 17, no. 1, pp. 1–6, Jan.-Mar. 1993.

[16] M. Scaife and Y. Rogers, "External cognition: how do graphical representations work?" *International Journal of Human-Computer Studies*, vol. 45, no. 2, pp. 185–213, 1996.

[17] D. Peuquet, "It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems," *Annals of the Association of American Geographers*, vol. 84, no. 3, pp. 441–461, Sep. 1994.

[18] C. Plaisant, R. Mushlin, A. Snyder, J. Li, D. Heller, and B. Shneiderman, "LifeLines: Using visualization to enhance navigation and analysis of patient records," in *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 1996, pp. 221–227.

[19] T. D. Wang, C. Plaisant, A. J. Quinn, R. Stanchak, S. Murphy, and B. Shneiderman, "Aligning temporal data by sentinel events: discovering patterns in electronic health records," in *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 2008, pp. 457–466.
[20] K. Wongsuphasawat and B. Shneiderman, "Finding comparable temporal categorical records: A similarity measure with an interactive visualization," in *Proceedings of the IEEE Symposium on Visual Analytics Science & Technology*, 2009, pp. 27–34.
[21] N. Kadivar, V. Chen, D. Dunsmuir, E. Lee, C. Qian, J. Dill, C. Shaw, and R. Woodbury, "Capturing and supporting the analysis process," in *Proceedings of the IEEE Symposium on Visual Analytics Science & Technology*, 2009, pp. 131–138.
[22] J. Scholtz, "Beyond usability: Evaluation aspects of visual analytic environments," in *Proceedings of the IEEE Symposium on Visual Analytics Science & Technology*, 2006, pp. 145–150.
[23] M. Tory and T. Möller, "Evaluating visualizations: Do expert reviews work?" *IEEE Computer Graphics and Applications*, vol. 25, no. 5, pp. 8–11, 2005.
[24] P. Isenberg, A. Tang, and M. S. T. Carpendale, "An exploratory study of visual information analysis," in *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 2008, pp. 1217–1226.
[25] A. C. Robinson, "Collaborative synthesis of visual analytic results," in *Proceedings of the IEEE Symposium on Visual Analytics Science & Technology*, 2008, pp. 67–74.
[26] D. Gotz and Z. Wen, "Behavior-driven visualization recommendation," in *Proceedings of the International Conference on Intelligent User Interfaces*, 2009, pp. 315–324.
[27] K. S. Park, A. Kapoor, and J. Leigh, "Lessons learned from employing multiple perspectives in a collaborative virtual environment for visualizing scientific data," in *Proceedings of the ACM Conference on Collaborative Virtual Environments*, 2000, pp. 73–82.
[28] P. Saraiya, C. North, V. Lam, and K. A. Duca, "An insight-based longitudinal study of visual analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 6, pp. 1511–1522, 2006.
[29] J. Scholtz, E. Morse, and T. Hewett, "In depth observational studies of professional intelligence analysts," in *Human Performance, Situation Awareness and Automation Technology Conference*, 2004.
[30] i2 Incorporated, "Analyst's Notebook," http://www.i2inc.com/, 2010, accessed March 2011.
[31] J. S. Yi, Y. ah Kang, J. T. Stasko, and J. A. Jacko, "Toward a deeper understanding of the role of interaction in information visualization," *IEEE Transactions of Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1224–1231, 2007.
[32] D. Kirsh, "The intelligent use of space," *Artificial Intelligence*, vol. 73, pp. 31–68, 1995.
[33] E. Hutchins, *Cognition in the Wild*. Cambridge, MA: MIT Press, 1995.
[34] Z. Liu, N. J. Nersessian, and J. T. Stasko, "Distributed cognition as a theoretical framework for information visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1173–1180, 2008.
[35] A. Tang, M. Pahud, S. Carpendale, and B. Buxton, "VisTACO: Visualizing tabletop collaboration," in *Proceedings of the ACM Conference on Interactive Tables and Surfaces*, 2011, pp. 29–38.

Bum chul Kwon is a Ph.D. candidate in the School of Industrial Engineering at Purdue University in West Lafayette, IN, USA. His research interests include information visualization, visual analytics and their applications in consumer decision making and the healthcare industry. He received his Master of Industrial Engineering in 2010 from Purdue University in West Lafayette, IN, USA. He is a student member of the IEEE.

Waqas Javed is a Ph.D. candidate in the School of Electrical and Computer Engineering at Purdue University in West Lafayette, IN, USA. His research interests include information visualization, visual analytics, and human-computer interaction. He received his Bachelor of Science in Electrical Engineering in 2007 from the University of Engineering and Technology, Lahore in Pakistan. He is a student member of the IEEE.

Sohaib Ghani is a Ph.D. candidate in the School of Electrical and Computer Engineering at Purdue University in West Lafayette, IN, USA. His research interests include information visualization, visual analytics, and human-computer interaction. He received his Bachelor of Science in Electrical Engineering in 2007 from the University of Engineering and Technology, Lahore in Pakistan. He is a student member of the IEEE.

Niklas Elmqvist is an assistant professor in the School of Electrical and Computer Engineering at Purdue University in West Lafayette, IN, USA. He was previously a postdoctoral researcher at INRIA in Paris, France. He received his Ph.D. in 2006 from Chalmers University of Technology in Göteborg, Sweden. He is a member of the IEEE and the IEEE Computer Society.

Ji Soo Yi is an assistant professor in the School of Industrial Engineering at Purdue University in West Lafayette, IN, USA. He received his Ph.D. in 2008 from Georgia Institute of Technology in Atlanta, GA, USA. He is a member of the IEEE.

David Ebert is the Silicon Valley Professor in the School of Electrical and Computer Engineering at Purdue University in West Lafayette, IN, USA. He received his Ph.D. in 1999 from the Ohio State University in Columbus, OH, USA. He is a fellow of the IEEE.