Supporting Communication and Coordination in Collaborative Sensemaking

Narges Mahyar and Melanie Tory

Abstract—When people work together to analyze a data set, they need to organize their findings, hypotheses, and evidence, share that information with their collaborators, and coordinate activities amongst team members. Sharing externalizations (recorded information such as notes) could increase awareness and assist with team communication and coordination. However, we currently know little about how to provide tool support for this sort of sharing. We explore how linked common work (LCW) can be employed within a 'collaborative thinking space', to facilitate synchronous collaborative sensemaking activities in Visual Analytics (VA). Collaborative thinking spaces provide an environment for analysts to record, organize, share and connect externalizations. Our tool, CLIP, extends earlier thinking spaces by integrating LCW features that reveal relationships between collaborators' findings. We conducted a user study comparing CLIP to a baseline version without LCW. Results demonstrated that LCW significantly improved analytic outcomes at a collaborative intelligence task. Groups using CLIP were also able to more effectively coordinate their work, and held more discussion of their findings and hypotheses. LCW enabled them to maintain awareness of each other's activities and findings and link those findings to their own work, preventing disruptive oral awareness notifications.

Index Terms—Sensemaking; Collaboration; Externalization; Linked common work; Collaborative thinking space.

1 Introduction

Supporting collaborative sensemaking has been identified as an important challenge in collaborative visualization [20]. Sensemaking in collaborative VA is a very time consuming and demanding process, requiring the analysts to iteratively exchange and discuss results to form and evaluate hypotheses, derive conclusions, and publish findings. Team members also need to maintain *awareness* of each other's work, including both activities that people are working on and results and evidence that they have found. Tools that provide *externalization* support (i.e., ability to record insights, questions, and findings, e.g., as text notes) can help teams to organize and share their results [6, 18, 22, 41], and those that provide awareness channels should enhance collaboration, communication and coordination [12]. However, to date, we have a very limited understanding of how to provide externalization and awareness support for collocated collaborative teams. How should such tool support look and behave within VA tools?

We investigate the use of *Linked Common Work* (LCW) to facilitate synchronous collaborative sensemaking. With LCW, common work elements such as similar findings are automatically discovered, linked, and visually shared among the group. We built this technique within a 'collaborative thinking space' that enables analysts to record, organize and schematize their externalizations. Linked common work reveals similarities in people's externalizations, enabling analysts to acquire awareness of each other's findings, hypotheses, and evidence. Moreover, each individual analyst can review and merge others' work from within his/her workspace. Our results demonstrate that applying LCW to externalizations, and providing the ability to integrate collaborators' findings together within one view, noticeably improve team awareness, coordination, communication, and analytic outcomes.

Our work focuses on supporting teams of investigative analysts, for example in the domain of intelligence analysis. Intelligence analysts need to sift through large document collections, determine which pieces of data are relevant, and gradually build up an explanation supported by evidence. Field studies have revealed that professional analysts need to share sources and data, view each other's work, and combine findings together in order to build common ground, resolve conflicts, and validate each other's findings and hypotheses [8, 25].

The *sensemaking* process of intelligence analysts has been studied in some depth, and has been described as involving two iterative loops: the information foraging loop and the sensemaking loop [34]. The information foraging loop involves searching for relevant data and reading, filtering, and extracting information, whereas the sensemaking loop involves iteratively developing a mental model, forming and evaluating hypotheses, and publishing the results. We focus primarily on supporting later stages of the sensemaking process (i.e., the sensemaking loop), when teams are more likely to work together in a synchronous, collocated fashion [25]. This synthesis phase is reported to be the most difficult and time-consuming phase of analysis [25].

We are exploring the design of visual thinking spaces that support the sensemaking loop in collaborative VA. A collaborative thinking space should enable analysts to record and organize findings, evidence, and hypotheses; moreover, it should facilitate the process of sharing findings amongst collaborators, to minimize redundant work and help investigators identify relationships and build a shared understanding. In this paper, we examine the value of employing LCW to relate and integrate team members' visual thinking spaces. The notion of LCW closely resembles *collaborative brushing and linking* [21] in which certain actions of each investigator are visible to collaborators through their own views. However, collaborative brushing and linking was only applied to search queries and retrieved documents and did not cover externalizations. It also focused on supporting only information foraging activities. In contrast, our work facilitates later stages of the collaborative sensemaking process (i.e., the sensemaking loop), by applying the linking concept to people's externalizations (i.e., recorded findings and notes). We anticipate that enabling analysts to see how their findings relate to each other should make it easier to maintain awareness of each others' work, build common ground, and solve analytic problems. We address the following research questions (RQs):

- RQ1: Does linking collaborators' externalizations lead to better analytic outcomes?
- RQ2: Does linking collaborators' externalizations improve communication?
- RQ3: Does linking collaborators' externalizations help collaborators to coordinate their work more effectively?
- RQ4: Does linking collaborators' externalizations increase collaborators' awareness of each others' findings and activities?

To answer these questions, we designed and implemented CLIP, a visual thinking space to support collaborative sensemaking. CLIP allows analysts to record their findings in the form of a node-link graph and

- *Narges Mahyar and Melanie Tory are with the University of Victoria. E-mail: {nmahyar, mtory}@uvic.ca*

Manuscript received 31 March 2013; accepted 1 August 2013; posted online 13 October 2013; mailed on 4 October 2013.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

timeline, add evidence to facilitate evidence marshaling, and add free form text to record hypotheses, questions, to-do-lists, etc. Most importantly, CLIP incorporates LCW to relate and integrate the findings of different collaborators. We assessed the value of LCW by comparing CLIP to a baseline tool (BT) without the LCW features. Results of our user study demonstrated that LCW led to more effective group coordination and communication as well as better analytic outcomes.

2 Related Work

The design of CLIP draws upon prior research on sensemaking and collaboration support. For individual work, many tools have been developed to support both phases of sensemaking (e.g., [4,19,42]). However, in the context of collaborative sensemaking much less has been done. Most of the existing tools either focus on the information foraging loop [21] or asynchronous collaboration [7,41].

In the remainder of this section, we summarize the existing guidelines on how to support collaborative sensemaking activities, and how those relate to CLIP's design and our experimental study. To gather these guidelines, we reviewed relevant field and observational studies to extract features that focused on the sensemaking loop.

2.1 Externalization, Schematizing, and History

Mahyar et al. [27] demonstrated the critical importance of externalization during collaborative analysis, and the lack of support for this process in many current visualization tools. Correspondingly, Kang and Stasko [25] suggested that supporting history of previous discoveries and sanity checking could save time during report-writing. Their field observations showed that analysts spent substantial time returning to original sources to find the supporting references and rationale behind their statements. Vogt et al. [38] and Pirolli and Card [34] similarly pointed out the need to record findings, hypotheses and evidence. Several studies in the intelligence analysis domain signified the importance of schematizing results [8,22,24]; in other words, organizing results and other externalizations into a structured format. Various structured formats can be useful, including timelines, spreadsheets, lists, and networks [8,22]. For instance Zhang [43] discussed the nature of external representations in cognition and mentioned diagrams, graphs, and pictures as a few typical types of external representations. For meetings or tasks that require flexibility, such as brainstorming and collaborative design, freeform graphical input could be a better option to support flexibility [23,32]. Other structures include casual loop diagrams, mind maps, diagrams, graphs, and pictures [23, 31, 32]. We expect schematizing to be even more critical for collaborative work, since the structure may also help with communication. CLIP therefore includes node-link graph and timeline schemas for representing findings. We anticipate that integrating collaborators' schematic views will help them to build common ground and relate findings.

2.2 Communication and Coordination

Various studies have found that the closeness of groups' collaboration styles directly affected outcomes [5,22,25,38]. Isenberg et al. [22] and Vogt et al. [38] found that teams who collaborated more closely were more successful, and Bradel et al. [5] reported that members in more successful teams understood each other's work better. Groups should not always work in a closely coupled fashion, however. Kang and Stasko [25] found that in a long term project, collaboration was loose during information collection but tight when synthesizing findings and writing a report. A good collaborative system, then, should encourage groups towards closer collaboration styles when there are relevant findings to be connected, but allow loose collaboration at other times.

One way to encourage closer collaboration is through awareness mechanisms that provide information to each investigator about their collaborators' activities and findings. Paul and Reddy [33] suggested providing support for both action awareness and activity awareness, showing actions that led to a particular activity. Various techniques have been developed to help collaborators maintain awareness. For instance, Hugin [26] provides awareness support in a mixed-presence setting through a layer-based GUI design. CoSpaces [28] places each person's information in a separate view. Users must then compare and reconcile different views, a potentially cumbersome process. People may also miss relevant changes since they are often hidden from view. Nonetheless, Mahyar et al. [28] reported that separate views were useful for exploring other people's work in a non-disruptive way.

In contrast, integrating everyone's information into one view could cause disruption to individual work as the view constantly updates. Brennan et al. [6] implemented a visualization of externalizations and explored ways to merge collaborators' content. Similar to CLIP, Brennan et al.'s approach provides common ground to support collaboration. While CLIP does not address confidence ratings as explicitly as their system (though they could be added to notes), the simpler visual design of CLIP should make it easier to communicate: consistency in visual encoding should make it easy to understand other users' perspectives and LCW should make it easy to find commonalties. Similarly, CoMotion [9] enabled analysts to share data views and notes. However, neither project evaluated whether the shared view was helpful in practice. Another related system is Cambiera [21], which provided awareness cues about related searches conducted by a collaborator during information foraging. Isenberg et al. [22] reported that these cues, which they termed *collaborative brushing and linking*, encouraged closer collaboration. We emphasize that Cambiera did not consider how linking could be applied to externalizations. Our work extends the linking idea to externalizations and support for the sensemaking loop. Note that we use the term 'linked common work' instead of 'collaborative brushing and linking' to generalize the linking idea to cases where there is no active 'brushing' action to initiate the link.

Which is more important, having everyone's work visible in a single shared space, or reducing disruption to individual work by keeping notifications and visual changes to a minimum? Our study addresses this question by comparing CLIP, which allows everything to be integrated within one view, to a baseline tool with a separate view approach.

It is also important to support task coordination. Mark and Kobsa [29] identify coordination support as a challenge for designers and emphasize the need to choose appropriate visual representations to help teams during VA tasks.

2.3 Shared and Individual Workspaces

The importance of providing both individual and group workspace is well known [15,21,22]. For example, Wallace et al. [39] demonstrated that groups had better outcomes when they were able to share their results together. The need for shared and individual workspaces applies equally well to externalization spaces. Even though collocated teams often assign one note taker [27, 36, 38], individuals still occasionally need to record private notes [27]. This suggests that it is important to provide both shared and individual notetaking spaces. Mahyar et al. [27] revealed that taking notes on paper reduced the note taker's awareness of other activities, suggesting that notetaking tools should be integrated with the investigative application. However, a separated digital view can suffer from the same problem. For instance, in Bradel et al.'s [5] study, only one user could take notes in the shared space; others had to work on separate views and this reduced awareness. Kang and Stasko [25] recommended "*promoting individual workspaces as well as...the ability to share sources and data, view and comment on others' work, and merge individual work together.*" Brennan et al. [6] similarly suggested providing individual perspectives on a shared space. McGrath et al. [30] introduced Branch-Explore-Merge, a structure for transitioning between individual and group workspaces. We note that their concept was applied to direct views of data, not within a collaborative thinking space. With CLIP, we address the need for individual workspace plus awareness of others' work by providing each user with flexible control over how much of the collaborators' information is shown in their view.

2.4 Studies of Collaborative Thinking Spaces

Most similar to our work are studies of collaborative visual analytics work involving thinking spaces, particularly studies that explore how to integrate or provide access to different users' views. What level of integration is appropriate for different kinds of shared information?

Chen et al. [7] built a tool that enabled asynchronous collaborators to record and share insights, and demonstrated that people could learn from others' past insights. Similarly, Willett et al. [41] demonstrated that asynchronous collaborators benefited from the ability to classify their text comments using tags and from the ability to link comments and views; however, in their system, users had to create links manually. Neither of these studies examined the value of automatically linking common work, nor did they look at synchronous collaboration.

For synchronous collaboration, most evidence suggests that highly integrated views should be a good approach. Balakrishnan et al. found that a shared view of a network diagram supported better performance at a collaborative intelligence task than separated views [2]. We note, however, that their network diagram was a simple social network rather than a dynamic thinking space for recording and organizing externalizations. We are motivated by their later study [3], where they suggested that it might help to additionally visualize partners' activities and externalizations. Bradel et al. [5] used Jigsaw's tablet view to allow analysts to record their findings in the form of notes, graphs, and timelines. They found that Jigsaw's tablet view was inadequate for collocated collaboration, because participants wanted to use it as a shared note space, but it accepted only one person's input at a time. Their study did not examine separate linked views of the thinking space, but they suggested that it may be a good idea.

3 Linked Common Work (LCW)

The LCW technique employed in CLIP reveals similarities between collaborators' findings by discovering, linking, and visually representing the common work. This approach is based on research in social interaction such as Clark and Brennan's [10], that showed the importance of a shared understanding for effective collaboration. Subtle visual cues enable analysts to gain awareness of each other's findings, hypotheses, and evidence with minimal disruption [16, 21]. This is what we called *partial merging*. Then, if an analyst wants to more closely monitor others' work the *full merging* option is available to integrate others' findings directly into his/her workspace. Partial and full merging are described in detail below. When views are fully merged, the layout computation only updates positions of common nodes (those in common between the user's graph and their collaborator's). Common nodes are overlaid and their edge shapes are recomputed as necessary. All remaining nodes (i.e. uncommon nodes) are placed where they originally were, and the user can move them if necessary. We avoid making automatic changes in the layout of the local graph in order to preserve the user's mental map of nodes and relationships.

As a proof of concept, we used Jigsaw to extract evidence items and lists of entities (People, Locations, Organizations, Chemicals, Events, etc.) from the document corpus to ensure perfect matching between participants' externalizations. There are many ways to improve both the visual representation and the merging algorithm used for LCW. For example, there are algorithms to merge entities that are named differently but semantically related (e.g., lexicon chains to find synonyms and related words). However, the main goal at this stage of our work was to demonstrate value of LCW. In future work we discuss how to improve and extend the LCW technique.

4 System Design

CLIP, our Java-based prototype, was designed to facilitate collaborative analysis by providing a space for teams to record and share hypotheses, conjectures, and evidence. CLIP's design (Figure 1) takes into account the design guidelines outlined earlier. For example, CLIP enables analysts to record externalizations in structured formats (graph, timeline, and notes), addressing the need for externalization and schematizing. CLIP also enables analysts to work individually but merge their findings, supporting the principle of shared and individual workspaces. CLIP facilitates validation of results by enabling people to add evidence to each finding and by visually representing evidence both around the nodes as well as in the evidence cloud. Most importantly, we implemented LCW to enhance awareness by revealing relationships between collaborators' externalizations. Collectively, these features help analysts to see who did what and follow the trail of how other analysts came to a particular conclusion. To facilitate same-time collaboration, CLIP supports awareness through user-controlled sharing of work amongst team members, with colour coding to indicate who did what. The specific scenario for which we designed CLIP is to support team-based analysis of a document collection for solving a mystery task. However, we envision that the design ideas in CLIP could be applied to other analysis scenarios. Rather than directly visualizing the document collection, CLIP visualizes and links the team members' externalizations relevant to their analysis. Interesting entities (e.g., people, locations, or events described in the documents) can be externalized as nodes, and relationships between events as links. Each recorded entity can be optionally linked to free form text notes and to a timeline. In our study, participants took the role of intelligence analysts solving a mystery task from the VAST 2006 challenge. In the following sections, we begin with a scenario related to our study task, illustrating CLIP's use. This is followed by a description of CLIP's features related to externalization and awareness support.

4.1 Scenario

Laura, Alex and Mary are reviewing a set of documents to solve a mystery task. Laura has been focusing on a suspicious event at the 'Silo'. From the article she finds that 'George Prado' is up to some illegal activity, maybe running a meth lab. She suspects 'George Prado' is the key person, so she records him as the main suspect and creates a node with his name. In addition, she creates a note containing her hypothesis that "*George is probably running a meth lab*". This could also bring George Prado to her collaborators' attention. Then she starts gathering evidence to support or refute her hypothesis about him. Later, she finds an interesting article about the 'Silo' and she creates a node for 'Silo'. As soon as the node is created, the visual glyphs on the node inform her that that Mary has also been investigating the 'Silo' event and has recorded information about it. From there, Laura gets interested to see what else Mary has found so far. She opts to merge Mary's entire graph into her own view. She discovers that Mary has found interesting relationships involving the 'Silo' event. Tracing Mary's work, she finds out that both Alex and Mary have collected substantial evidence of the terrorist group's links to George Prado.

Laura decides to also view Alex's full work by merging it with her own. Looking at notes made by Alex, she realizes that 'George Prado' is up to something bigger than running a meth lab. She opens a discussion about 'George Prado' running the 'FFE farm' and from there, they connect the facts and validate their hypotheses. At the same time, Mary records a new finding that shows George's brother is a security guard at the 'Annex', a chemicals warehouse. As soon as Laura and Alex see this new finding in their own views, they start talking to Mary and sharing all their findings so far; in this discussion they realize that the Prado family are probably supplying chemicals to the terrorists.

4.2 Externalization

The importance and benefits of recording externalizations (notes, findings, etc.) have been emphasized by many researchers (e.g., [14, 22, 27]). CLIP provides space for recording and visually representing important entities and relationships, the time order of events, and free form text notes. Each recorded entity takes the form of a node in a node-link graph (Figure1A). Each item is indicated by a unique colour corresponding to the owner. Initially, each user logs into their instance of CLIP by selecting a username and a colour. Evidence can be attached to each node. According to previous research [25], returning to original sources and checking the references can be very tedious and time consuming. Attaching evidence to nodes in CLIP helps analysts easily return to original sources to verify accuracy of reported findings. Each node represents an entity and has six main attributes: text, type, note, image, date, and evidence list. Only the text and evidence list (at least one evidence document) are required for node creation; other node attributes are optional and can be updated any time.

To add a new node to the graph, a user enters information in a popup dialog (Figure 2). Each list in the dialog is pre-populated by the values that were extracted from the document corpus. To enter values

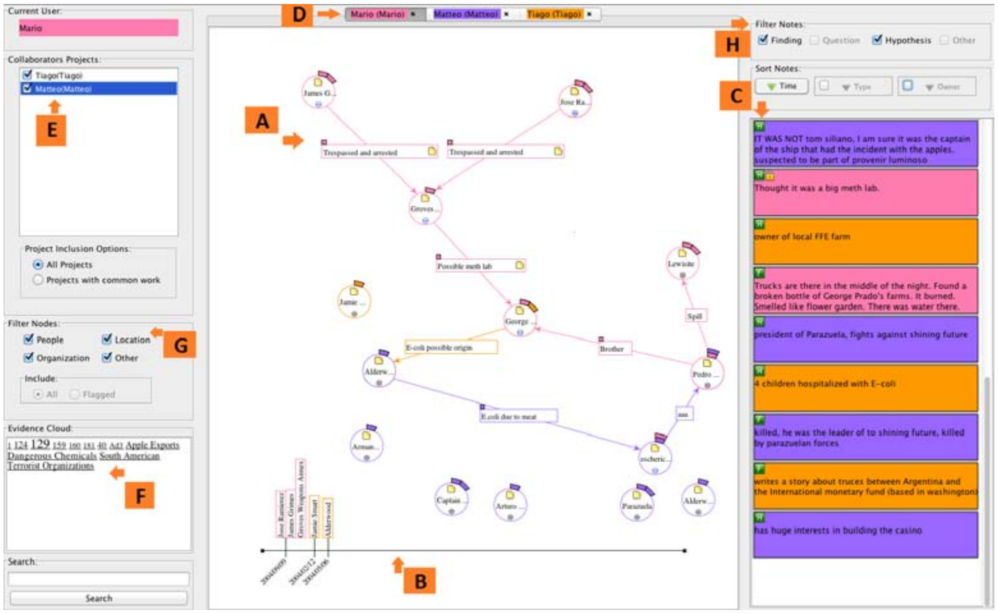

Fig. 1: Screenshot of CLIP. A) Graph pane, to create a network diagram of people, locations, and events, B) Timeline, to see the timeline of events, C) List of notes, easy review of all notes in one location, D) Tabs, to see collaborators' views, E) Merge option, to choose a collaborator's work to be merged with your own, F) Evidence cloud, to see the list of evidence and their frequencies, G, H) Filtering and sorting options.

that do not belong to any of these categories, there is a text box labeled "Other". To assist users with data entry and improve the discovery of common work, we implemented an auto-fill feature for this text box that provides suggestions (name values such as chemicals, events, etc.). A date stamp (null by default) can be attached to the node (Figure 2B). Any node with a non-null date stamp will automatically appear on the timeline. A date stamp can be attached to any node type (i.e. Person, Location, Organization, or Other). The rationale behind this design was to enable users to associate entities of all types with time if needed. Interestingly, in our study we observed that many participants associated people and location names with dates, perhaps as a shorthand way to represent an event. To attach evidence to a node, a user can select documents from the evidence list (Figure 2C). The current design permits attaching up to twelve documents to a node as evidence. The dialog also contains a text area for recording a note (Figure 2D). Note content type (Hypothesis/Question/Finding/Other) and scope (Public/ Private) can be set as well. To assist a user to easily identify private versus public notes, a lock icon appears on the private notes. Finally, an image can be attached to a new node (Figure 2E).

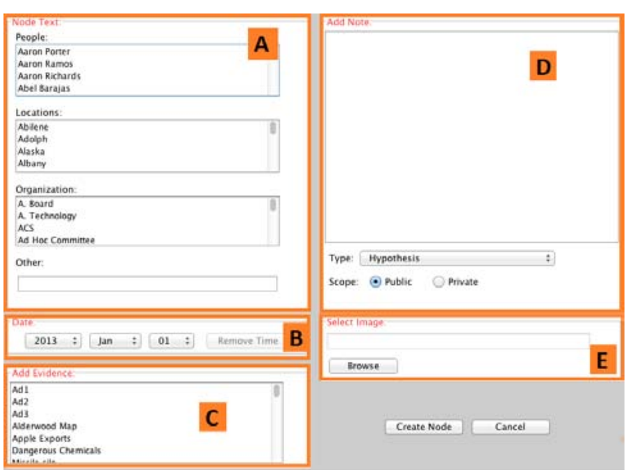

Fig. 2: Dialog for creating a new node. A) Select node's text, B) Add date, C) Add evidence, D) Free form note, E) Add image.

Each graph node has a toggle button (+ or -) that controls the visibility of the node's children (if any). Collapsing a node collapses all the branches that stem from the node. This feature makes it possible to collapse/ expand the graph from any given node, improving scalability. Users can filter nodes based on type (Figure 1G). This enables them to hide parts of graph if required (improving the scalability) or quickly locate nodes of specific types.

Links represent relationships between captured entities. Each link has three main attributes: text, note and evidence list. These attributes mirror those of the nodes. Unlike a node's evidence list, a link's evidence list can be null. Figure 3 shows a node's design. Node text is placed in the middle of the circle. A yellow note icon above the text indicates that a note is attached to the node (Figure 3A). If the evidence list is not empty, there are segments drawn on the outside of the node circle, one for each evidence document attached to the node. These visual cues provide a quick overview of each node. Segments in a collaborator's colour (Figure 3B and 3C) represent evidence found by a collaborator, one of the ways we reveal LCW. By default, all notes are

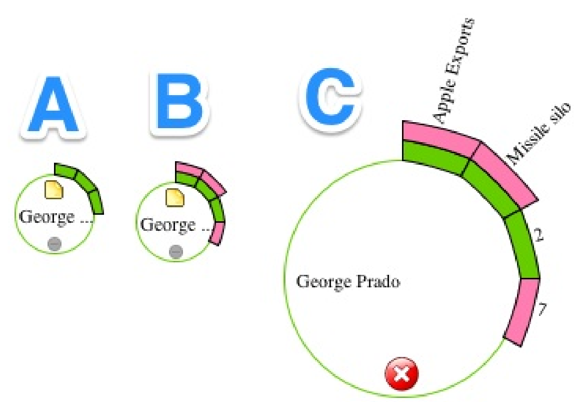

Fig. 3: Node details. A) View of a node before a collaborator creates the same node and B) after a collaborator creates the same node. Visual cues (colour coded segments) indicate it is a common node and reveal the common and different evidence. C) Enlarged node to see the details. Enlarging a node also highlights related items in other views, such as notes in the note list, timeline items, and evidence.

placed in the note panel (Figure1C). Each note can be closed by the user later. When closed, the yellow note icon (inside the related graph node) changes to red as a visual indication. Users can sort and filter notes by note type, time, and owner (Figure 1H). When a node with a

Date-stamp attribute is created, the system automatically places a box with the node's text on the timeline (Figure1B). Items on the timeline are ordered chronologically from left to right and items with similar date stamps are grouped.

The evidence cloud (Figure 1F) is a tag cloud of the documents attached to nodes as evidence. Font size indicates the frequency of attachment as evidence. Content of this view is based on information included in the workspace. If a user includes all collaborators' graphs, the evidence cloud includes all evidence items across the group. This view identifies documents that have been noted as relevant and reveals document importance based on frequency.

The implementation of CLIP supports full coordination of all views. When a node is enlarged to view details, the corresponding note and/ or timeline item (if present) are highlighted by fading out other items, enabling the user to quickly identify related items. Similarly, selecting a note or timeline item highlights corresponding items in other views. Clicking on a document name in the evidence cloud highlights nodes that contain the selected document as attached evidence.

4.3 Awareness Support

To support awareness of collaborators' activities, instances of CLIP that are running on different machines communicate in real-time to share information. Distinctive colours are used to distinguish work by different people, similar to Cambiera [21].

Partial merging: If another user has a node with the same name, then the local node is changed to notify the user that there is similar work. To keep changes in the local node subtle and yet noticeable, the only visual alteration is in the evidence list decorated around the node (colour coded segments). Evidence lists of the local node and the collaborator's node are combined, and repeating evidence segments are stacked up. Figure 3 shows a node 'George Prado' before and after a collaborator adds evidence. The colour of the local node is green and the colour of the collaborator is pink. Figure 3A shows the node before partial merging. Figure 3B depicts the same node after partial merging. In this case, the collaborators have two evidence items in common and other evidence items identified by only one person or the other. In addition to these visual cues, CLIP automatically combines all the collaborators' notes related to the node (ordered chronologically by default). By right clicking on a node, users can enlarge it (Figure 3C) to reveal more detail. Enlarging a node automatically highlights all related items (i.e., timeline items, notes, and evidence).

Tabs: Each tab in CLIP (Figure1D) encompasses a view of the analysis work in progress in another copy of CLIP. Tabs are labeled with the collaborator's colour and username to enable fast recognition of who owns the work. Tabs show a node-link layout that is identical to the node-link layout created by the owner of that information.

Full merging: Figure 1 depicts an example of a fully merged view. The merged design enables the viewer to easily gain an understanding of how their collaborators' work relates to their own (e.g., what entities their collaborators are interested in and why, and what evidence they have found). Figure 1E is a list of collaborators' names that can be used to decide whose work to merge with your own. Checking the box next to a collaborator's name merges all of the collaborator's work into the local view. CLIP re-computes the graph layout and unites nodes with the same name. The primary user's layout is maintained as much as possible in order to preserve their mental map.

5 User Study

We conducted a user study to gain a better understanding of LCW's effects on collaborative sensemaking. We employed a between-subjects experimental design to compare CLIP to a baseline tool (BT) with the LCW features removed.

5.1 Participants

We recruited 48 participants (16 groups of three, 8 groups / condition), who were graduate and senior undergraduate students from varied disciplines. To simulate work situations and create a comfortable environment, group members were required to know each other and have previous joint teamwork experience in a school or work project. To mitigate the impact of using students, we targeted participants who had some experience with data analysis. Participants' age ranged from 20-60 (Avg=28). There were 37 grad students and 28 males out of 48. We assigned groups randomly to each condition. Participants were compensated with $20 each. To encourage active participation, we provided a small financial reward for the team with the highest score.

5.2 Dataset and Scenario

We employed the "Stegosaurus" dataset from the VAST 2006 challenge [40]. This synthetic dataset contains approximately 240 documents, mostly news articles plus a few maps and supporting pieces of information. The documents describe approximately 3000 entities. The scenario involves finding a hidden chemical weapon production. We chose Stegosaurus because it is a standard task to evaluate visual analytics tools, and has been used in other studies [1,22,35]. We were also careful to select a task that represents a real life scenario but can be solved by non-experts. The dataset contains a scenario that reveals the first clue. From there, analysts are challenged to work through the dataset and iteratively search and filter to find the ten most relevant documents. Similar to real life scenarios, there are distractors that could point analysts in the wrong direction. While dataset authors estimated that the plot could be solved in about 2-6 hours with standard tools [40], we ended all of our sessions at 90 minutes as in [22].

5.3 Apparatus

Our experimental setup included two iMacs and one 17" MacBook Pro, arranged as shown in Figure 4. Participants were collocated and therefore could speak to each other and look at each other's screen if they wished. The physical arrangement was determined through pilot studies where we experimented with several different options to find an arrangement that was comfortable for participants. First we tried arranging the group members within a U-shaped table, so they could easily look at each other's screens; however we noticed that they did not have much discussion. Then we arranged them around a table with three laptops to simulate most current work practices, but we received many complaints about the small display size. This led us to the final larger-screen setup. We expected participants to complain about the screens blocking their views in this configuration, but they reported that it was very practical. We compared CLIP against a Baseline tool (BT). All participants in a group used the same version, either CLIP or BT. BT was identical to CLIP, except that we removed the LCW features (i.e., partial and full merging, as defined in section 4.3). We emphasize that BT still contained some awareness features; specifically, collaborators could still examine each other's work through tabs. We kept the tabs because they are similar to what many systems provide currently; for instance, Jigsaw's Tablet view [13] allows analysts to take notes in schematic form but offers no collaborative options to share between concurrent users. This approach also allowed us to specifically investigate the effects of the LCW technique.

Fig. 4: User study setup and physical arrangement.

5.4 Procedure

We assigned groups to conditions at random. The procedure for both versions was the same. We began with a tutorial on the system's features (15 minutes for BT, 20 minutes for CLIP). We asked participants to try out the system's features with a different sample data set. An observer was present to answer their questions and help them to experiment with all the features. Then, participants received background information about the task and started by reading the scenario, which provided the first clue. Documents were all digital, and all participants had access to all documents. Participants used Mac's Spotlight to search the text corpus. To search within a document, they used Microsoft Word's search functionality. They recorded their results into CLIP or BT. We ended the study whenever the teams were confident and ready to present their results, or at 90 minutes, whichever came first. Then we asked groups to write a report of their findings and hypotheses. Following the task, we conducted an open-ended interview with each group to discuss the system's features, their challenges, and suggestions to improve the system.

5.5 Measures and Hypotheses

In this section we summarize measures and hypotheses related to each of our research questions. We gathered data from five different sources: videos, interaction logs, the final written report submitted by each group, screen shots of visual elements created in CLIP or BT, and notes taken by the observer. All the sessions, including debriefing and interview, were audio and video-recorded. In total, we gathered 96 hours of video that includes the 16 groups' analysis and follow-up interviews. We used Transana [11] to analyze the videos and measure the total conversation time for each group.

5.5.1 Performance

To measure performance (RQ1), we analyzed groups' written reports. Using the same scoring scheme as Isenberg et al. [22], groups received positive points for facts they had connected (maximum of 11) and negative points for wrong hypotheses. 11 was the maximum possible score (i.e., all the facts were successfully discovered and connected) and a negative score means that the group uncovered few facts and produced incorrect hypotheses. In addition, and similar to [22], we also counted the number of discovered relevant documents as an indicator of performance. Successful completion of the task was partly related to participants' ability to find the 10 most relevant documents in the corpus and connect the facts within them. We analyzed the screen shots and logs to obtain the number of relevant documents discovered by each group. We hypothesized that CLIP groups would have better results on performance measures, as follows:

- H1: CLIP groups will have higher task scores and find a greater number of relevant documents than BT groups.

5.5.2 Communication, Coordination, and Awareness

We transcribed all the conversations to quantitatively measure communication effectiveness (RQ2), coordination (RQ3), and awareness (RQ4). Using an iteratively built coding scheme, we categorized each instance of conversation. We define an instance of conversation as one or more consecutive statements by a single individual. We chose to code instances of conversation because other possible units, such as sentences, are difficult to clearly delineate in oral conversation. The coding scheme was comprised of seven different categories (DH, RV, CO, SA, VF, QF, and RU). Table 1 depicts each code, along with its definition and example. Conversations were coded as DH whenever group members were engaged in a discussion trying to connect the facts and generate hypotheses. This was different from VF (verbalizing findings) when they were not actually connecting facts, they were only stating findings that they found interesting. This usually involved reading parts of a document out loud or reporting a summary of a finding. Referring to the visualization tool (RV) represented instances where participants orally referenced visualization elements such as nodes or notes. The seventh code, RU (Relevant but otherwise uncategorized), was used for any instance of conversation that was related to the case but did not fit within any of the former six codes. Sometimes a single instance reflected more than one code. For example, there were instances when participants were referring to the visualization and then they started to have a discussion about their findings and tried to connect them together. We coded these instances as both RV and DH. Other instances of double coding included RV and CO. Therefore, counts of the codes are not mutually exclusive. Over 2800 instances of conversation were coded using the scheme. We did not code conversations between group members and the experimenter.

Two independent coders coded the conversation data. We assigned groups randomly to each coder. Each coded 10 groups (5 CLIP and 5 BT groups), with 4 overlapping ones. Inter-coder reliability was 0.91, calculated using Krippendorff's alpha.

Research [2] has shown that fully sharing the work across the group can trigger discussions that are focused on solving the problem. Referring to the visualization also can enhance communication [17]. Because LCW should enable collaborators to more easily integrate their findings and discuss a shared view of their externalizations, we expected that CLIP groups would discuss more facts and hypotheses (DH) and refer to the visualization more often (RV):

- H2: CLIP groups will have more instances of DH and RV than BT groups.

We coded coordination (CO) utterances as those where collaborators tried to coordinate the group activities by dividing the task, documents, the search (e.g.,"*You search for flowers and I will search for apples*"), etc. According to prior research [12], we expected CLIP groups to better coordinate their work. We argue that if the tool supports better awareness, collaborators will be able to coordinate their actions at a much lower level of granularity. That is, instead of simply doing a high-level division of work at the beginning and then sharing findings at the end, collaborators will be able to continually adjust their task division as the work progresses. Therefore, we expected to see more CO instances with CLIP than with baseline:

- H3: CLIP groups will have more instances of CO than BT groups (because they will coordinate at a lower level of granularity).

In order to measure awareness, we coded conversations that were basically for seeking or sharing awareness about each other's activities and findings. For example, questions such as "*Are you guys going forward?*" or "*What have you found so far?*" were coded as seeking awareness (SA). Questions about another group member's finding(s) were coded as (QF), and verbalizing one's own findings as a way of sharing was coded as (VF). The rational behind this coding was that we noticed baseline groups spent more time interrupting other members to ask questions about findings or activities, and more time announcing their findings out loud. These questions and verbalizations could be easily eliminated if they could see each others' findings at a glance (the way they could see everyone's results in a merged view in CLIP).

- H4: CLIP groups will have fewer instances of SA, VF, and QF than BT groups (because they will be less reliant on the verbal channel for awareness).

To further explore awareness, we analyzed responses to the interview question about the extent to which participants were aware of each others' work. We also considered checkpoints, when in the middle of the session we stopped them and asked each individual to explain their findings and hypotheses. Then we asked them whether findings of one group member were surprising to others.

6 Results

In this section, we present both quantitative and qualitative findings of the study including the usage statistics of CLIP's main features.

6.1 Quantitative Findings

Table 2 presents scores achieved by CLIP and baseline groups, as well as results of the communication analysis. It reports the number of instances of discussion of hypotheses (DH), referring to the visualization (RV), coordination (CO), seeking awareness (SA), verbalizing findings (VF) and asking questions about another group member's findings (QF).

6.1.1 Task Performance

CLIP groups achieved considerably higher scores than baseline groups, strongly supporting H1. As shown in Table 2, scores of CLIP groups ranged from 5 to 11 (Avg=8.25, SD=2), whereas baseline

Code	Description	Example
DH	Having discussion or generating hypotheses	"US government is supplying the rebels with Lewisite."
RV	Referring to the visualization tool	"Link that with your apples. I will make a new node, linking Parazuelan."
CO	Coordinating the group	"Let's divide the work now, I will search for apples you look for flowers."
SA	Seeking awareness	"What do you guys got?"
VF	Verbalizing findings	"Former farm worker Francisco Dorado formed Shining Future in 1988."
QF	Questions about findings of another group member	"What did you find about apple bursting?"
RU	Relevant but otherwise uncategorized	"Oh okay found that article."

Table 1: Communication coding scheme.

groups were from -2 to 7 (Avg=2.75, SD=2.8). The maximum possible score was 11. A two-tailed t test showed a statistically significant difference between the average performance of CLIP and BT groups ($p < 0.001$). With the exception of group 3, all CLIP groups achieved 7 or higher. We believe the subpar performance of group 3 resulted from their strategy: they spent considerable time organizing the data chronologically before engaging in analysis.

With only one exception (G3, found 9 out of 10) all CLIP groups successfully found the 10 most relevant documents. On the other hand, only two baseline groups were able to find all of the relevant documents (G9 and G16). Even the top three ranked BT groups (9, 13 and 16) who found 10, 9 and 10 relevant documents respectively were not able to connect all the facts. A two-tailed t test showed a statistically significant difference ($p < 0.001$) between the average number of relevant documents found by CLIP (Avg=9.9, SD=0.4) and BT groups (Avg=6, SD=3). Task time was not an important factor. We found no correlation between scores and time (r^2 = 0.028), and no difference in average time between the conditions (CLIP Avg= 87.6 min, SD=88, BT Avg=86.8 min, SD=87), probably because the task was quite long and most groups used up nearly all the available time.

Tool	Group	Score	DH	RV	CO	SA	VF	QF
CLIP	12	11	185	178	57	0	15	5
	5	10	127	76	15	0	10	7
	8	10	124	26	23	1	1	6
	6	8	131	37	16	2	11	6
	15	8	123	15	10	4	4	3
	1	7	116	20	10	2	2	5
	11	7	102	20	20	1	6	4
	3	5	88	65	11	1	7	2
Avg	-	**8**	**116**	**30**	**15**	**2**	**5**	**4**
Baseline	9	7	116	17	10	9	38	27
	13	6	19	5	5	8	19	9
	16	5	114	6	14	7	10	5
	10	2	23	5	8	7	18	13
	4	2	20	5	9	11	21	15
	14	2	13	8	6	9	14	16
	2	0	11	4	5	4	5	2
	7	-2	25	9	4	1	3	5
Avg	-	**3**	**43**	**7**	**8**	**7**	**16**	**12**

Table 2: Comparison of performance, communication and coordination of CLIP versus Baseline groups.

6.1.2 Communication

H2 predicted that CLIP would foster discussion of facts and hypotheses (more DH). Our results strongly support this hypothesis (see Table 2). A two-tailed t test showed a significant difference in the number of DH utterances between CLIP and BT groups (CLIP Avg=116, SD= 28, BT Avg=43, SD=45, $p < 0.001$). Although there was no significant difference in the overall talking time between conditions, the difference in DH means that CLIP groups had significantly more discussions about hypotheses and connections between facts.

H2 also predicted that CLIP groups would refer to the visualization more often, and our results also confirmed this prediction. CLIP groups extensively referred to the visualization tool (RV), significantly more often than BT groups (CLIP Avg=30, SD=55, BT Avg= 7, SD=4, $p < 0.001$). We also observed that there was more discussion triggered by the system in CLIP groups. This was mostly when participants realized their teammate had done some related work. CLIP groups also had fewer awareness seeking conversations (see section 6.1.4).

6.1.3 Coordination

H3 predicted more instances of CO in CLIP groups than in BT groups (reflecting more detailed task division). We found a significant difference in the number of CO instances (CLIP Avg=15, SD=16, BT Avg= 8, SD=3, $p < 0.01$). In relation to this, we also noticed many instances where CLIP groups coordinated their work via the tool. To further analyze the effect of the visualization tool on coordination, we looked into RV examples that were double coded with CO. For instance, we coded this as RV and CO: "*Link my node with your apples and I will make a new node to link Parazuela*". This is an example of coordination where collaborators deliberately connected their results through the tool in order to solve the problem. We observed and recorded many of these instances for CLIP groups.

6.1.4 Awareness

H4 predicted that using LCW would help collaborators to maintain awareness of each other's work with less reliance on verbal communication. Conversation analysis strongly supported this hypothesis. CLIP groups had significantly fewer awareness seeking utterances (SA) (CLIP Avg=2, SD=1, BT Avg=7, SD=3, $p < 0.001$). CLIP groups reported that it was much easier to figure out who was doing what by looking at the merged view. CLIP groups also verbalized their findings significantly less than BT groups (CLIP Avg=5, SD= 5, BT Avg=16, SD=11, $p < 0.04$). There was a marginally significant difference in the number of QF (CLIP Avg=4, SD=2, BT Avg=12, SD=8, $p < 0.06$).

6.2 Qualitative Findings and Usage Statistics

Three primary awareness channels were available to participants: oral communication, LCW (CLIP only), and tabs. To complement and elaborate on our quantitative conversation analysis, in the following sections we report qualitative observations and the results of our post-task interviews for each awareness channel. In the interviews, all CLIP users reported being aware of their collaborators' work most of the time. They all attributed this to use of LCW features, especially full merging. They found partial merging cues to be an interesting notification of common work that helped them to understand who else had related results and evidence. However, all of the CLIP participants attributed their awareness to full merging. Two CLIP groups (G6, G1) indicated that showing collaborators' notes was another important feature that helped them to maintain awareness of each others' work. In contrast, many baseline groups mentioned that they were not aware of each others' work. Five out of eight groups reported oral communication as their main awareness mechanism. The rest reported that their awareness channels were oral communication as well as using tabs. These results are consistent with our RV, SA and QF findings. Table 3 shows the usage statistics of CLIP's main features other than LCW.

Tool	Node	Note	Link	Timeline	Tab	E. Cloud
CLIP	20(6)	22(6)	12(5)	10(3)	52(50)	15(5)
BT	10(7)	14(8)	7(8)	6(4)	71(40)	11(12)

Table 3: Usage statistics for CLIP and Baseline (AVG (SD)).

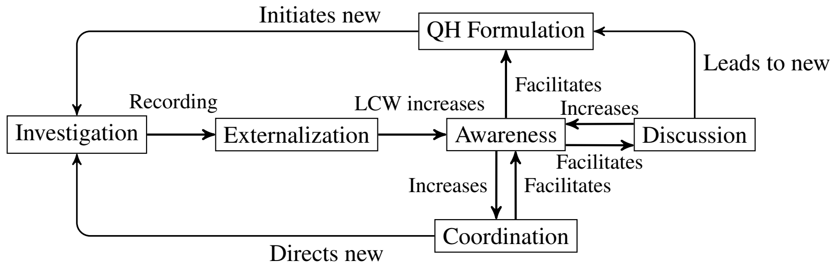

Fig. 5: Collaboration model emphasizing the role of LCW in increasing awareness and discussion among team members. Awareness leads to better coordination of activities and formulation of new questions/hypotheses, which in turn initiate and direct new investigation.

6.2.1 Oral Communication

CLIP groups' oral communication focused heavily on discussion of hypotheses, coordination and referring to the visualization. By contrast, for baseline groups, oral communication was the dominant awareness channel, and without it, participants were not aware of each others' activities most of the time. For instance, one member of group 10 asked, "*What are you guys doing, why you don't talk?*" and a member of group 4 stated, "*I felt I did not know what [Participant A] was doing [she was silent most of the time]*". A similar result was reported by Wallace et al. [39], who did not provide any form of thinking or note space for participants in their study. We observed two key problems associated with communicating only through the verbal channel. Sometimes sharing out loud disturbed others. One baseline participant asked her teammate twice to be quiet. Her teammate was trying to share his findings frequently and make sure that they were all aware of each others' work, but she wanted to focus. Instead she decided to write comprehensive notes to share with the others. Another participant stated, "*I couldn't read my stuff when others were telling me what they read, it was grabbing my focus*". The second problem with verbal sharing was that if the information was not recorded, it could be easily forgotten. Although speech was the fastest way to get updated about others' work, there were a few instances where key facts were shared verbally, but later on, the group did not report those key facts in their debriefing (and thus received a lower score than they might have). We noticed this in particular for two CLIP groups (G1 and G6); session logs showed that those key facts were never entered into CLIP.

6.2.2 LCW (CLIP only)

Participants reported merging individual work (all CLIP groups, 22 group members), LCW (6 out of 8 groups), and LCW of notes (6 out of 8 groups) as the three most useful features of the system. CLIP participants made extensive use of the LCW features, especially the ability to merge everyone's node-link graphs together. This could cut some unnecessary communication, reduce redundancy, and let team members focus on the task better. According to one CLIP participant, "*[Merging] made it faster because we knew what everyone was looking at. We could go on the same direction or do something else...it helped us to collaborate more closely, when you're not paying attention that much, especially I said, 'Hey, what about this?' and someone else is like, 'We've already done this,' and you can just look at their graph. And I can connect my stuff to theirs*".

Looking at the most successful groups (12, 5 and 8, all using CLIP), we observed that the strategies they had in common were a clear work division and extensive use of CLIP's LCW features, suggesting that these were good predictors of success. Even though these groups had different leadership styles, they constantly divided the workload. According to the system log, these groups also made intense use of CLIP's features to coordinate their activities. They systematically merged their partners' work into their own view to link their work together. Participants reported that the merged view helped them find important results in others' work. It also inspired confidence and helped them identify relevant keywords. Participants said, "*I noticed that some of my most powerful points... he also had them. I could see the two colours on it. That gave me confidence,*" and, "*For common nodes, I was looking at the evidence. If they were different from mine, I was checking them as well. Common items made me confident and helped me to keep going.*" CLIP groups became quite dependent on the shared node-link graph. For instance, in group 5, B was sharing her findings with A, and A said "*Put it down, create a node*". Later when B complained that the team ignored one of her findings ("*I found it before and I told you!*"), A explained why it had gone unnoticed, "*Because you did not connect it to my node, so I did not look at it!*".

We were curious to see how participants in CLIP groups would choose to use merging. Would they leave the default setting (partial merging) to keep their workspace uncluttered and avoid the disruption of constant updates from other participants' changes? Or would they choose to see everything? Answer: the latter. Most participants chose to set merging on from the beginning and kept it visible until the end.

It was interesting that participants reported that oral communication was disruptive, but CLIP updating the shared view was not. Instead, participants reported merging to be useful for collaboratively exploring the task, sharing important evidence, exchanging documents, and reducing redundant work. During the interview participants emphasized that merging was one of the most useful features of CLIP.

Five groups (eleven participants) in the baseline condition actually requested a merging feature that would put everyone's information in one view. For example, participants stated, "*I was not able to make a link to someone else's work, so I could not make a connection,*" and "*It is hard to remember what the others have registered by checking the tabs, so we would like to be able to draw links between nodes created by different people. It is also good to avoid redundancy.*" Only one participant reported a potential negative side of merging, stating, "*It was interesting, but it was a double-edged sword. It could help me or push me [in a] correct or incorrect direction.*"

6.2.3 Tabs and Notes

In addition to oral communication, most baseline groups also relied heavily on tabs for awareness. Some groups, however, used tabs only to quickly check what the others were working on; for example, "*I only looked at their tabs when I was trying to find something that they have read. I just wanted to refer to their work but not for everything*". Interestingly, tabs usage in CLIP was not much lower than BT (see Table 3), even though they also used the merging feature. One CLIP participant explained that she used tabs to see how other group members arranged their nodes (because CLIP's merge feature recomputed the layout for each individual to maintain their mental map).

Notes were valued in both CLIP and BT. Collaborators' notes were accessible via tabs in both tools, and via merging in CLIP. Participants stated that notes provided an overview, enabled them to remember why they had created graph nodes, and allowed them to copy important information from the documents. Several people reported that notes helped them to identify interesting information belonging to others. Participants in group 12 stated, "*[C:] The notes on the side. I got most info from them, to be honest. I would read the notes and go 'Wow, that's cool!' [B:] Yeah, other people were highlighting things that you should read.*" Similarly, another participant said, "*When someone didn't write a good note, I didn't look at what they were doing*".

7 Discussion and Future Work

LCW clearly supported groups in this collaborative sensemaking task. CLIP groups achieved significantly better scores (H1), coordinated and communicated more effectively (H2 and H3), and relied on LCW to maintain awareness of each others' work (H4). CLIP groups had significantly more discussion about hypotheses and evidence (DH) and were able to focus their oral communication on discussing the case and coordination activities rather than using oral communication as the main awareness channel. This research extends earlier work on LCW, by establishing its value in the sensemaking loop of collaborative analysis, not just the information foraging loop, and demonstrating how it can be applied to externalizations. CLIP also illustrates how the LCW concept can be employed within a collaborative thinking space.

To better explain the effects of using CLIP on teams' collaboration, we derived a collaboration model for CLIP groups based on our results (see Figure 5). Similar to the model in [12], our model shows how awareness plays a critical role that enhances communication and coordination activities. From Figure 5, we can see that recording externalizations and automatically sharing them via LCW increased awareness. Increased awareness in turn enabled groups to coordinate their work at a deeper level. Being able to see others' results triggered discussion and this assisted teams to better formulate their new questions and hypotheses (QH Formulation). There are mutual effects between awareness, coordination and discussion; i.e., each one will influence the other. QH Formulation and coordination of activities initiate and direct new investigation. With BT, LCW was missing. In our model in Figure 5, this means that the link between Externalization and Awareness was effectively broken. Collaborators still maintained some awareness through oral communication, but this mechanism was less effective. Reduced awareness in turn had detrimental effects on the teams' coordination, discussion, and investigative activities.

While CLIP provided an effective thinking space for the intelligence analysis task in our study, additional work would be needed to extend it for more general use. To begin with, CLIP aims to support the sensemaking loop, and therefore provides no explicit support for information foraging. Combining CLIP with a complementary tool like Cambiera [21] may be an effective way to support both phases, which may be especially crucial when dealing with a larger document set. Awareness cues related to information foraging could present collaborators' current activities (e.g., revealing that a collaborator is reading a document or entering a search query). Another limitation of CLIP is that it does not automatically identify entities or relationships between documents; we used Jigsaw to extract this information for the purpose of our study. We would like to integrate CLIP's thinking space and LCW features into a document analytics tool such as Jigsaw [37] that automatically extracts entities and relationships.

Another way to extend CLIP, as suggested by one group, would be to add visual indications that distinguish past work from current changes (e.g. colour saturation to indicate node age). Although the dynamics of the node-link graph show the evolution of the team's findings, it is not clear at any given time which changes are the most current. Another interesting feature suggested by participants was to create a summary evidence file to help with publication of the results.

Scalability is another important issue. In CLIP, collapsing a node collapses all the branches that stem from the node, improving overall scalability and flexibility. To scale this design to large and complex problems, however, different visual representations might be needed. The visual structures (e.g., node-link graph) may not scale well even for individuals, and with multiple analysts, keeping track of collaborators' changes and updates to such a large representation may be impossible. We predict that the 'share-everything' strategy that was successful for CLIP groups in our study might break down at a larger scale. A variation of Branch-Explore-Merge [30] might reduce the number of visual updates since they would only appear upon merging. For small thinking spaces, this may be a significant disadvantage since awareness notifications would be delayed. However, in large thinking spaces, providing awareness notifications in such chunks may cause less visual distraction and reduce the likelihood of small updates being missed. There are also other scalability issues in the current design. First, while the colour coding works well for small groups (our target), it should be reconsidered for larger groups. Also, for the specific task used in this study, decorating evidence around a node was enough. However, the design might need to change for larger datasets with more evidence items. One possible way to improve scalability could be to encode the quantity of evidence related to a node as the node size. Similarly, the size of notes could adapt to their length.

One interesting question that arises with visual thinking spaces is the potential that they may lead to *group-think*, a situation where the group fails to consider possible explanations because they too quickly follow one avenue of investigation. It is possible that sharing findings through LCW may discourage a healthy level of independent analysis. We do not have a good way to assess the level of group-think in our study, in part because avenues of independent thought are neither easy to categorize nor measure. One possible approach to avoid group-think could be to design tools that promote discussion of alternative hypotheses, perhaps by finding and highlighting disagreements in the findings. Research into causes of group-think and mechanisms to prevent it are an important area of future research. Nonetheless, the much stronger performance of CLIP groups in our study indicates that the awareness benefits of LCW outweigh costs such as group-think.

Future work should also examine the value of LCW in a field setting with professional analysts. Our participants were students because it is extremely difficult to find enough professional analysts for a lab experiment. We took care to recruit participants with some data analysis experience and chose a task that did not require domain-specific knowledge. Nonetheless, student behaviour will undoubtedly differ from that of experts. For example, professional analysts might have established coordination strategies and therefore be less reliant on tool support for coordination. We would also like to explore how LCW influences collaborative dynamics over a longer analysis period.

Another interesting future direction is to understand how CLIP could be used on a shared screen (e.g., a wall or tabletop). Finally, although we examined the value of LCW for collocated work, it might have even greater value for distributed or asynchronous scenarios. Maintaining awareness is generally more challenging in these situations because of the limited communication channels available to collaborators. LCW could play a critical awareness role in such situations, but this will need to be tested in future studies. It is quite possible that additional features will be needed (e.g., a more extensive note feature that enables threaded discussions) when verbal and / or non-verbal awareness communication channels are unavailable.

8 Conclusion

CLIP demonstrates how the concept of linked common work can be employed within collaborative thinking spaces to support the sensemaking loop during collaborative analytics. CLIP provides an environment for analysts to record, organize, share and connect results. Moreover, CLIP extends earlier thinking spaces by integrating LCW features that reveal relationships between collaborators' externalizations to increase awareness among team members. Our user study compared CLIP to a baseline version without LCW features. Results demonstrated that LCW significantly improved analytic outcomes at a collaborative intelligence task. Groups using CLIP were able to communicate and coordinate more effectively. They were able to use oral communication primarily to discuss the task, generate hypotheses, and coordinate their activities at detailed level, rather than employing it for disruptive awareness notifications. Most importantly, LCW enabled collaborators to maintain awareness of each other's activities and findings and link those findings to their own work.

9 Acknowledgements

We thank Leandro Collares and Wanda Boyer for their assistance with data analysis, and the VisID group at UVic for their helpful suggestions. Thanks also to Rock Leung and our other colleagues at SAP for their feedback on system design and to Mark Whiting for sharing the dataset. This research was funded by SAP, NSERC, and GRAND.

REFERENCES

[1] C. Andrews, A. Endert, and C. North. Space to think: large high-resolution displays for sensemaking. In *Proc. SIGCHI Conference on Human Factors in Computing Systems*, pages 55–64. ACM, 2010.

[2] A. D. Balakrishnan, S. R. Fussell, and S. Kiesler. Do visualizations improve synchronous remote collaboration? In *Proc. Human Factors in Computing Systems*, pages 1227–1236. ACM, 2008.

[3] A. D. Balakrishnan, S. R. Fussell, S. Kiesler, and A. Kittur. Pitfalls of information access with visualizations in remote collaborative analysis. In *Proc. Computer supported cooperative work*, pages 411–420. ACM, 2010.

[4] E. A. Bier, E. W. Ishak, and E. Chi. Entity workspace: an evidence file that aids memory, inference, and reading. In *Proc. Intelligence and Security Informatics*, pages 466–472. Springer, 2006.

[5] L. Bradel, A. Endert, K. Koch, C. Andrews, and C. North. Large high resolution displays for co-located collaborative sensemaking: Display usage and territoriality. *Int. J. Human-Computer Studies*, 71(11):1078–1088, 2013.

[6] S. E. Brennan, K. Mueller, G. Zelinsky, I. Ramakrishnan, D. S. Warren, and A. Kaufman. Toward a multi-analyst, collaborative framework for visual analytics. In *Proc. Visual Analytics Science And Technology (VAST)*, pages 129–136. IEEE, 2006.

[7] Y. Chen, J. Alsakran, S. Barlowe, J. Yang, and Y. Zhao. Supporting effective common ground construction in asynchronous collaborative visual analytics. In *Proc. Visual Analytics Science and Technology (VAST)*, pages 101–110. IEEE, 2011.

[8] G. Chin Jr, O. A. Kuchar, and K. E. Wolf. Exploring the analytical processes of intelligence analysts. In *Proc. Human Factors in Computing Systems*, pages 11–20. ACM, 2009.

[9] M. C. Chuah and S. F. Roth. Visualizing common ground. In *Proc. Int. Conf. on Information Visualization*, pages 365–372. IEEE, 2003.

[10] H. H. Clark and S. E. Brennan. Grounding in communication. *Perspectives on socially shared cognition*, 13(1991):127–149, 1991.

[11] C. Fassnacht and D. Woods. Transana (version 2.20)[computer software]. *Madison, WI: School of Education at the University of Wisconsin-Madison*, 2007.

[12] V. L. Gava, M. d. M. Spinola, A. C. Tonini, and J. C. Medina. The 3c cooperation model applied to the classical requirement analysis. *JISTEM-Journal of Information Systems and Technology Management*, 9(2):235–264, 2012.

[13] C. Görg, Z. Liu, and J. Stasko. Reflections on the evolution of the jigsaw visual analytics system. *Information Visualization*, 2013.

[14] D. Gotz, M. X. Zhou, and V. Aggarwal. Interactive visual synthesis of analytic knowledge. In *Proc. Visual Analytics Science and Technology (VAST)*, pages 51–58. IEEE, 2006.

[15] C. Gutwin and S. Greenberg. Design for individuals, design for groups: tradeoffs between power and workspace awareness. In *Proc. Computer supported cooperative work*, pages 207–216. ACM, 1998.

[16] A. H. Hajizadeh, M. Tory, and R. Leung. Supporting awareness through collaborative brushing and linking of tabular data. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2189–2197, 2013.

[17] J. Heer and M. Agrawala. Design considerations for collaborative visual analytics. *Information visualization*, 7(1):49–62, 2008.

[18] J. Heer, F. van Ham, S. Carpendale, C. Weaver, and P. Isenberg. Creation and collaboration: Engaging new audiences for information visualization. In *Information Visualization*, pages 92–133. Springer, 2008.

[19] M. S. Hossain, C. Andrews, N. Ramakrishnan, and C. North. Helping intelligence analysts make connections. In *Proc. Scalable Integration of Analytics and Visualization*, 2011.

[20] P. Isenberg, N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, and H. Hagen. Collaborative visualization: definition, challenges, and research agenda. *Information Visualization*, 10(4):310–326, 2011.

[21] P. Isenberg and D. Fisher. Collaborative brushing and linking for co-located visual analytics of document collections. In *Computer Graphics Forum*, volume 28, pages 1031–1038. Wiley Online Library, 2009.

[22] P. Isenberg, D. Fisher, S. A. Paul, M. R. Morris, K. Inkpen, and M. Czerwinski. Co-located collaborative visual analytics around a tabletop display. *Visualization and Computer Graphics, IEEE Trans.*, 18(5):689–702, 2012.

[23] E. Kandogan, J. Kim, T. P. Moran, and P. Pedemonte. How a freeform spatial interface supports simple problem solving tasks. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, pages 925–934. ACM, 2011.

[24] Y.-a. Kang, C. Gorg, and J. Stasko. Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study. In *Proc. Visual Analytics Science and Technology (VAST)*, pages 139–146. IEEE, 2009.

[25] Y.-a. Kang and J. Stasko. Characterizing the intelligence analysis process: Informing visual analytics design through a longitudinal field study. In *Proc. Visual Analytics Science and Technology (VAST)*, pages 21–30. IEEE, 2011.

[26] K. Kim, W. Javed, C. Williams, N. Elmqvist, and P. Irani. Hugin: A framework for awareness and coordination in mixed-presence collaborative information visualization. In *ACM Intl. Conference on Interactive Tabletops and Surfaces*, pages 231–240. ACM, 2010.

[27] N. Mahyar, A. Sarvghad, and M. Tory. Note-taking in co-located collaborative visual analytics: Analysis of an observational study. *Information Visualization*, 11(3):190–204, 2012.

[28] N. Mahyar, A. Sarvghad, M. Tory, and T. Weeres. Observations of record-keeping in co-located collaborative analysis. In *Proc. Hawaii Int. Conf. on System Sciences (HICSS)*, pages 460–469. IEEE, 2013.

[29] G. Mark and A. Kobsa. The effects of collaboration and system transparency on cive usage: an empirical study and model. *Presence: Teleoperators and Virtual Environments*, 14(1):60–80, 2005.

[30] W. McGrath, B. Bowman, D. McCallum, J. D. Hincapié-Ramos, N. Elmqvist, and P. Irani. Branch-explore-merge: facilitating real-time revision control in collaborative visual exploration. In *Proc. Interactive tabletops and surfaces*, pages 235–244. ACM, 2012.

[31] D. R. Millen, A. Schriefer, D. Z. Lehder, and S. M. Dray. Mind maps and causal models: using graphical representations of field research data. In *CHI'97 Extended Abstracts on Human Factors in Computing Systems*, pages 265–266. ACM, 1997.

[32] T. P. Moran, W. Van Melle, and P. Chiu. Spatial interpretation of domain objects integrated into a freeform electronic whiteboard. In *Proc. ACM Symp. User Interface Software and Technology*, pages 175–184. ACM, 1998.

[33] S. A. Paul and M. C. Reddy. Understanding together: sensemaking in collaborative information seeking. In *Proc. Computer Supported Cooperative Work*, pages 321–330. ACM, 2010.

[34] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proc. Int. Conf. on Intelligence Analysis*, volume 5, pages 2–4, 2005.

[35] J. M. R. and H. Kasper. Collaborative work on a high-resolution multi-touch wall-display. *Trans. Computer-Human Interaction*, 21(2):340, 2014.

[36] A. C. Robinson. Collaborative synthesis of visual analytic results. In *Proc. Visual Analytics Science and Technology (VAST)*, pages 67–74. IEEE, 2008.

[37] J. Stasko, C. Görg, and Z. Liu. Jigsaw: supporting investigative analysis through interactive visualization. *Information visualization*, 7(2):118–132, 2008.

[38] K. Vogt, L. Bradel, C. Andrews, C. North, A. Endert, and D. Hutchings. Co-located collaborative sensemaking on a large high-resolution display with multiple input devices. In *Proc. Human-Computer Interaction, INTERACT 2011*, pages 589–604. Springer, 2011.

[39] J. R. Wallace, S. D. Scott, and C. G. MacGregor. Collaborative sensemaking on a digital tabletop and personal tablets: prioritization, comparisons, and tableaux. In *Proc. Human Factors in Computing Systems*, pages 3345–3354. ACM, 2013.

[40] M. A. Whiting, C. North, A. Endert, J. Scholtz, J. Haack, C. Varley, and J. Thomas. VAST contest dataset use in education. In *Proc. Visual Analytics Science and Technology (VAST)*, pages 115–122. IEEE, 2009.

[41] W. Willett, J. Heer, J. Hellerstein, and M. Agrawala. Commentspace: structured support for collaborative visual analysis. In *Proc. Human Factors in Computing Systems*, pages 3131–3140. ACM, 2011.

[42] W. Wright, D. Schroh, P. Proulx, A. Skaburskis, and B. Cort. The sandbox for analysis: concepts and methods. In *Proc. Human Factors in computing systems*, pages 801–810. ACM, 2006.

[43] J. Zhang. The nature of external representations in problem solving. *Cognitive science*, 21(2):179–217, 1997.