

# Supporting Collaborative Information Analysis: A Classroom Study

AUTHOR, Institution Address City State Postcode Country

AUTHOR, Institution City Arunachal Pradesh Country

AUTHOR, Institution City Arunachal Pradesh Country

---

Information analysis involves dynamic modeling and exploration of complex information spaces, and is fundamentally collaborative at any non-trivial scale. That said, the tools supporting information analysis either do not support such dynamic workflows, or they do not support collaboration at all. Along these lines, we developed *CAnalytics* to address these two critical challenges by: enabling interleaving of data modeling and analysis within a single workspace; and supporting collaboration through shared annotation and increased activity awareness. This paper describes a classroom study of analysts-in-training and examines their tool usage over multiple usage sessions. We report an interleaving work patterns students spontaneously employ with our tool, as well as the analytic behaviors that drive the interleaving pattern. We also note the distinctions among collaborative strategies, and project improved awareness design to enable a more interleaving workflow.

CCS Concepts: •**Computer systems organization** → **Embedded systems**; *Redundancy*; Robotics; •**Networks** → Network reliability;

Additional Key Words and Phrases: Information analysis, collaboration, visualization, classroom study, activity awareness

## ACM Reference format:

Author, Author, and Author. 2010. Supporting Collaborative Information Analysis: A Classroom Study. *ACM Trans. Web* 9, 4, Article 39 (March 2010), 19 pages.

DOI: 0000001.0000001

---

## 1 INTRODUCTION

Collaborative information analysis is a form of sensemaking wherein a team of analysts identify and evaluate causal relationships in a complex corpus of documents. A common example of such an analysis is intelligence analysis, where a variety of putative facts are assembled, e.g. financial records, witness observations and interviews, and the social connections between persons of interest. Armed with these different assembled facts investigators collaboratively: assess means, motives, and opportunities; articulate and investigate further hypotheses and deductions; and develop one or more theories of the crime. Other examples include business intelligence, scientific research, social constructivist learning, etc.

---

To be added.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2009 Copyright held by the owner/author(s). Publication rights licensed to ACM. 1559-1131/2010/3-ART39 \$15.00

DOI: 0000001.0000001

According to Pirolli and Card's model [?], information analysis starts with modeling data from textual documents, followed by representing these data models in various information artifacts, and developing them into hypotheses. Systems that currently support information analysis are aimed at a single phase and therefore only support part of the overall analysis workflow. This imposes a clear boundary between each of these phases on the analysts. For example, Analysis of Competing Hypotheses (ACH) assumes that data has been modeled, and that relevant evidence can be adduced appropriately to various hypotheses, but provides no structured support for either. Other techniques, such as Information Extraction and Weighting (IEW), help structure evidence modeling, but do not extend utilization of evidence to hypothesis generation. The unintended boundary between phases has the consequence that data modeled in one software cannot be effectively utilized in hypothesis development in another system. And analysts have to handoff, often via replicating the data in the new system, information between software systems, making it difficult to revisit and revise the data model. We therefore are motivated to develop an integrated workspace in which analysts can model and analyze data in one place, and we utilize the system as an experimental instrument to investigate analytic behavior afforded by such integration.

Any information analysis activity, at least at a non-trivial scale, is fundamentally collaborative [?]. The intelligence community puts great value on collaboration. A report from the Director of National Intelligence, *Vision 2015*, called for "*a dramatic shift from traditional emphasis on self-reliance toward more collaborative operations*" [?, p.13]. However, most analytic tools that are widely used in the intelligence community (e.g. Analyst's Notebook [?] and PARC ACH [?]) do not support collaboration. The community has to rely on separate collaboration tools (e.g. email, Intellipedia [?]), which lack serious support for analytics, for team coordination [?]. Thus, analysts must coordinate their work outside of their tool support, manually sharing their analytic products.

Supporting collaboration in information analysis is challenging, and perhaps more than in other situations (e.g. collaborative writing, wiki) because the task itself can be extremely complex. A team could be working with much more complex data structure (e.g. spatial data, temporal data, and relational data, as opposed to text only), coordinating through multiple analytic artifacts (e.g. visualizations as opposed to document only), and making sense of different levels of analysis, assumptions, and hypotheses, both synchronously and asynchronously throughout a long-term course of collaborative interaction. Thus teams must not only stay aware of what other members are doing, but *why* they are doing that in a specific context of analysis. This study aims to investigate how technology can mediate team analysis and to understand what awareness is needed beyond team actions.

This paper reports specifically on a collaborative information analysis task situated in the intelligence domain, one of the most complex information analysis areas. We performed our study in an intelligence analysis course. A classroom study provides a natural environment in which participants engage in multi-session, relatively complex class projects. Due to the difficulty in accessing professional analysts due to security and confidentiality issues, studying *analysts-in-training* provides us a chance to include them in a longitudinal design loop. These students already have knowledge and experience with state-of-the-art analytic techniques and tools and are thus more likely to provide insightful feedback. Besides, the students are young learners that are willing to employ new work practices supported by features in tools. They are important parts of the future intelligence community. In some sense, their practice can be treated as a view into the future of practice of the community [?].

We have developed a tool that includes annotation for data modeling, interactive visualization for data analysis, and collaboration features. While this paper reports a collaborative task in a specific domain, findings regarding team process and breakdowns meet the interest of the broader

CSCW community. This study makes three contributions: 1) we observed a spontaneously adopted interleaving workflow and characterized how such interleaving occurred as well as its consequence; 2) we distinguished three labor division strategies teams employed with our collaboration support, and our results implied that a more interleaving workflow could be enabled by awareness of uncertainty, context of insight, and contribution value.

## 2 RELATED WORK

### 2.1 Information analysis workflow

Pirolli and Card's Think Loop Model [?] is a widely used model of information analysis. The model describes the process of information foraging and sensemaking in which raw evidence is successively modeled, filtered, and synthesized into a best hypothesis. The model is a bottom-up process of structure building, but also includes a local feedback loop at each stage. Thus, analysts can reconsider propositions in the evidence file, asking how they are related, or a given hypothesis, asking what schemata it rests upon.

Empirical studies of information analysis suggest that the iterative looping can have a wider scope than is obvious in the Pirolli and Card's model [?]. For example, Chin et al. [?] observed five professional intelligence analysts working both individually and as a team. They found that analysts often need to review the original documents even at advanced stages of analysis. The scope of these reconsiderations is not consistent with the local feedback architecture of the Think Loop Model. It seems more consistent with a parallel or multi-phased model [?] in which structure building occurs at a variety of levels in parallel. Similar findings were reported by other empirical studies [?, ?, ?]. A report [?] from a workshop of professional intelligence analysts listed "dynamic data processing and visualization" as one of top requirements in computational support for intelligence analysis, emphasizing the need for an integrated environment for data modeling and analysis.

However, tools supporting information analysis are often designed targeted at a single phase of activity, and thus not supporting the whole workflow. For example, research efforts have been made to understand information collection and modeling [?, ?], but little support is provided to extend these models to analysis. Techniques such as Information Extraction and Weight (IEW) helps structures data evidence but offers no structure to turn the evidence to hypothesis development. Similarly, tools supporting the activity of data analysis assume data has been modeled. Analytic tools such as interactive visualization emphasize present data in insightful means but provides no utility to data re-modeling [?].

Similar calls were made in other data intensive task domains as well. For example, in interactive machine learning, researchers [?, ?] call for an all-in-one environment in which machine learning practitioners can tune model parameters and evaluate model performance through visualization in one place. In the area of visual analytics, Ware [?] warned of the "*asymmetry in data rates*" [p.382], pointing out that visual analytic tools emphasized data flowing from systems to users far more than from users to systems. Functionalities are mostly designed to adjust data representation rather than modeling, which are in fact equally important. Our work aligns with these efforts, and contribute to the design and evaluation of an integrated workspace in supporting information analysis tasks.

### 2.2 Collaboration and awareness in information analysis

Collaboration is critical in information analysis. The intelligence community has called for collaboration, and indeed closer collaboration, as opposed to merely coordinating draft products by the end [?]. Empirical studies [?, ?] reported that analysts' practice was fundamentally collaborative. Effective collaboration helps analysts pool different knowledge and expertise and reduces the effect of confirmation bias [?].

Significant efforts have been made to understand and support collaboration in the CSCW community, and many of them address, as it is in our study, tasks with large and complex data, such as collaborative modeling [?, ?], collaborative information seeking [?, ?], and collaborative visualization [?, ?]. While these works provide insight in supporting teamwork with complex artifacts, less is known in terms of how technology-mediated collaboration occurs in a complete analytic process, and what awareness is needed for analytic specific purposes.

The state-of-the-art computational tools supporting intelligence analysis either do not support collaboration, or they do not integrate serious analytic support. On one hand, most analytic tools (e.g. Analyst's Notebook and PARC ACH) are designed for single user only, with only a few exceptions (e.g. Te@mACH [?]). On the other hand, as Treverton [?] reviewed state-of-the-art collaboration tools in the intelligence community, those tools only support coordination for general situations, without specific analytic functionalities. For example, Intellipedia [?] is a wiki platform for sharing of intelligence reports and documents, yet it does not integrate functions for data modeling and analysis at all.

A key enabler for effective collaboration is *activity awareness*, defined as team's awareness of its own sustained collaborative activity [?, ?]. Derived from Activity Theory [?], activity awareness encompasses information covering all aspects of collaboration, such as partner presence, mediating artifacts, group actions, social interactions, shared information, and group values and norms. Activity awareness has been utilized as a design concept in guiding and evaluating collaboration features.

Many lab studies have been reported to investigate specific awareness features to support collaboration in information analysis. For example, Convertino et al. [?] examined the use of public and private views for role based collaboration. Goyal and Fussell [?] studied the effect of hypotheses sharing on sensemaking. Mahyar and Tory [?] designed a visualization to connect collaborators' common findings and evaluated its support for team performance. Hajizadeh et al. [?] explored how sharing teammate's interactions affects awareness. These studies provide evidence to validate hypotheses of specific design features. However, due to time constraint (mostly within one hour), these studies had to employ a simplified task with reduced content and complexity. Artifacts created by participants were thus relatively simple and superficial (e.g. with a single artifact or few items in an artifact). More complex task would have pushed participants to create more sophisticated artifacts (e.g. multiple views or cluttered display that requires filtering) and to try balancing between team coordination and individual reasoning, which would have provided more insights into team-based analytic process. Our study aims to examine tool usage in a classroom context, which allows for higher task complexity and longer teamwork time in a more realistic environment.

### 3 CANALYTICS FEATURES

As shown in Figure 1, we developed a web application tool, *CAanalytics* (standing for "C-ollaborative Analytics"), to support teams of analysts in identifying, visualizing, integrating and assessing facts from multiple sources. Two particular design objectives are 1) to build an integrated environment for intelligence analysis, and 2) to support collaboration with shared data and awareness functions. The design is informed by earlier paper prototype studies [?, ?], in which the researchers examined team's communication patterns and spontaneously created artifacts in a crime scenario. We also take into account findings from empirical studies conducted by Chin et al. [?] and Kang and Stasko [?] when making design decisions.

*Coordinated multiple views for data modeling and analysis.* To enable an integrated environment for different functionalities of analysis, CAanalytics employs a multiple-view interface, with each

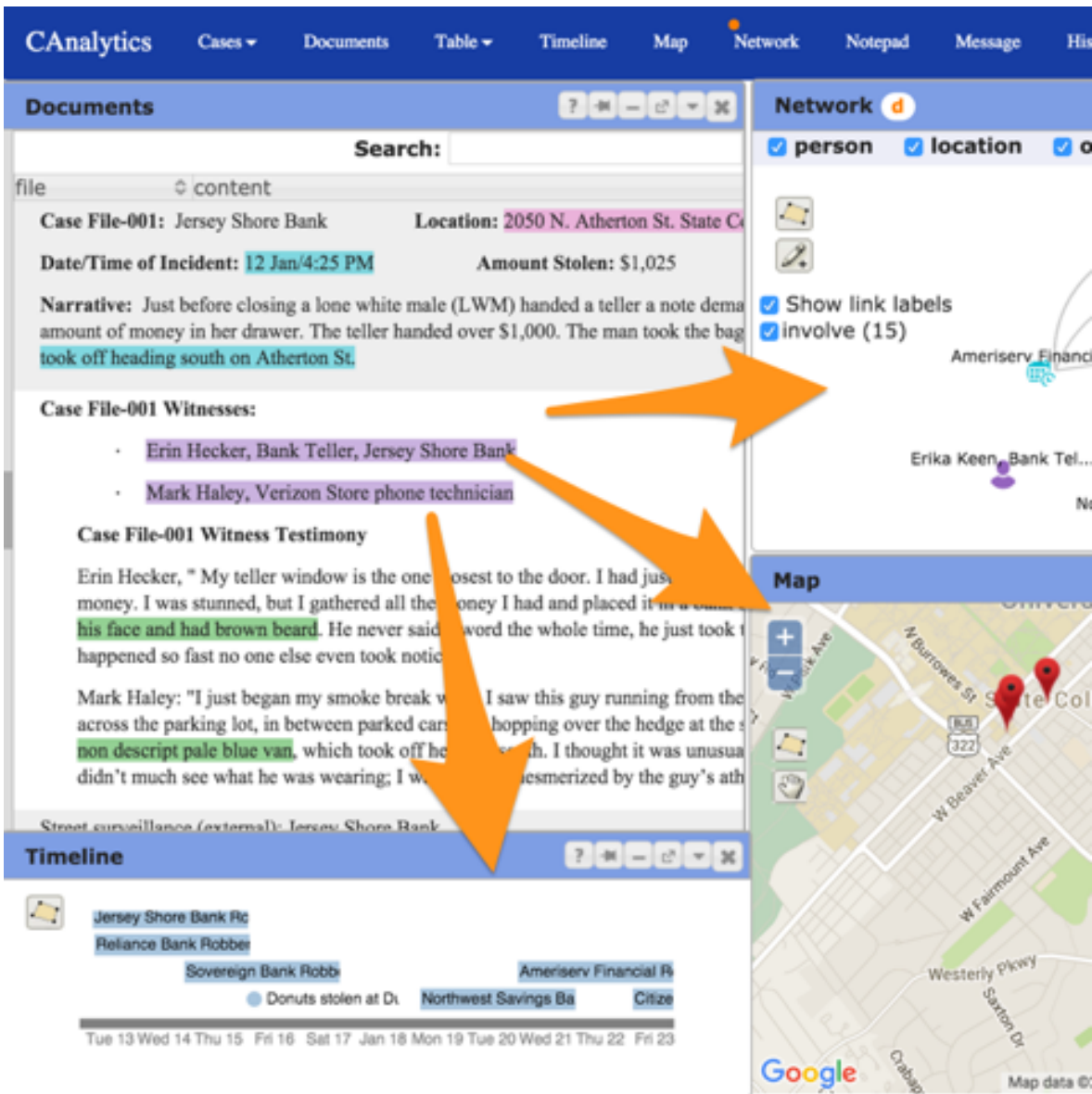


Fig. 1. CAnalytics user interface. Each window is closable, movable, and resizable. Shown here are *document* view (top left), *timeline* (bottom left), *network* (top right), *map* (bottom middle), and *message* (bottom right). Other windows include *table*, *notepad*, and *history* view.

view in a floating, closable window. Views are coordinated and share the underlying data pool so that products of different activities are not fragmented. The following paragraphs describe features of each view.

The *document* view supports evidence modeling through annotation. Annotation provides a basic structure to assist analysts in containing and framing new data. In the document view users

can select and highlight a snippet of text and annotate it as a type of entity such as a person, location, event, etc., or as a relationship between entities. Unlike other entity-based systems such as [?, ?], we use annotations to allow analysts to manually create evidence objects of interest. Manual annotation allows for greater user control in terms of information of interest and granularity that best suits their ad-hoc analytic needs. Users can add attributes to the annotated object, e.g. time in an event, and coordinate in a location. Users can also make reference to other objects in the attribute; for example, users can add people objects to an event indicating that these people were involved in the event.

These annotations turn text into structured data objects, which are then displayed in visualizations in the same workspace, including *table*, *timeline*, *map*, and *network* view—artifacts that are frequently constructed to hold attribute data, temporal data, spatial data and relational data respectively [?]. Figure 1 shows an example: when an annotation is created in the document view with information about time, location, participants, and their relationships, a new event is created in the timeline view, a new location is created in the map view, and new people are added to the network graph with a labeled edge representing the relationship (or new edges are added to existing nodes). Hovering the mouse over an entity will activate an entity detail window that displays attributes in detail, and analysts can modify, or re-model the entity in situ.

The views afford brushing and linking interactions; that is, when users brush entities in one view, related entities are displayed in other views. Thus the analyst can narrow down entities to their interest by: specifying a time range using the functionality provided on the timeline; making a spatial query with map filter; or selecting a cluster of entities by drawing a bounding area in the network view.

*Collaboration and awareness features.* To support collaboration, CAnalytics affords real-time collaborative editing, similar to the Google Tools. Users can open several concurrent editors to collaboratively edit multiple entity objects. Entities and annotations are immediately shared within a team and rendered in teammates' corresponding views.

In addition to real-time data sharing, CAnalytics supports other awareness features. A *notification system* sends individual's actions to the team in the form of a text box located in the top right corner of the workspace. An iconic indicator on top of a view window, which we call *tool coordinator*, shows who else is also working on this view. A *message* tool is a real time chat window that enables team communication with persistent message history. The system also maintains a traceable log of time-stamped individual activities in the *history* tool. Users can learn team activities about who did what to which object at what time. Entities mentioned in the message tool and history tool are hyperlinked and will trigger pop-up detail window when the user moves the mouse over them. With these awareness features, users who work synchronously can be informed of others' activity continuously; users who work asynchronously will be able to use the history to reconstruct their work status and become aware of changes beyond the point of their last interaction.

## 4 SETTING AND METHOD

The setting for this study was an undergraduate course in an intelligence training program in a US university. The program was designed to train students to become professional intelligence analysts. A key characteristic of the course is to emphasize hands-on practice on team-based intelligence analysis.

Of the 98 students enrolled in the course (from two sections), 73 consented to participate in the study. All of the 73 students held major in the program of Security and Risk Analysis. Most (75%) of them were in the third academic year ( $M = 3.06$  years,  $SD = .44$ ), indicating that participants in





Fig. 2. Classroom setting

our study had relatively advanced experience and knowledge in intelligence analysis. Participants' age ranged from 19 to 28 ( $M = 20.4$ ,  $SD = 1.18$ ). 77% of the participants were male.

During the first ten weeks of the course students learned several analytic techniques, including IEW, ACH, timeline analysis and network analysis, as well as state-of-the-art tools to implement these techniques, including Analyst's Notebook and PARC ACH. Students also practiced applying these techniques in two hands-on projects before our study. A typical workflow started with IEW to extract and model evidence from documents, followed by building analytic artifacts such as an ACH Matrix in PARC ACH, and a timeline and a network graph in Analyst's Notebook. Since data from IEW table could not be shared or extended to other tools, students had to manually replicate data for each different tool. Most tools they used lacked serious collaboration support (except that some teams used Google Doc to construct an IEW table). Analysts were unable to contribute simultaneously (an issue known as production blocking [?]). Students divided their work by tools: each person picked a tool and created and analyzed an artifact with the tool on their own. This had the consequence that findings and hypotheses be made without integrating collective efforts and diverse knowledge. Analysts coordinated work by manually sharing documents or graphs through email or cloud storage service (e.g. Dropbox), resulting in a scattered placement of results, requiring repeated manual resynchronizing to identify redundant or missing pieces of information, analysis of information, and analytic hypotheses. In summary, the students in our study had learned and practiced with tools lack of support for collaboration, and were aware of the shortcomings.

Students were given a tutorial on CAnalytics a week before the project began. One of the authors walked through the features of CAnalytics, without enforcing or implying any strategic use of the tool. Students then accomplished a small case analysis at their own pace. Although students were encouraged to make full use of CAnalytics, to ensure a naturalist environment students were always free to employ any other tools that they believed useful.

Our study began on the 10th week of the course and lasted one week. The analysis that students performed on our tool was the investigation of seven bank robberies fabricated by the course instructor. Teams were provided a set of documents pertaining to the robberies, including police reports, witnesses reports, video records, and news media. The analysis was designed to be open-ended, meaning that there was no single, definitive answer. The instructor explained that the task was to simulate real world scenarios, in which analysts always reasoned in the circumstances of uncertainty, ambiguity, and complexity. The instructor told the students that a total of 6 hours of

workload was expected, including in-class and outside-class work. In the end of the project students were required to submit a team report, describing their hypotheses, assumptions, conclusions and supporting evidence. Students were graded on their ability to understand and enact the professional practices of intelligence analysis. This strong normative emphasis on problem solving practices provides us an appropriate evaluation context for new interactive tools: Tools are only valuable to the students insofar as they actually support better practices and better outcomes.

Students were randomly assigned into 25 teams (23 three-person teams and 2 two-person teams). To avoid the effect of group size on team behavior analysis, we excluded the two-person teams from our analysis in this paper. When in class, students used a 27-in Macintosh desktop. Students could use any equipment when outside class.

We employed a number of data collection approaches as suggested by prior researches [?, ?]. We administrated a post-study questionnaire. The questions used 7-point likert scale and included nine items measuring individual's self-reported awareness, seven items for communication quality, six items for collective efficacy, and three items for perceived performance [?]. We also used NASA-TLX [?] to measure cognitive load. The end of the questionnaire included open-ended questions asking how the tool helped or impeded their work. We captured user interactions with system logs. Instead of simply logging low-level events like mouse click and keyboard strokes, we recorded actions such as creating an annotation and deleting an entity. Finally, we reviewed team reports and graded them as an indicator of team performance. Since the task was open-ended, there was no single right answer. We constructed an assessment rubric together with the course instructor by listing all possible hypotheses and evidence from the documents, with a full score of 16. The first author and a research assistant graded the reports independently. If the grades differ by less than 2, an average is set as the final grade (14 out of 22 reports). Otherwise (the rest 8 reports), the two graders review the reports together and make an agreement.

One limitation in our classroom study is that we were unable to conduct control group comparisons. Ethically it is difficult to assign students to education conditions that may be disadvantaged. Indeed, the instructor we worked with wanted all of his students to experience the same educational opportunities. This is a direct conflict between our interest in using the classroom context as a larger-scale testbed CAnalytics, and the students/instructor interest in experiencing and learning about the effect of technology on collaborative intelligence analysis. The classroom is surely a special case of the "real world", but it is the real world relative to a lab study context. We often cannot run control conditions in workplaces.

## 5 RESULTS

While participants could access the tool any time, teams had three most intensive usage sessions over the week: two were in class and one was outside class before the team report deadline. 22 teams self disclosed that they used CAnalytics as the principle analytic tool throughout the project although they were free to use others; one team reported that they mostly used Google Doc because they "*felt more used to*". Nine other teams reported having used Google Doc, but only for composing the final team report. Seven teams reported using GroupMe and other instant message outside class. They used these tools for instant communication and coordination of meeting. Thus the CAnalytics system logs captured most of team's analytic sessions, although their communication data could be partially missing.

Over the week, teams created 1805 entities and 1529 relationships in total. The number of entities teams created ranged from 24 to 223 ( $M = 82$ ,  $SD = 39.9$ ), and the number of relationships ranged from 7 to 237 ( $M = 69.5$ ,  $SD = 51.0$ ), showing a large variety. The big range was related to different team data modeling strategy, which will be discussed in detail later.



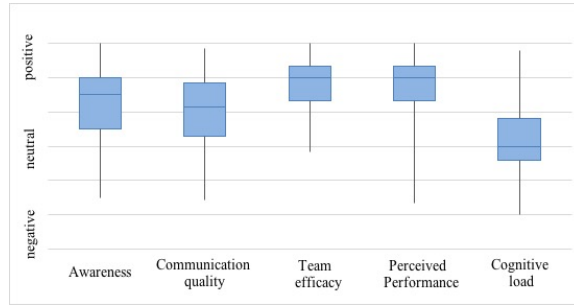


Fig. 3. Survey responses (box shows Q1-Q3 and median; ends of whiskers show maximum and minimum)

CAnalytics was generally well received by the students. An overview of the related survey items (shown in Figure 3) shows that students positively rated all aspects of the tool except cognitive load, towards which they had a close to neutral feeling.

### 5.1 Interleaving data modeling and analysis

We examined the pattern of data modeling and analysis first by qualitatively looking at a visualization of the entire interaction log (e.g. Figure 4a shows one team's interaction). All teams worked intensively on data modeling as they started off the project. This was the phase when teams were getting themselves familiar with the documents and made initial annotation input into CAnalytics. Starting from certain time point, all teams started analysis on visualizations, followed by frequent transitions between data modeling and analysis. 11 teams started data analysis in the first usage session, while the other 11 teams had this transition in the second usage session. In average, the transitions occurred in 47.6 minutes after the project began. The earliest transition occurred in 14 minutes after the team started the project, and the last team had the transition around 104 minutes, later in the second session. We also found performance difference among teams that started analysis early and those late. Teams that started analysis in session one had higher performance ( $M = 8.6$ ) than teams that started from session two ( $M = 6.7$ ), although the difference was not statistically significant.

The fact that participants returned to making annotations after analysis indicated that they did not wait to start analysis till they had finished modeling. Indeed, the activity of data modeling and data analysis were highly interleaved throughout the project (as shown by the interleaving color bar in Figure 4a). Participants switched from one activity to the other activity frequently. The state transition diagram (Figure 4b) demonstrates the interleaving in an aggregated way, in which we encode the number of transitions as width of the link. This result confirms our design expectation that data modeling and analysis should not be supported as separate staged activities, and that an integrated environment should streamline the workflow.

After confirming the existence of interleaving workflow, we explored further what drives the switching between data modeling and analysis. We looked into team behavior in modeling and analysis respectively through log analysis and artifact analysis, and describe the results below.

### 5.2 Data modeling: filtering vs. accretion

We noted a distinction between filtering and accretion strategies in data modeling, similar to what was reported in a paper prototype study [?]. Filtering is selectively modeling of data and adding to an artifact. Users must decide what information is relevant, and thus what is to be excluded, as well as what granularity of information is to model. Filtering requires more team coordination, because

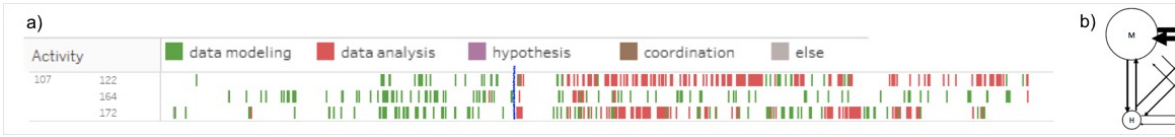


Fig. 4. (a) Visualization of interaction logs of Team 107. Each row of colored marks indicates the sequence of top level activities a participant performed. (b) State transition diagram of interaction logs of Team 107. Each node is an activity, whose size represents the time spent on the it (M: data modeling; A: data analysis; H: hypothesis development; C: coordination); a link represents a switch from one activity to another, whose width encodes the number of switches. We see highly frequent transitions between data modeling and data analysis

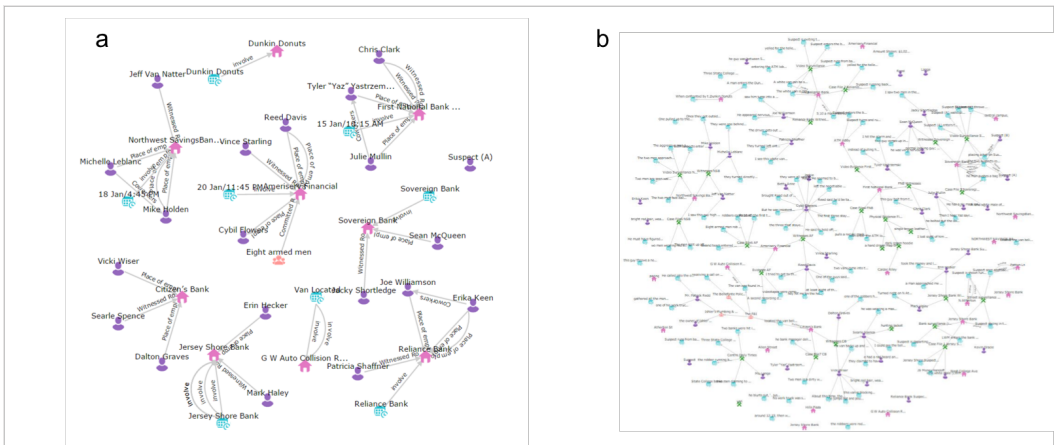


Fig. 5. Network artifact comparison: filtering (a) vs. accretion (b)

teammates must reach a common ground of the current problem as well as information needed to answer the problem. Figure 5a is an example of network built with filtering strategy. It only represented key information of robberies and thus provides a good overview of the whole story.

Accretion is an attempt to comprehensively represent the problem by adding information of all levels of details to an artifact. Users extract every fact from the document, regardless of its immediate relevance to the problem. Accretion requires less coordination as it is relatively mechanical note taking. A disadvantage of accretion is that the produced artifact could be fairly complex. An example is Team 108, who modeled every step the suspects took in a robbery. Such artifact provides a detailed view of the data, and can be useful when a step-by-step comparison is needed to identify robbery patterns. Yet it also resulted in far more entities (223) than the average (82) and more cluttered network view (Figure 5b). This accounted for the large range of entities created across teams mentioned beforehand. Participants also realized the problem. They reflected that they spent too much time in detail events, and many did not help their analysis at all:

*I felt that after we were done annotating, we hadn't really accomplished anything and that we were no closer to solving the case than when we had started. In the end it didn't really help that we had annotated the data. (P86)*

Filtering and accretion strategies are not exclusive; in fact many teams applied a mix of two strategies. Depending on the need of analysis, filtering provides an overview of the whole story,

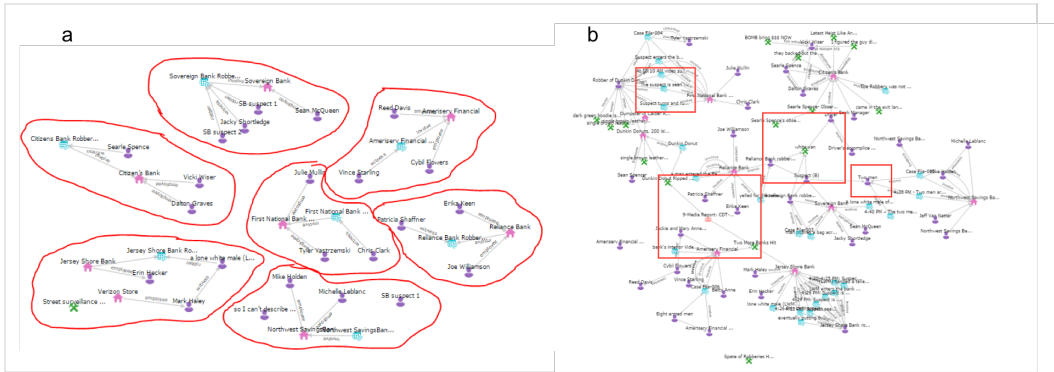


Fig. 6. Network artifact comparison: separate clusters (a) vs. connected clusters (b). The parts highlighted in red squares in (b) are key evidence that connects clusters

and accretion helps investigate a specific robbery in detail. Modeling data at different levels opens different analytic threads, which accommodates multilevel analysis, including examining attributes of an individual entity, comparing two related events, and developing a hypothesized story on the whole dataset.

### 5.3 Artifact analysis: fact vs. inference

We examined the analytic artifacts teams created, the network graph in particular because social relationships played the most critical role in this specific scenario and teams spent most time on network analysis (as reflected from the log). We found a clear distinction among the network artifacts. For example, networks from 8 teams (mean performance score = 7.8) consist of separate clusters (Figure 6a). Nodes within a cluster are connected, representing information space of a robbery case; Nodes between clusters are nonetheless not connected, indicating each robbery is a self-contained case. However, these teams did not miss the connections between robberies, as these teams still discussed these robbery connections in their report. It turned out that these teams documented any possible relationships between robberies in the notepad tool as a list, separate from the network graph; that is, these teams distinguished information content and synthesized them in different artifacts.

In contrast, 6 other teams (mean performance score = 8.3) created networks composed of connected clusters. While a cluster is still a representation of a robbery, some of them are connected through an evidence node. An example is Figure 6b, in which we mark in red four *connectors* that link the clusters. These connectors were key evidence that led the teams to hypothesize that those robberies were related and might be committed by the same criminal group (e.g. the white van shown up in two robberies). These teams represented all information in one artifact.

By comparing these two types of networks, we found that links within a cluster were typically *factual* relationships modeled from raw documents (e.g. a white van was witnessed at a location), and links between clusters were often *inferences* beyond literally documented (e.g. a white van at location A is the same van witnessed at location B). Teams creating separate clusters represented only facts in the network and held evidence with uncertainty in a separate artifact. One advantage of distinguishing facts and inferences is that teams can always be aware of assumptions made when making a conclusion. And since all inferences are held in one place, teams are forced to confront them and review their uncertainty iteratively in the process. However, the strategy also

Element	Example
Social awareness <i>who is present?</i>	CAalytics helped me stay aware,of my teammates activities because I could see who was lo top,right corner (P123)
History awareness <i>Who has done what?</i>	The way you are able to view when and where your teammate made or updated annotations/i the key to staying aware of what your team has done. It is a great tool in respects to that. For able to view the changes my team made while I was not using the CAalytics tool at the same using the history tab. (P171)
Information awareness <i>What is being changed?</i>	CAalytics was very helpful in keeping us updated on what was being changed/noted/amende when. This was very beneficial for staying on the same page and knowing what changes were no one individual was out of the loop. (P157)
Action awareness <i>Who is doing what?</i>	I liked how you could always see what your teammate were viewing on the website. For exampl on the bluf when my teammates were working on the network part of the program. If I were t piece of information that I thought might be helpful to them I would just tell them. My tea same thing in return. (P51)
Intention awareness <i>Who is going to do what?</i>	CAalytics showed what tab [tool] my teammates were working on which helped me be awa were working on. For example, if I saw that one of my teammates was on the network tab, I were attempting to connect the information that was relevant to one another.,I would then be any new findings I had that could influence their work (P160)

Table 1. Subject feedback of awareness instances

adds difficulty to analysis as analysts may overlook or fail to combine evidence scattered in different artifacts.

On the contrary, in connected-cluster networks, facts and inferences overlaid in one artifact together drive the layout of the network, are better synthesized, and give analysts a clearer picture in one place. Teams may discuss and evaluate the level of uncertainty of inferences to decide whether to add them to the network. This strategy makes analysis more interactive among teammates: they need to negotiate, evaluate, and reach consensus on the value and validity of inferences. However, a problem with mixing facts and inferences is, to some extent teams might forget whether a link is factual or inferred, and ask whether conclusion derived from the visualization can be trusted under uncertainty.

Uncertainty in analysis represents a gap between analyst's knowledge and data needed to address the problem. The existence of uncertainty drives analysts to switch to data modeling and to collect more data or remodel existing evidence in order to resolve the uncertainty.

#### 5.4 Collaboration and awareness

One recurring theme in the subject feedback we collected was that the collaboration features were helpful for solving the problem. In the survey 88% of the students positively rated their group awareness. Participants appreciated that the tool complemented traditional analytic tools, describing CAalytics as Analyst's Notebook with real time synchronization features similarly to Google Docs, or described it as Google Doc with added visual analytics. To quote one participant, "CAalytics is like an analysts notebook that multiple people could work on at once [... and] an analysts version of a Google Doc" (P65).

Participants reflected that they could now contribute simultaneously without concerns of interference and could have everything in one place instead of manually sharing documents via a cloud service.

*It was much easier to coordinate as a team with CAnalytics because we could all work on the same system at the same time. Without CAnalytics, we were forced to do the work separately and compile all the work onto one system after we had finished. (P156)*

Students also reported that being able to see teammate's status made the task more motivating and engaging:

*During class I wasn't sure if my teammates were doing work for that class or another thing but then seeing their dot [tool indicator] switch between applications on the software and updates pop up on my screen I knew they were doing work for 231. (P141)*

*The fact that you can see what other teammates are doing and they can see what you are doing creates a sense of accountability in terms of separating the work load. (P51)*

The motivating effect of awareness might account for, at least partially, the fact that teams were participating equally. We measured the equality of participation in terms of number of created entities and time spent on CAnalytics. We refer to Olson [?] in calculating equality: one minus the variance of proportions times 100 (for better readability). Thus the score ranges from .67 (when only one person contributes in a three-member team) to 1.00 (when everyone participated exactly equally), and higher score indicates higher balanced participation. The resulted equality of created entities and time was .96 and .99 in average respectively. This indicates participants contributed fairly evenly.

Another repeated theme was the awareness features helped assure all teammates were executing the team plan. Participants reflected on their experience that a common team breakdown was misunderstanding of the team plan, and that they did not realize the misunderstanding until everyone had spent significant efforts finishing their "perceived" job. CAnalytics made plan execution assured because they could always see where teammates were working and what they were generating; and if anything unexpected happened, they could communicate immediately rather than in the end of the project.

Participants reported many other instances of awareness they realized using CAnalytics. We categorized them based on the element of awareness, or the essential problem of awareness of *what* [?], into social awareness, information awareness, action awareness, history awareness, and intention awareness, as shown in Table 1.

When asked what features helped them stay aware of team activities, 28 participants mentioned the tool coordinator, 24 mentioned the notification system, 19 mentioned the history tool, 14 mentioned the real-time update of user-generated data, 12 mentioned the collaborative editor, and 7 mentioned the message tool. Although the number of mentions does not simply indicate tool usefulness, it suggests users were explicitly aware of these awareness features and appreciated their support.

Students' positive feedback on awareness was further corroborated by interaction logs. For example, we measured the number of entities accessed by other collaborators. While data generated by users is automatically shared, it is up to collaborators whether to read/edit the shared information or ignore information altogether. The log showed that on average, 77.6% of the created entities were *read* by at least one other teammate. Further, We measured how many entities were *edited* by collaborators, a phenomenon we argue requires higher awareness, because the collaborator must not only realize the creation of the entity, but also understand its content. We defined *peer editing*, manipulated as the ratio of editing other's entities over editing those created by oneself. We found that all teams edited collaborator's entity objects, with a peer editing value equal to .83 ( $SD = .45$ ). The result suggests that teams had little difficulty accessing and modifying partner's created data objects.

One major critique is the lack of support for sharing intermediate analytic insights. An insight is revealed and contextualized by a specific arrangement of views, e.g. a reduced data view of interest through filter, a highlighted entity representing the analytic focus, and a clustering layout of network to demonstrate a specific relationship. While teammates share the same data pool, they are likely to have different views of data, and thus different *interpretations* toward the data. A dynamic view together with its interpretation represents user's intermediate analytic status. Sharing these insights could inspire team analysis [?]. With CAnalytics participants complaint that they could not easily make such communications. The team could “*be looking at the same information but arranged in completely different ways*” (P131).

### 5.5 Labor division strategies

We noted different labor division strategies from interaction logs. Seven teams followed *document-based collaboration* (DBC): they divided their work by evenly distributing the different documents among team members (as shown in Figure 7a). Each member read through and made annotations on their own set of documents. An advantage of this strategy is that individuals get less workload and thus have more time to think deeply about their own documents. Individuals also implicitly take over the responsibility of their assigned documents to gain insights and share them when the team synthesize findings in later analysis. When the team needs information from one document, they rely on the “document owner” to share his/her finding. The failing of this strategy is thus in case an individual fails to identify or convey evidence in the document, the team may overlook the information altogether [?].

Four teams followed an *entity-based collaboration* (EBC) strategy. Instead of dividing by documents, they divided work by entity types: each individual went through all documents but only annotated entities of certain types, e.g. teammate A only annotated persons and teammate B annotated locations (as shown in Figure 7b). This strategy seems to save teammate's time in the data modeling phase. And since each person focuses only on certain entities, they are more likely to identify recurring patterns, for example, the white van used in multiple robberies. However, focusing on certain entities could lead individuals to superficial syntactic scanning of documents instead of deep reading. This could further lead to extremes of annotating all entities of the type, whether they are related to the problem at hand or not, similar issues with accretion strategy discussed beforehand. Indeed we found from the interaction log that EBC teams created more entities ( $M = 101$ ) in average than DBC teams ( $M = 94.5$ ). Moreover, with emphasis on certain entities, individuals are likely to know only partial aspects of a robbery and hence have difficulty connecting and synthesizing facts to deduce any conclusion. As one participant reflected, “*we broke up by entity type, which reduced our individual involvement in each other's entity types*” (P99). The result indicated that the average performance of EBC teams ( $M = 7.63$ ) was lower than that of DBC teams ( $M = 9.25$ ), although the difference was not statistically significant.

The rest (eleven) of the teams did not show specific labor division patterns. Indeed, teams did not necessarily have to divide their work in order to collaborate, especially when the collaborative tool provides possibility to work closely together. Teammates could read and annotate the same document because they could see new annotations by others in real time and build on other's annotation. Figure 8 shows an example where one team worked on the same document simultaneously. The three team members exhibited high synchronicity in which document to analyze, which we believe was not by accident. This has the advantage that teammates are always on the same page and can discuss hypotheses throughout the analysis process. Participants did have concern for possible duplication. As one participant complaint, “*we could not actively see the changes our teammates were making until well after they had made them*” (P46). This was because an annotation

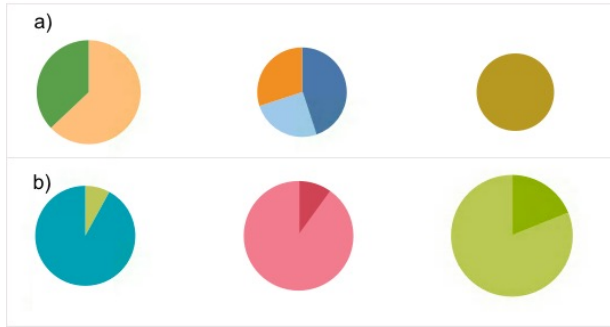


Fig. 7. Pie charts showing different labor division strategies. Each pie chart corresponds to one team member. (a) Document-based collaboration. The pie chart shows the documents one team member annotated, color coded by document ID. (b) Entity-based collaboration. The pie chart shows the entity types one team member created, color coded by entity types.



Fig. 8. Graph showing the timeline of one team creating annotations. Each row corresponds to one team member. Each bar represents an annotation, color coded by document ID. The red blocks highlights the periods when all teammates worked on the same documents simultaneously.

was shared only *after* it was created, yet another teammate might already be drafting an annotation on the same text snippet in the meantime.

## 5.6 Interaction with team performance

To systematically evaluate factors that influenced team performance, we conducted a multiple linear regression between team performance as the dependent variable and team collaboration characteristics (i.e. equality of entities and time respectively, peer editing, number of entities created, shared entities, switch time from data modeling to analysis) as independent variables. We treated a team as a system [?] and modeled team-level relationships. The relationship of individual level variables (e.g. survey items) could not be simulated with regression directly because data within a team was interdependent. The analysis was performed using *R* [?].

The result is shown in Figure 9. The regression model was found to be statistically significant ( $F(5, 16) = 4.58; p < .01$ ), with an *R* squared equal to .59, meaning 59% of the variability in team performance was explained by the model. Five team-level variables contributed significantly to prediction of team performance. The model suggested that balanced contribution of entities predicted higher team performance scores ( $\beta = 45.79, p < .05$ ), but balanced participation time did not show such effect. More peer editing led to better performance ( $\beta = 3.89, p < .01$ ), implying that analysts should be encouraged to model data as a team, and to remodel other's entity objects as needed. Longer elapsed time before a team started analysis (*switch time* in the figure) predicted lower performance ( $\beta = -.05, p < .05$ ), which suggested that teams should shorten the time of pure data modeling and start analysis earlier. Larger number of entities teams created also predicted higher performance ( $\beta = .03, p < .05$ ). This can be interpreted that teams providing sufficient model input did benefit from system support. However, we are not ready to claim more entities are always



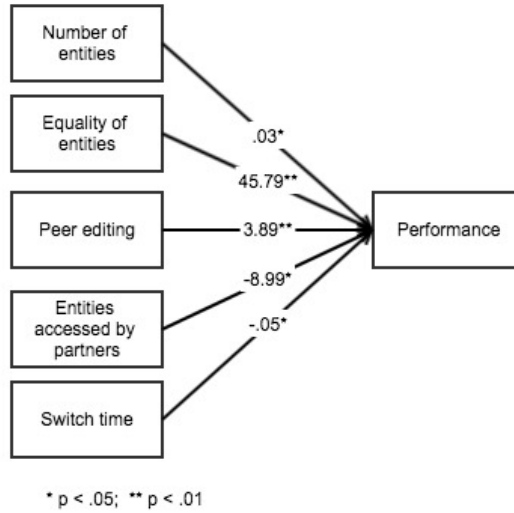


Fig. 9. The relationship between collaboration characteristics and team performance.

better because we should also be cautious that entities irrelevant to the team problem bear no value to team analysis, as discussed beforehand. Somewhat surprisingly, the number of shared entities negatively predicted performance ( $\beta = -8.99, p < .05$ ). We considered this might also relate to the issue of entity value: sharing entities without relevance to the team problem leads to reduced team efficiency and creates distraction, which in turn would lead to worse team performance. Another possibility is that teammates may not necessarily access all entities partners created. The theory of Transactive Memory [?] suggests that teammates do not have to know everything in the team but should know “who knows what”. Individuals specialize in their own field (e.g. one specializes in event analysis while another in social relationship examination) and reach for collaborators when other knowledge is needed. In such case, common knowledge in shared entities would actually be redundant. Individuals do not need to access content of entities created by others, but only keep aware of the creator.

## 6 DISCUSSION

The goal of the study is to explore design opportunities to support collaboration in information analysis by evaluating tool usage in a natural environment over multiple usage sessions. Our work builds upon prior empirical studies (e.g. [?, ?, ?, ?]) and embodies their design implications in our tool. The study also complements research that only tests tools in short term lab studies (e.g. [?, ?, ?]) by investigating tool usage over multiple usage sessions.

This study adopts an *analyst-centered design* approach. A critical requirement of developing tools that meet user needs is to understand their needs and practices. When these needs and practices are specialized (as is the case in intelligence analysis), it is particularly important to include the target user population in the design process [?]. Our classroom study provided an opportunity with deep access to analysts in training. These analysts have been trained with knowledge and skills of intelligence analysis, and have experience with state-of-the-art analytic tools such as Analyst’s

Notebook and PARC ACH. In their reflections, participants often compared CANalytics to those tools. Their multi-session usage of CANalytics also allows them to adapt to the tool and learn to appropriate it to the best of their team purpose [?]. Therefore, while their feedback is admittedly not the same as an experienced professional, their feedback does provide a deeper insight into strengths and weaknesses of CANalytics.

The study provides encouraging results on supporting collaboration in intelligence analysis. Participants appreciated an all-in-one environment where they could share raw documents, evidence snippets, views of evidence and hypotheses in one place. They liked the fact that they could contribute simultaneously without blocking or interfering each other. Another benefit of the collaborative tool is to keep teammates aware of each other's activities. Staying aware of teammates not only helps establish a common ground for planning team strategy, but also ensure everyone is following the plan as expected. Moreover, results suggest the awareness features provide positive *social facilitation* [?]: individuals found the task motivating and engaging with awareness of each other's activity. We also measured collaboration characteristics that impacted team performance, and found that balanced contribution, peer editing, and earlier switching from modeling to analysis predicted higher team performance. Most importantly, we documented the interleaving workflow enabled by the integrated environment, and explored momentum in modeling and analysis behaviors that drove the activity switching. Below we discuss design implications that could potentially enable a better collaborative experience in information analysis tasks.

*Scaffold a structured interleaving workflow.* A misconception about information analysis is that data modeling and data analysis are two staged activities. This is akin to the *waterfall* software development model, which features a sequential process that flows downwards through the phases of requirement conception, software design and implementation, and testing and maintenance. Critics have argued against this approach and instead called for an iterative design process that leads to reframe user requirements, redesign, redevelopment, and retesting.

Our result demonstrates a similar iterative and dynamic process in intelligence analysis. The result is striking especially because our participants have been trained with tools that impose a waterfall model. They could have followed their old waterfall practice with our tool, yet instead all teams spontaneously switched to an iterative manner. We projected that modeling on multiple granularities drives analysis on different levels, and uncertainty in analysis pushes analysts back to collecting additional data.

Realizing that, we probably could shape analysts into a more interleaving workflow with a more structured approach. Structured techniques help users perform analysis in a systematic and transparent manner. For example, IEW aims to structure evidence modeling, and ACH proceduralizes evaluation of analytic hypotheses. Yet the way to switch between the two activities is loosely defined [?]. Our system implies a structured modeling through annotation and a structured analysis by visualizing data in multiple views, and by sharing the data structure we enable a smooth switching between the two stages. Our result implies the role of multilevel modeling and analysis uncertainty in driving the switch. Based on that, we could build a scaffolding process to assist analysts in connecting data and analysis to enable a more interleaving flow. For example, when user adds a new data object, the system could suggest possible connections to existing evidence in the context of an appropriate level of views, which is likely to help analysts discover new patterns. When a user creates a new hypothesis with uncertainty, the system could highlight associated evidence, which would prompt the analyst to re-examine the data and look for more data. Such scaffolding provides a basic structure to link stages of analytic activities that analysts can take on without imposing a specific fixed workflow.

A smoothier interleaving workflow could also be made by increasing team awareness of partner's modeling and analysis. That is, uncertainty in one's analysis not only motivates oneself to data modeling, but also drives partners (who is aware of the what and why of the uncertainty) to collect more data. Similarly, one's modeling could influence and drive partner's analysis, given the partner is fully aware of what is modeled and how the new data connects to the level of existing data. Such "team-level" interleaving could make teamwork more interactive and close coupling, but also requires support of higher awareness, especially of hypothesis uncertainty and data model context. Improved design for multi-granularity modeling, uncertainty representation, as well as team awareness of these features opens up possibilities for coordinating interleaving at the team level. We discuss these design implications in further detail in the following paragraphs.

*Build collapsible views for multi-level modeling.* We observed that analysts built data models in multiple granularities and engaged and coordinated among different levels of details. For example, analysts jumped quickly from digging into details of a single event, to comparing between two events, and to overviewing all robberies as a complete story. When sharing these data, a critical requirement is to represent them in an appropriate context in order to ensure teammates understand them correctly. A collapsible data view could be a solution to accommodate such multi-level modeling. This can help draw teammate's attention to a specific item while keeping the global context available. Analysts can focus on a certain level of detail at a time while conveniently switching between levels. A collapsible view also reduces the problem of cluttered view when data volume increases. and when analysts dig into greater details (e.g. representing suspect's all actions to identify patterns of common actions in two robberies). An analyst can overview all robberies and only unfold detailed actions when investigating into a specific robbery.

*Design a richer graphic language for uncertainty.* Uncertainty is prevalent in analysis and more prevalent in collaborative analysis because each teammate becomes a source of uncertainty when they contribute a piece of information [?]. Teams in our study spontaneously employed two different approaches to deal with uncertainty: either to mix them for better synthesis or to separate them for better clarification. Representing uncertainty appropriately assures teammates when they build analysis upon other's insight. This helps reduce the problem emphasized in Chin et al.'s work [?] that analysts do not trust partner's conclusion. It also helps teammates stay aware of issues that need to be addressed, thus decide what extra data is needed. We propose that a richer graphic language and interaction be designed so that analysts can encode uncertainty into the views. For example, links and entities with different uncertainty can be visualized in different transparency. Users can *filter* by uncertainty so that users can choose to consider evidence only with high credibility, or to review all inferences and decide what extra data is needed to consolidate them.

*Share views as team resource.* View sharing is important for sharing and understanding analytic result [?]. A common solution is to enable a public view which always keeps synchronized for all teammates [?, ?]. However, a single public view does not meet the need in intelligence analysis because analysts may want to share multiple views in an iterative analysis process. We propose that views should be treated as a *team resource*, just like data, which is sharable, extensible, and reusable. For example, an analyst can save their current view as a shared resource when they feel it useful to collaborators. Other people can reuse the view to their need. Shared views should be interactive rather than static images, so that analysts can perform all interactions including filtering and highlighting, and are able to evolve the view with collective team efforts, a critical requirement emphasized in [?].

*Distinguish visible vs. valuable contributions.* Finally, we noted cases where teams created far more entities than needed with an accretion strategy. Strikingly, while similar data modeling strategy was reported in the paper prototype study [?], users with CAnalytics seemed far more tempted to accretively add information, with far more entities and cluttered views. For example, the extreme team created as many as three times entities than the rest teams in our study, much more than the difference in the paper prototype study. Why did this happen? We guess both the context of classroom study and the system design contributed. Unlike in the lab study where teams are temporarily assembled, teams in a class evaluate peers either consciously or unconsciously and value how themselves are being evaluated. Such social pressure motivate individuals to make contributions, and indeed to make *visible* contributions, more than *valuable* contributions. That is, participants noticed that their work activity was visible to their partners, and accordingly prioritized doing more visible work over doing less visible work. In some cases, this led to a new problem of easy and less valuable contributions that were highly visible - such as creating and therefore sharing data entities that were not particularly important, and subsequently made data models seem cluttered. For example, creating and therefore sharing an entity gets immediately notified to the team whereas weighing the importance and relevance of an entity goes silent in the system. We need to investigate approaches to making significant contributions more visible, or perhaps making it more immediately visible that less important contributions are indeed less important.

## 7 CONCLUSION

This study investigates the feasibility, effectiveness and consequence of supporting collaborative information analysis in the context of classroom study. Our study provides encouraging results on design of an integrated collaborative workspace. Analysts spontaneously employed an interleaving workflow and making use of awareness features provided by the tool to stay team awareness. We also identify new opportunities in awareness support; awareness includes not only team actions, but information uncertainty, contribution visibility, and context for insight sharing. We propose design implications to enable a more interleaving workflow and possibly at a team level.

While this study was performed with an emphasis on intelligence analysis, we argue that CAnalytics can also be utilized for a broader domain of collaborative information analysis, wherein a team examines a complex information space of facts and relationships. As collaborative information analysis is increasingly a typical and chronic task, it is important for research to examine, understand, and provide effective tools and environments for these long-term, real-world CSCW interactions. This requires situating research in more complex work activity contexts, and directly investigating interactions, experiences, and outcomes in those contexts.

To be added

Received February 2007; revised March 2009; accepted June 2009