Information Visualization (2008) 7, 49 – 62

www.palgrave-journals.com/ivs

Design considerations for collaborative visual analytics

Jeffrey Heer[1]
Maneesh Agrawala[1]

[1]*University of California, Berkeley, CA, U.S.A.*

Correspondence:
Jeffrey Heer, Soda Hall # 1776, Computer Science Division, EECS Department, University of California, Berkeley, CA 94720-1776, U.S.A.
Fax: +1 510 642 5775
E-mail: jheer@cs.berkeley.edu

Abstract

Visualizations leverage the human visual system to support the process of sensemaking, in which information is collected, organized, and analyzed to generate knowledge and inform action. Although most research to date assumes a single-user focus on perceptual and cognitive processes, in practice, sensemaking is often a social process involving parallelization of effort, discussion, and consensus building. Thus, to fully support sensemaking, interactive visualization should also support social interaction. However, the most appropriate collaboration mechanisms for supporting this interaction are not immediately clear. In this article, we present design considerations for asynchronous collaboration in visual analysis environments, highlighting issues of work parallelization, communication, and social organization. These considerations provide a guide for the design and evaluation of collaborative visualization systems.
Information Visualization (2008) **7,** 49 – 62. doi:10.1057/palgrave.ivs.9500167

Keywords: Visualization; visual analytics; sensemaking; social data analysis; collaboration

Introduction

Visualizations leverage the human visual system to support the analysis of large amounts of information. Although most visualization research to date focuses on the interaction between a single user and an interactive display, visual analysis is rarely a solitary activity. Analysts must share and communicate their findings. They may disagree on how to interpret data and contribute contextual knowledge that deepens understanding. As participants build consensus or make decisions, they learn from their peers. Furthermore, some data sets are so large that thorough exploration by a single person is unlikely. Such scenarios regularly arise in business intelligence,[1] intelligence analysis,[2,3] and public data consumption.[4] Consequently, the design of visual analysis technologies could benefit by considering social interaction in addition to perceptual and cognitive processes. In this spirit, a recent report[3] names the design of collaborative visualization tools as a grand challenge for visualization research.

The social aspects of visualization have taken on new importance with the rise of the Internet, enabling collaboration between participants acting in different geographic locations and at different times. This distributed, asynchronous style of collaboration introduces new challenges for visualization research. Most existing research on collaborative visualization has focused on synchronous scenarios: users working together at the same time to analyze scientific results or discuss the state of a battlefield. Collocated collaboration usually involves shared displays, including large wall-sized screens and table-top

Received: 1 December 2007
Accepted: 5 January 2008
Online publication date: 21 February 2008

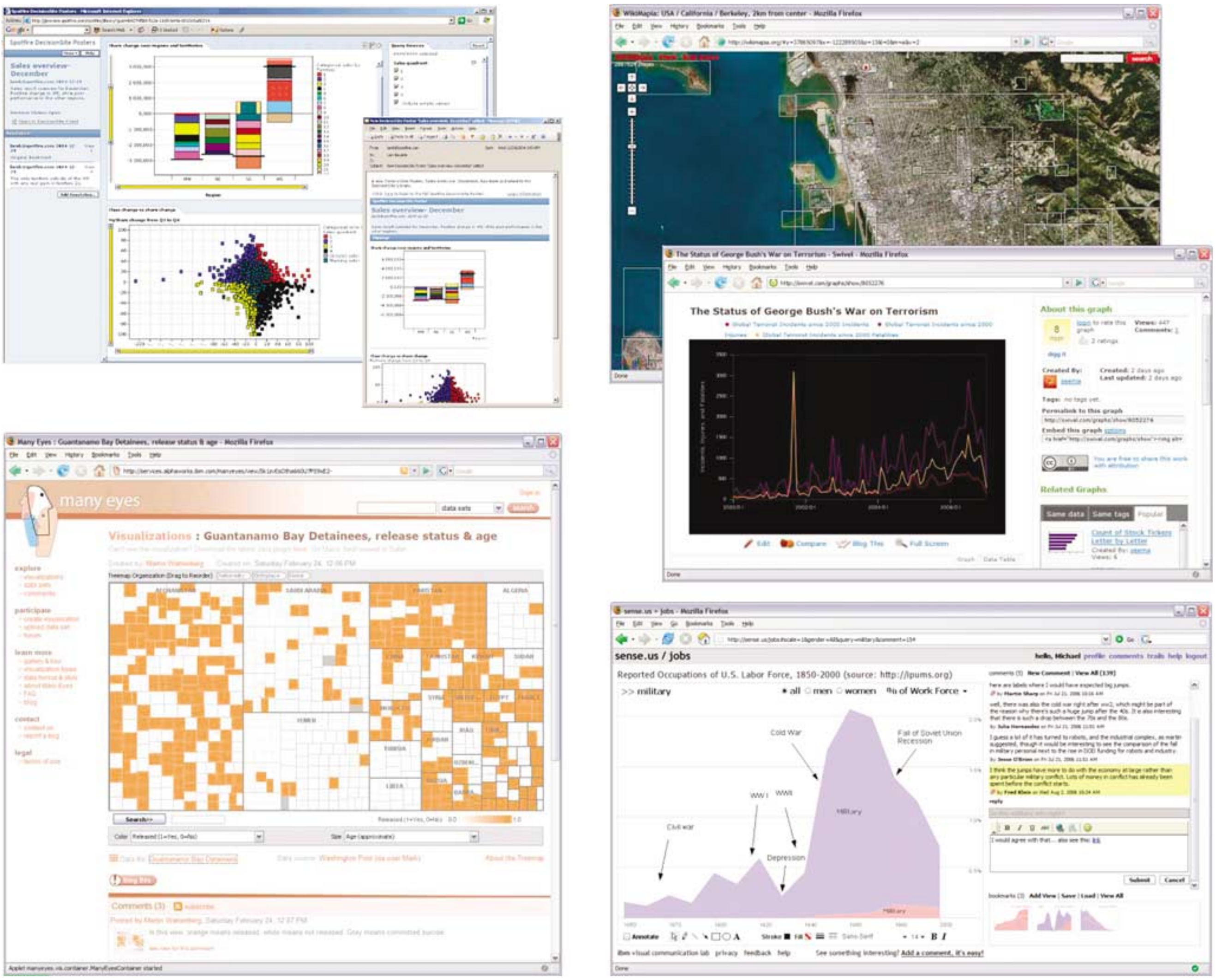

Figure 1 Asynchronous collaborative visualization systems. Clockwise from top-left, Spotfire Decision Site Posters,[16] Wikimapia,[19] Swivel,[17] Sense.us,[15] and Many Eyes.[18] These systems support varied levels of sharing, discussion, and annotation of visualized data.

devices.[5,6] Systems supporting remote collaboration have primarily focused on synchronous interaction,[7–9] such as shared virtual workspaces[10] and augmented reality systems that enable multiple users to interact concurrently with visualized data.[11,12]

In contrast, relatively little research attention has focused on asynchronous collaboration around visualizations.[13] Yet, by partitioning work across both time and space, asynchronous collaboration may provide greater scalability for group-oriented analysis. There is evidence that, due in part to a greater division of labor, asynchronous decision making can result in higher-quality outcomes – broader discussions, more complete reports, and longer solutions – than face-to-face collaboration.[14]

One challenge to achieving the benefits of asynchronous collaborative analysis is determining the appropriate design decisions and technical mechanisms to enable effective collaboration around visual media. Creating effective collaborative visual analysis environments raises a number of design questions. How should collaboration be structured, and what shared artifacts can be used to coordinate contributions? What are the most effective communication mechanisms?

One source of design guidance comes from examining existing systems. A handful of recent web-based collaborative visualization systems[15–19] provide features for sharing and discussing visualized data (Figure 1). Geographic visualization systems such as Wikimapia[19] support end-user annotation of satellite imagery to identify landmarks and points of interest. Commercial visualization tools such as Spotfire Decision Site Posters[16] allow visualizations created in desktop applications to be posted to web pages and discussed through text comments. The public web sites Many-Eyes.com[18] and Swivel.com[17] allow users to upload data sets and visualize them using a palette of supported visualizations. Collaboration is supported through application bookmarks (in the form of URLs) into specific visualization states, blog-style text comments, and posting linked screenshots to external blogs. The research prototype sense.us[15] provides new techniques for linking commentary with visualization states and supports graphical annotation and the construction of tours through multiple visualization states. Observing usage of these systems provides numerous examples of group sensemaking in action[15] cycles of observation, question, and hypothesis; social navigation to interesting or controversial data; and identification of problematic or incorrect data values.

Still, these systems represent only the first steps in a larger design space; many challenges remain for providing effective, scalable forms of collaborative analysis. We wish to better support the observed user behaviors of these existing systems by grounding design decisions in both practical and theoretical knowledge of social interaction. A theoretically grounded design framework can be applied to contrast the existing offerings and guide the future research and development of social visual analysis systems. Towards this aim, we review research in analytics, social psychology, sociology, organizational studies, and computer-supported cooperative work to identify a set of design considerations to inform the development of asynchronous collaborative information visualization systems.

The goal of this article is to identify key issues to guide work in collaborative visual analytics. We have grouped our design considerations into seven inter-related areas: division and allocation of work; common ground and awareness; reference and deixis; incentives and engagement; identity, trust, and reputation; group dynamics; and consensus and decision making. In each of these areas, we discuss the aspect underlying effective collaboration and suggest specific mechanisms by which they could be achieved. We conclude by summarizing the various design considerations presented and suggesting avenues for future research and development in collaborative visual analytics.

Division and allocation of work

A fundamental aspect of successful collaboration is an effective division of labor among participants. This involves both the segmentation of effort into proper units of work and the allocation of individuals to tasks in a manner that best matches their skills and disposition. Primary concerns are how to split work among multiple participants and meaningfully aggregate the results.

Benkler[20] describes the role of modularity, granularity, and cost of integration in the peer production of information goods, drawing on examples such as online discussions, open source software, and Wikipedia. *Modularity* refers to how work is segmented into atomic units, parallelizing work into independent tasks. The *granularity* of a module is a measure of the cost or effort involved in performing the task. The optimal granularity of modules is closely tied to the incentives for performing the work. For example, in online scenarios where the incentives tend to be small and non-monetary, a small granularity helps facilitate participation, encouraging people to participate in part due to the ease of contributing. A variety of granularities enables different classes of contribution to emerge.

The third aspect of Benkler's model is the *cost of integration*: what effort is required to usefully synthesize the contributions of each individual module? Collaborative work will only be effective if the cost of integration is low enough to warrant the overhead of modularization while enforcing adequate quality control. There are a number of mutually inclusive approaches to handling integration: automation (automatically integrating work through technological means), peer production (casting integration as an additional collaborative task given to trusted participants), social norms (using social pressures to reduce vandalistic behavior), and hierarchical control (exercising explicit moderation).

Collaborative visual analytics can similarly be viewed as a process of peer production of information goods. Such goods may include the observations, questions, and hypotheses generated in the analysis process as well as tours or presentations of analysis results. Questions for collaborative visualization include how to facilitate the modularization of work. The first step is determining the units (modules) of contribution and their granularity. Existing frameworks for aiding this task include structural models of visualization design and sensemaking processes. Once modules have been identified, one can then attempt designs that reduce the cost structure for these tasks. Another important concern is the proscription of particular task types or roles – what aspects should be formally inscribed in the system and what should be left open to negotiation and definition by work groups themselves?

The information visualization reference model

One model for identifying modules of contribution is the *information visualization reference model*,[21,22] a general pattern for describing visualization applications (Figure 2). The model decomposes the visualization process into data acquisition and representation, visual encoding of data, and display and interaction. Each phase of this model provides an entry point for collaborative activity. Contributions involving data include uploading data sets, cleaning or reformatting data, moderating contributed data (e.g., to safeguard copyright or privacy concerns), and affixing metadata (e.g., providing keyword tags). Additional contributions of varying granularity lie in the application of visual encodings. Examples include matching data sets with existing visualization components, editing visual mappings to form more effective visualizations,

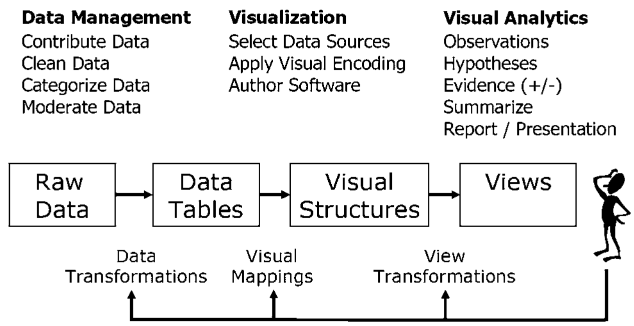

Figure 2 The information visualization reference model. Source data is mapped into tables that are visually encoded and presented in interactive views.[21,22] Collaboration may occur at the level of data management, visualization, or analysis.

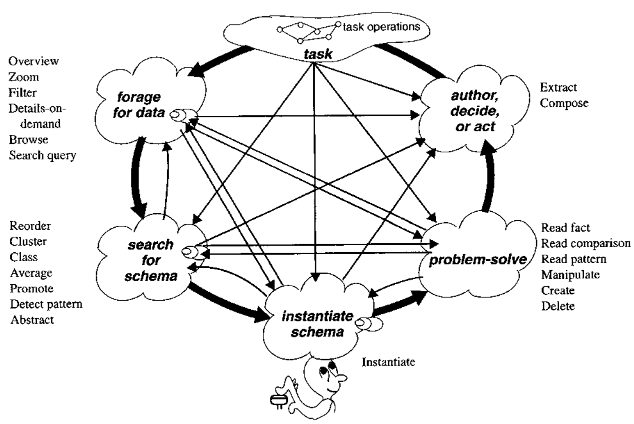

Figure 3 The sensemaking cycle. The diagram depicts the various phases and loops of the sensemaking process, annotated with common tasks. The image is taken from Card *et al.*[21]

and authoring visualization software components. Both Many Eyes[18] and Swivel[17] enable contribution of data sets and visual mappings. Important issues for future work include the accuracy and provenance of contributed data sets. The primary focus of this article, however, is at the level of interaction, where we consider how collaborative visual analysis and exploration can be conducted most effectively.

The sensemaking model

To better understand analytic contributions, we consult the sensemaking model,[21,23] which grounds the use of information visualization in a theory of how people search for, organize, and create new knowledge from source information. Social issues accrue at each phase of the model: how do people communicate, how do they judge the contributions of others, how are groups formed, and what motivates contributions? We touch on each of these issues in subsequent sections. As indicated by the numerous interconnections in Figure 3, the sensemaking process has a much higher degree of coupling than the information visualization reference model, carrying implications for the granularity, and integration of contributions.

Intelligence analysis provides examples of both co-operative and competitive models of work.[3] In cooperative scenarios, modules may be of fine granularity and pooled such that collaborators can immediately make use of others' work. Examples include finding relevant information sources, identifying connections between sources, and positing hypotheses. Such work may involve tightly coupled collaboration, requiring awareness and communication among participants. In competitive scenarios, work is not integrated until a later stage of sensemaking, such as detailed, evidence-backed hypotheses or recommended actions. While lacking the benefits of resource pooling, this approach encourages individual assessment and can reduce groupthink bias. Accordingly, it may benefit collaborative visualization systems to support both fine-grained and coarse-grained work parallelization.

If adopting a competitive model, the main concern is with integrating the end results of the sensemaking process. How can analytic conclusions or suggested actions be presented, compared, and evaluated? If co-operative models are used, either across all collaborators or within teams, we should consider social issues affecting each phase of sensemaking.

Information foraging The first phase of sensemaking is information foraging.[1] Given the underlying metaphor of foraging for food, an activity often performed by social packs of animals, social information foraging[24] seems a natural extension. Technologies for collaborative foraging could help pool findings, such as discovery of relevant information, and support notification updates. Design challenges include how to structure and categorize shared findings, such as identified trends or outliers, and provide task-sensitive retrieval mechanisms by which others can access them. Furthermore, systems could make the foraging behaviors of others visible by analyzing and displaying activity traces, facilitating social navigation[25] of data sets. Visualizing aggregate foraging behaviors is metaphorically similar to the scent trails left by ants foraging for food. In this form, general usage can be treated as an implicit collaborative contribution, a possibility discussed further in the section Awareness and social navigation.

Information schematization The next phases of sensemaking concern the construction and population of information schemata, in which findings from the foraging process are organized. Schematization could be conducted in a collaborative fashion by enabling information organization and discussion among collaborators. One challenge is to synthesize the contributions of various collaborators in a manner that reduces the cost of integration, resulting in accessible forms such as summaries of arguments and evidence. To this aim, asynchronous collaboration can be structured through shared external representations[26] for manipulating the information. For example, discussion forums aggregate contributions through the accretion of comments and replies in a sequential fashion, where as wikis (e.g., Wikipedia) and open source software rely on human editing backed by a revision management system to integrate and moderate contributions. Alternatively, systems with highly structured input such as NASA ClickWorkers[20] or von Ahn's 'games with a purpose'[27] rely on statistical aggregation. Clearly, the form of the collaborative artifact strongly affects the cost of integration; it may be more costly to find relevant information in a sprawling discussion forum than in a group-edited document or statistical summary. As the number of collaborators or the complexity of contributions increases, the need for mechanisms facilitating aggregation becomes more acute. Future research is needed to devise and evaluate new external representations for structuring collaborative visual analytics.

Some existing research suggests mechanisms for representing and integrating analytic contributions. The analyst's sandbox[28] provides a visual environment for spatially organizing hypotheses and positive and negative evidence. Coupled with revision management tools, the analyst's sandbox might also serve as a shared editing environment for collaborative analysis. Brennan *et al*.[29] introduce a tool for comparing and integrating the work of independent analysts. Their system uses a logic programming approach to merge network diagrams of collected evidence. Billman *et al*. CACHE[30] system supports the analysis of competing hypotheses; each analyst maintains a matrix of hypotheses and evidence and provides numerical measures of the reliability of evidence and assessments of the degree to which evidence confirms or disconfirms the hypotheses. The CACHE system statistically aggregates these ratings to form a group assessment. Argumentation systems such as Zeno[31] allow users to graphically structure an argument into claims, constraints, and evidence. Similar to CACHE, Zeno can then automatically evaluate the current level of support for the provided claims. While each of these systems suggest possible approaches to structuring the creation of information schemata, further investigation is needed to compare (and potentially hybridize) these approaches. Usability and expertise are also important concerns; techniques that work well for professional analysts may not be appropriate for supporting collaborative visualization for a general audience on the Web.

Problem solving, decision making, and action The final phases of sensemaking involve problem solving and action. These phases may or may not take place within the collaborative analysis environment. Furthermore, the analysts themselves may not be decision makers, thus mechanisms for presenting and disseminating analytic findings to others are often crucial components of collaborative work. Findings gained from the analysis may serve as input to collaboration in other media, suggesting the need to both facilitate external access to the contents of the visual analysis environment and extracting content for use in other systems. If collaborators conduct problem solving and decision making within the system, aforementioned issues regarding communication, discussion, and consensus must be addressed.

Common ground and awareness

Inspired by linguistics, social psychologists have investigated fundamental prerequisites for successful communication. Clark and Brennan[32] describe the concept of *common ground*, the shared understanding between conversational participants enabling communication. Through shared experience and discussion, people constantly monitor their mutual understanding. For example, facial expressions, body language, and backchannel utterances such as 'uh-huh' and 'hmm?' provide *grounding* cues of a participant's current level of understanding. Both positive evidence of convergence of understanding and negative evidence of misunderstanding are used to establish common ground.

Surprisingly, an imperfect shared understanding is often sufficient. *The principle of least collaborative effort* states that conversational participants will exert just enough effort to achieve successful communication.[33] Collaborative effort may be applied during both a *planning* stage, in which a participant formulates their next utterance, and an *acceptance* stage, in which a participant ascertains if partners have understood the utterance. This principle serves as an evaluation guide for collaboration mechanisms, as different mechanisms may affect the amount of effort needed for collaborators to effectively communicate. For example, multiple studies have shown that the media of communication affects the cost structure of collaborative effort.[34,35] views of a shared visual environment minimize the need to verbally confirm actions that can be assessed visually. However, media effects such as latency can hamper the efficiency benefits of such cues.[35]

At both general and detailed levels, grounding theory provides a useful guide for design decisions. When collaborating around visualizations, participants must be able to see the same visual environment in order to ground each others' actions and comments, suggesting the need for mechanisms for bookmarking or sharing specific states of the visualization. Collaborators must be able to share views to specific visualization states both within the visualization environment itself and across other media. For example, the results of visual analysis might be shared more effectively as part of a web page or blog, where a dedicated readership and familiarity with collaborators better establishes the necessary common ground with respect to the subject matter. At minimum, the ability to easily pass around pointers (e.g., URLs) to specific views is indispensable, and therefore collaborative visualizations must be able to explicitly represent and export their internal state space.[15,18,36]

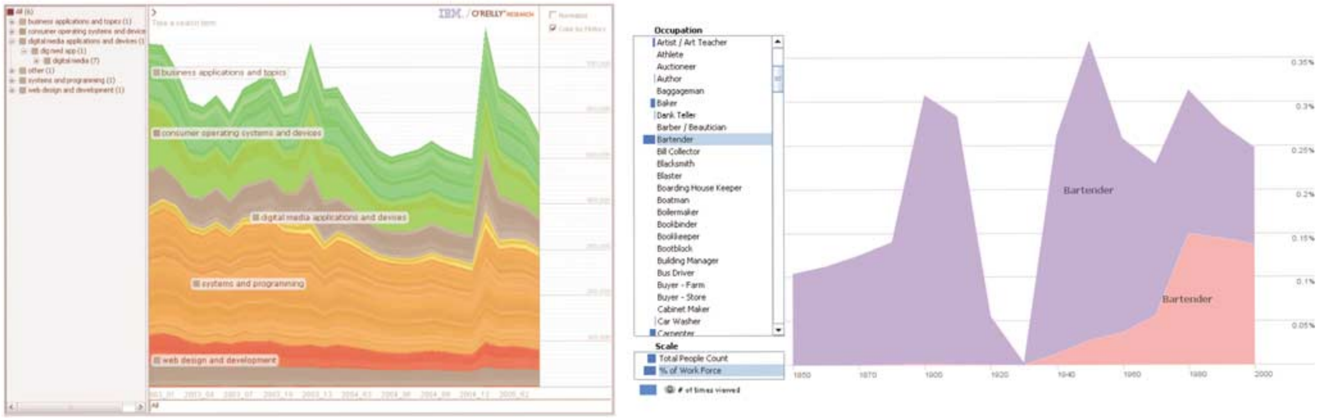

Figure 4 Visual social navigation cues. Left: Grayed-out regions of the visualization indicate data regions that have previously been visited.[36] Right: Bar charts embedded within navigation components indicate the relative visitation rates for visualization views reachable from the current state.[41]

Discussion models

Given the ability to access a shared viewpoint, one must still determine the forms of discussion and annotation around that view. For example, one could use visualization bookmarks within a standard discussion forum, is form of *independent discussion* is unidirectional, linking from text to the visualization. Most existing systems, including Decision Site Posters,[16] Many Eyes,[18] and Swivel,[17] provide support for independent, unthreaded comments. Another approach is *embedded discussion*, placing conversational markers directly within the visualization, such as comments over annotated geographic regions in Wikimapia.[19] This approach provides unidirectional links that point from the visualization to text.

Grounding might be further facilitated by more deeply tying discussion to the visualization state space. *Doubly linked* commentary[15] allows comments to link to specific views as in independent discussion, while also enabling all such discussions to be retrieved *in situ* as visualization views are visited. Our hypothesis is that directly associating commentary with specific states of the visualization will facilitate grounding by disambiguating the context of discussion, while also enabling serendipitous discovery of relevant discussion during exploration. Evidence for this hypothesis could take the form of simplified referential utterances or facilitation of reader comprehension. Future research might consider more complicated linking structures, such as tying discussion to multiple views, as well as conducting formal evaluations of the effects of varied discussion models.

Awareness and social navigation

Another important source of grounding comes from *awareness* of others' activities, allowing collaborators to gauge what work has been done and where to allocate effort next.[37,38] Within asynchronous contexts, participants require awareness of the timing and content of past actions. The need for coordination suggests that designs should include both *history* and *notification* mechanisms (e.g.,[39]) for following actions performed on a given artifact or by specific individuals or groups. Browseable histories of past action are one viable mechanism, as are subscription and notification technologies such as Really Simple Syndication and Atom (Figure 4).

User activity can also be aggregated and abstracted to provide additional forms of awareness. *Social navigation*[25,40] involves the use of activity traces to provide navigation cues based on the behavior of others, allowing users to purposefully navigate to past states of high interest or explore less-visited regions (termed 'anti-social navigation' by Wattenberg and Kriss.[36]) Navigation cues may be added to links to views with low visitation rates or to action items such as unanswered questions and unassessed hypotheses. One approach is to add visual cues to user interface widgets indicating aggregated social activity, such as the number of visitations to a particular visualization state or the number of comments linked to that state. For example, Willet *et al.*[41] embed small bar charts indicating either visitation or comment counts into dynamic query widgets. In a controlled experiment, they found that both comment and visitation cues increase revisitation to states visited by other users. They also found that visitation cues also led to significantly more unique discoveries when the data was unfamiliar to users, but that discovery rates equalized over subsequent trials.

Reference and deixis

A vital aspect of grounding is successfully referring to artifacts, people, places, or other items. As both Clark[42] and Brennan[34] explain, reference can take on many different forms. We focus on reference in spatial contexts.

When collaborating around visual media, it is common to refer to specific objects, groups, or regions visible to participants. Such references may be *general* ('north by northwest'), *definite* (named entities), *detailed* (described by attributes, such as the 'blue ball'), or *deictic* (pointing to an object and saying 'that one', also referred to as *indexical* reference). Once the referent has been successfully established and grounding has been achieved between participants, collaboration can move forward.

Clark[42] surveys various forms of spatial indexical reference, grouping them into the categories of *pointing* and *placing*. Pointing behaviors use some form of vectorial reference to direct attention to an object, group, or region of interest, such as pointing a finger or directing one's gaze. Placing behaviors involve moving an object to a region of space that has a shared, conventional meaning. Examples include placing groceries on a counter to indicate items for purchase and standing across from the teller to indicate that you will be the purchaser. In addition to directing attention, indexical reference allows patterns of speech and text to change. Participants can use deictic terms like 'that' and 'there' to invoke indexical referents, simplifying the production of utterances along the principle of least collaborative effort.

Hill and Hollan[43] further discuss the various roles that deictic pointing gestures can play, often communicating intents more complicated than simply 'look here.' They describe how different hand gestures can communicate angle (oriented flat hand), height (horizontal flat hand), intervals (thumb and index finger in 'C' shape), groupings (lasso'ing a region), and forces (accelerating fist). While other forms of reference are often achieved through speech or written text, deictic reference in particular offers important interface design challenges for collaborative visualization. Our hypothesis is that methods for performing nuanced pointing behaviors can improve collaboration by favorably altering its cost structure. Hill and Hollan make this claim explicitly, arguing for 'generally applicable techniques that realize complex pointing intentions' by engaging 'pre-attentive vision in the service of cognitive tasks.'

Brushing and dynamic queries

A standard way to point in a visualization is through *brushing*[44,45] and *dynamic queries*[46] selecting and highlighting a subset of the data through direct manipulation of the display or auxiliary query controls. Naturally, these selections should be sharable as part of the state of the visualization. In addition, a palette of visual effects richer than simple filtering and highlighting can let users communicate different intents. For example, a user selecting a range of values in a chart might have one of any number of intents. If the user is interested in the specific points selected, those points should be prominently highlighted. However, if the user is primarily interested in the range of the contained values, the range interval should be given visual prominence.

Brushing-based forms of pointing have the advantage that the pointing action is tied directly to the data, whether modeled as a vector of selected tuples, a declarative query, or both. 'Data-aware' representations allow a pointing intention to be reapplied in different views of the same data, enabling reuse of references across different choices of visual encodings. Data-aware annotations could also enable users to search for all commentary or visualizations that reference a particular data item. As data-aware annotations are machine-readable, they might also be used to export subsets of data and help steer automated data mining (e.g.,[47]). Finally, machine-readable selections might be used as input for achieving more generalized forms of reference. For example, one might deictically refer to a particular object, but formulate a broader selection by abstracting from the properties of that object (e.g., 'select all items blue like this one'). In this way, other forms of reference might be achieved in both human and machine readable form.

Graphical annotation

Freeform *graphical annotations*[15] can provide an expressive form of pointing. Drawing a circle around a cluster of items or pointing an arrow at a peak in a graph can direct the attention of remote viewers. The angle of the arrow or shape of the hand-drawn circle may communicate emotional cues or add emphasis. Although such drawing and vector graphic annotations allow a high degree of expression, without any explicit tie to the underlying data they only apply to a single view in the visualization. Freeform annotations can persist over purely visual transformations such as panning and zooming, but if they are not data-aware they may become meaningless in the face of data-oriented operations such as filtering or drill-down. One future research direction is to develop hybrid approaches that combine aspects of both brushing and graphical annotation. The resulting techniques could create graphical annotations that are tied to data points so that they can be displayed in other views of the data and remain meaningful (Figure 5).

Ambiguity of reference

An additional concern is ambiguity of reference. Clark *et al.*[48] demonstrate how people's common ground can affect ambiguity resolution: two people with greater familiarity might successfully communicate using ambiguous references, while a third participant remains confused. Asynchronous collaboration may be more susceptible to ambiguities than synchronous collaboration because participants do not receive immediate feedback or grounding cues from other collaborators. As a result, designers of pointing interactions must also consider the ease with which pointing actions can be interpreted unambiguously. The implicit interplay between gesture and text, often fluidly performed and subconsciously interpreted in synchronous interactions, may need to

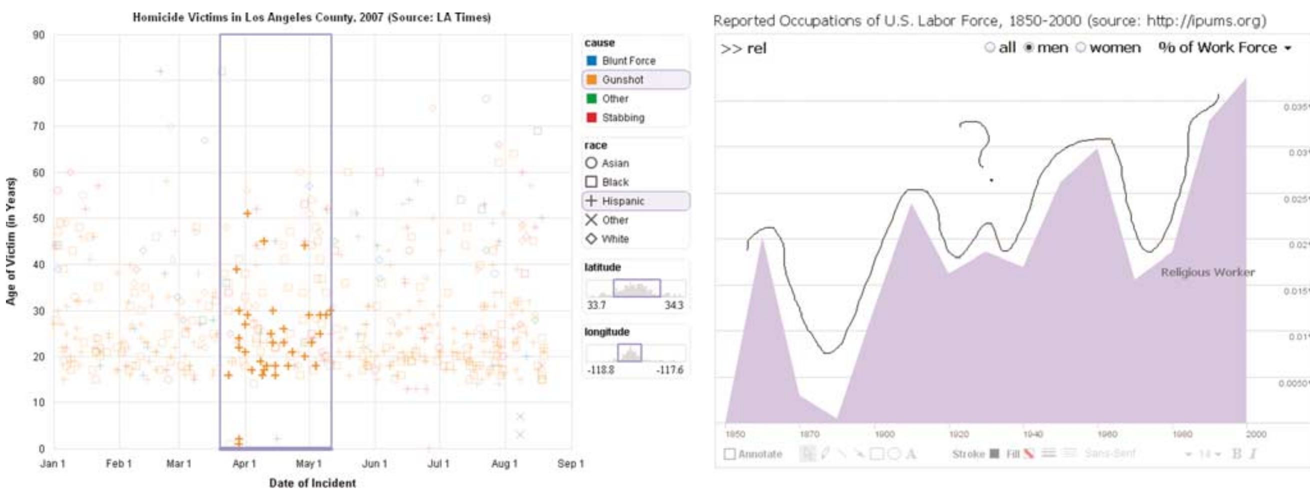

Figure 5 Annotation and selection in visualizations. Left: Brushing and dynamic query techniques enable forms of pointing and selection tied to the underlying data. Right: Freeform graphical annotation enables a greater range of expressiveness.[15]

be linked more concretely in asynchronous settings. For example, a text comment involving multiple deictic terms may need to link those terms explicitly to visual annotations, as the gestural cues used in face-to-face communication are not available for disambiguation.

Incentives and engagement

If collaborators are professionals working within a particular context (e.g., financial analysts or research scientists) there may be existing incentives, both financial and professional, for conducting collaborative work. In a public goods scenario, incentives such as social visibility or sense of contribution may be motivating factors. Incorporating incentives into the design process may increase the level of contribution, and could provide additional motivation in those situations that already have established incentive systems.

Benkler[20] posits an incentive structure for collaborative work consisting of monetary incentives, hedonic incentives, and social-psychological incentives. *Monetary* incentives refer to material compensation such as a salary or cash reward. *Hedonic* incentives refer to well-being or engagement experienced intrinsically in the work. *Social-psychological* incentives involve perceived benefits such as increased status or social capital.

Personal relevance

A number of observations of social use of visualization have noted that visualization users are attracted to data that they find personally relevant.[49,13,36] For example, in collaborative visual analysis of the occupations of American workers,[15] people often start by searching for their own profession and those of their friends and family, similar to the way people search for names in the popular NameVoyager visualization.[36] The hypothesis is that by selecting data sets or designing the presentation such that the data is personally relevant, usage rates will rise due to increased hedonic incentive. For example, geographic visualizations may facilitate navigation to personally relevant locations through typing in zip codes or city names, while a visualization of the United States' budget might communicate how a specific user's taxes were allocated rather than only listing total dollar amounts.

Social-psychological incentives

In the case of social-psychological incentives, the visibility of contributions can be manipulated for social effects. Ling *et al.*[50] found that users contributed more if reminded of the uniqueness of their contribution or if given specific challenges. In one experiment, Ling *et al.* also found that participants contributed more when given group goals than when given individual goals, a finding at odds with existing social-psychological theory. Cheshire[51] ran a controlled experiment finding that, even in small doses, positive social feedback on a contribution greatly increases contributions. He also found that visibility of high levels of cooperative behavior across the community increases contributions in the short term, but has only moderate impact in the long term. These studies suggest that social-psychological incentives can improve contribution rates, but that the improvements depend on the forms of social visibility. One incentive for visual analysis may be to prominently display new discoveries or successful responses to open questions. Mechanisms for positive feedback, such as voting for interesting comments, might also foster more contributions.

Game play

Finally, it is worth considering game play as an additional framework for increasing incentives. For example, von Ahn's 'games with a purpose'[27] reframe otherwise tedious data entry tasks as actions within online games, successfully leveraging game dynamics to engage users in the construction of information goods. Heer[49] discusses various examples in which playful activity contributes to visual analysis, applying insights from an existing theory of playful behavior[52] that analyzes the competitive, visceral, and teamwork building aspects of play. For example, scoring mechanisms create competitive social-psychological incentives. Game design might also be used to allocate attention. For example, creating a team-oriented 'scavenger hunt' analysis game could focus participants on a particular subject matter. Salen and Zimmerman[53] provide a thorough resource for the further study of game design concepts.

Identity, trust, and reputation

Aspects of identity, reputation, and trust all influence the way people interact with each other. Within a sensemaking context, interpersonal assessments affect how people value, consider, and respond to the individual contributions of others. Other things being equal, a hypothesis suggested by a more trusted or reputable person will have a higher probability of being accepted as part of the group consensus.[54] For social sensemaking in a computer-mediated environment, design challenges accrue around the various markers of identity and past action that might be transmitted through the system. For example, Donath[55] describes how even a cue as simple as one's e-mail address can lead to a number of inferences about identity and status.

Identity presentation

Many theorists try to understand interpersonal perception via the signals available for interpretation by others. Goffman[56] distinguishes between *expressions given* and *expressions given off* to indicate those parts of our presentation of self that are consciously planned (e.g., the content of our speech) or unconsciously generated (e.g., a wavering of voice indicating nervousness), each of which is interpreted to form opinions of a person. Donath[55] classifies such signals into *conventional signals* – low-cost signals that are easy to fake (e.g., talking about going to the gym) – and *assessment signals* – more reliable signals that are difficult to fabricate (e.g., having large muscles).

Other researchers have focused on the way media with different capacities for transmitting such signals affect interpersonal assessment. For example, most computer-mediated communication filters out non-verbal cues, stripping many of the signals 'given off' by participants. Despite these missing cues, Walther[57] argues that online relationships can be as deep and meaningful as face-to-face interactions through explicit sharing of personal information. However, due to diminished cues and asynchronous interaction, such online relationships can require longer time spans to develop. These diminished cues allow for a greater role of imagination and speculation when assessing another person. Furthermore, many researchers find that such diminished cues give rise to 'deindividuation' effects that have both desirable and undesirable consequences. For example, people who are shy in face-to-face interactions often show greater rates of participation, but anti-social 'flaming' is also more prevalent online.[57,58]

When considering the implications of identity assessment for collaborative visualization systems, designers should also take the context of deployment into account. If collaborators are already familiar to each other, they may require little additional support to make assessments of identity and reputation, instead of relying on existing channels through which assessments can be made. It may be enough to simply identify collaborators' individual contributions with recognizable names. Still, it may prove valuable for visual analysis environments to interface with external communication channels, both for sharing and interpersonal assessment. Many organizations maintain online personnel directories to aid awareness and collaboration; visual analysis systems should be able to leverage such existing artifacts.

On the other hand, if collaborators begin as strangers, mechanisms for self-presentation and reputation formation need to be included in the system design. Possible mechanisms include identity markers, such as screen names, demographic profiles, social networks, and group memberships. Considerations include the type of personal information germane to the context of visual analysis; for example, is a playful or professional environment desired? Attributes such as age, geographic location, interests, and skills might help assess a collaborator's background knowledge, affecting the confidence one places in hypotheses. Of course, this picture is complicated if such measures are self-reported, because such self-reports may be subject to fabrication.

Reputation formation

Considering how interpersonal assessment develops over time gives rise to questions of reputation and trust formation. In the case where participants only interact through the system itself, means of gauging a user's past actions or contributions are needed not only to aid awareness (c.f. section on common ground and awareness) but also to facilitate reputation formation. Observations of past actions provide *implicit* means of reputation formation, allowing collaborators to make inter-personal judgments grounded in past activity. One challenge for design is to consider what pieces of information are most informative for reputation formation.

Some systems instead provide *explicit* reputation mechanisms, such as seller ratings in online markets such as eBay.[59] In a visual analysis environment, collaborators might rate each other's contributions according to their interestingness or accuracy. Such ratings may help surface contributions with higher relevance, provide a reputation metric for contributors, and additionally constitute a social-psychological incentive for high quality contributions.

Group dynamics

The makeup of collaborative groups is another aspect important to social sensemaking. Many scenarios, such as business and research, may involve work groups that are already well established. In such cases, standard group management and communication features common to many collaborative applications may be sufficient. However, when organizing effort in public goods scenarios, explicit mechanisms for assisting group formation may aid collaborative visualization efforts.

Group management

At a basic level, formal *group management* mechanisms must present means for addressing issues of scalability and privacy. Group management mechanisms can support the coordination of a work group on a specific task within a larger collaborative environment, providing notification and awareness features at the group level. Groups also provide a means of filtering contributions, improving tractability and reducing information overload for participants who may not be interested in the contributions of strangers. Finally, groups provide a means of limiting contribution visibility, providing one mechanism for individual privacy within large-scale online scenarios.

An alternative approach to explicit group management is to support groups already formed in other mediated environments. Such support requires a decentralization of the analysis process, enabling collaborative visual analysis technologies to be embedded in external media. Examples include embedding an interactive visualization into a blog entry or introducing visual analysis applications into existing social environments such as Facebook. This strategy is common with existing social data analysis sites like Swivel[17] and Many-Eyes[18] the longest and deepest discussions tend to occur around visualization screenshots posted to an external blog.

Group size

One challenge for group management is the choice of group size. Larger groups may be able to achieve more through a larger labor pool, but can incur social and organizational costs.[60] For example, larger groups are more likely to suffer from the *free rider* problem[61] or *social loafing*[50] due to diluted accountability. Pirolli[24] describes a mathematical model of social information foraging that measures the benefit of including additional collaborators in information gathering tasks. His analysis finds that beyond certain sizes, additional foragers provide decreasing benefits, suggesting that an optimal group size exists, dependent on the parameters of the foraging task. An important future experiment would be to evaluate Pirolli's model through application to real visual analysis teams.

Group diversity

Another issue facing group formation is the diversity of group members. In this case diversity can include the distribution of domain-specific knowledge among potential participants and differences in attributes such as geographical location, culture, and gender. Organizational studies[62,63] find that increased group diversity can lead to greater coverage of information and improved decision making. However, diversity can also lead to increased discord and longer decision times.

Various measurements of diversity may be applied to suggest a set of group members that will provide adequate coverage for an analysis task. Such measurements might come from analyzing differences between user profiles and structural features of the social networks of the participants.[64] Such networks may be explicitly articulated or inferred from communication patterns, such as the co-occurrence of commenters across discussion threads. Wu *et al.*[65] study of organizational information flow found that information spreads efficiently among homophilous (similar) group members but not across community boundaries, further suggesting the value of identifying structural holes and directing bridging individuals in the social network towards particular findings. By constructing user profiles based on demographic data, social connectivity, and prior usage, automated systems may be able to help suggest relevant tasks to appropriate community members.

Consensus and decision making

The need to establish group consensus arises in many phases of the sensemaking cycle. Examples include agreement about the data to collect, how to organize and interpret data, and making decisions based upon the data. Collaborators may reach consensus through discussion or through the aggregation of individual decisions.

Consensus and discussion

Mohammed[54] combines various contributions in social psychology and organizational studies to posit a model for cognitive consensus in group-decision making. Mohammed's model takes into account the assumptions, category labels, content domains, and causal models possessed by each participant, and how they might evolve through discussion. One tangible recommendation that comes from this work is to clearly identify the points of dissent, creating focal points for further discussion and negotiation. From a design perspective, collaborators need communication mechanisms that allow points of dissent to be labeled and addressed. Collaborative tagging[66] is one potential candidate. Formalizing contributions in structured argumentation systems[30,31] may be another avenue. For example, a content analysis of contributions to the sense.us system[15] found that users primarily used free-text comments to post observations, questions, and hypotheses. These categories could be formally represented to help structure discussion and voting.

Scheff[67] notes that consensus requires more than participants simply sharing a belief; participants must think that their beliefs are the same, and achieve realization that others understand one's position. Users need feedback loops to gauge mutual understanding. Along these lines, it may be useful to consider the effects of multiple communication channels on decision processes. Collaborative visualization environments that provide messaging, in either synchronous or asynchronous forms, might provide backchannels for negotiation and non-public discussion. The integration of instant messaging into the GMail e-mail service is an example of weaving different communication channels into a single system.

The value of different forms of consensus can vary based on the task at hand. Hastie[68] found that group discussion improved accuracy when decision tasks had demonstrably correct solutions because groups could evaluate their output. When task outcomes are open-ended, consensus through discussion is harder to evaluate. In a simulated graduate admissions task, Gigone and Hastie[69] found little value in discussion, as group decisions were well-matched by simply averaging prior individual decisions.

One design implication that again arises is to use *voting* or *ranking* systems. Mechanisms for expressing support or disdain for hypotheses could aid data interpretation and further identify controversial points. For example, Wikimapia[19] users can vote on the accuracy of labeled geographic regions and Swivel[17] supports ratings of interestingness. A game-like variation on this approach is the creation of *prediction markets*[70] individuals can be given a limited amount of 'points' or 'currency' that they can use to vote for hypotheses they find most promising. Hypotheses that are later validated could reap payback rewards for their supporters.

Information distribution

An important dimension of group consensus is the distribution of information across group members. Both Stasser[71] and Gigone and Hastie[69] find that it is difficult for groups to pool information effectively, and therefore, decision-making is biased in the direction of the initial information distribution. They hypothesize that the lack of effective pooling may be due to the persistence of individual decisions made prior to discussion or to information shared prior to the group meeting. The prior decisions and information set a common ground for discussion and bias conversation against the unshared information. Thus, improving collective information foraging may help inform group decision-making by changing the information distribution. Collaborative analysis environments may facilitate better information pooling by providing a record of findings and opinions that can be surveyed prior to decision-making and discussion.

Presentation and story-telling

Common forms of information exchange in group sensemaking are reports and presentations. Narrative presentation of analysis 'stories' is a natural and often effective way to communicate analytic findings, and is a primary use of Spotfire Decision Site Posters.[16] Furthermore, users of Swivel,[17] sense.us,[15] and Many Eyes,[18] all use external media such as blogs and social bookmarking services as external communication channels in which to share and discuss findings from visualizations. The challenge to collaborative visualization is to provide mechanisms to aid the creation and distribution of presentations. For example, sense.us[15] allows users to construct and share trails of saved views and can thus provide tours spanning multiple visualizations. GeoTime Stories[72] supports textual story-telling with hyperlinks to visualization states and annotations. However, neither system yet allows these stories to be exported outside the respective applications. In future work, such mechanisms could be improved with support to build presentations semi-automatically from interaction histories, apply pointing and annotation techniques, and embed the resulting presentations in external media. Viewers of analysis stories may also find value in conducting follow-up analysis and verification on parts of the story, enabling presentations to serve as a catalyst for additional analysis.

Conclusion and future directions

This article presents design considerations for collaborative visual analytics, attempting to identify the aspects underlying successful collaboration and suggest mechanisms for achieving them. Highlights include a list of collaborative visualization tasks, techniques to improve shared context and awareness, and suggestions for increasing engagement and allocating effort. Many of these considerations are summarized in Table 1. The overarching goal is to design socio-technical systems that improve our analytic capabilities by promoting an effective division of labor among participants, facilitating mutual understanding, and reducing the costs associated with collaborative tasks.

Visiting these considerations also provides an agenda for future research in collaborative visual analytics, surfacing hypotheses in need of study and suggesting new technical mechanisms. For example:

- What is the effect of different discussion models (e.g., independent, embedded, and doubly linked) on participation and the establishment of common ground?
- Beyond textual discussion, what external representations will effectively support collaborative analysis? How do such artifacts affect grounding and the cost of integration?

Table 1 Selected design considerations for collaborative visual analytics. The table lists many of the individual design considerations visited in this article, providing a brief description and noting the relevant sections that discuss the issue in detail

Design consideration	*Description*	*Section headings*
Modularity and granularity	Identify appropriately-scoped units of work that form basic analytic contributions.	Division and allocation of work
Cost of integration	Synthesize work while attempting to lower integration costs and maintain quality.	Division and allocation of work
Shared artifacts	Structure collaboration through shared, editable representations.	Division and allocation of work, Common ground and awareness
Artifact histories	Provide histories of actions performed on artifacts.	Common ground and awareness
View sharing / bookmarking	Enable lightweight sharing of views across media with bookmarks (e.g., URLs)	Common ground and awareness
Content export	Support embedding of annotated views in external media (e.g., email, blogs, and reports)	Common ground and awareness
Discussion	Support commentary; consider implications of discussion model on common ground.	Common ground and awareness, Consensus and decision making
Notification	Support notification subscriptions for views, artifacts, people, and groups.	Common ground and awareness
Action flags	Mark needed future actions: unanswered questions, need for evidence, etc.	Common ground and awareness, Consensus and decision making
Social navigation	Make activity patterns visible, determine popular, and neglected data regions.	Division and allocation of work, Common ground and awareness
Recommendation	Suggest related views, comments, and data to current points of interest.	Common ground and awareness
Pointing techniques	Support nuanced pointing through selection techniques and visual effects.	Reference and deixis
Personal relevance	Increase engagement by increasing personal relevance of data sets.	Incentives and engagement
Social-psychological incentives	Increase engagement by surfacing unique individual contributions.	Incentives and engagement
Game play	Game design elements can provide incentives and be used to direct effort.	Division and allocation of work, Incentives and engagement
Identity markers	Enable identification of collaborators in a contextually-appropriate manner.	Identity, trust, and reputation
User profiles	Support awareness of others' backgrounds and skills.	Identity, trust, and reputation
Activity histories	Personal action histories allow past contributions to be assessed.	Common ground and awareness, Identity, trust, and reputation
Activity summaries	Activity indicators or summaries aid reputation and visibility of contributions.	Common ground and awareness, Incentives and engagement, Identity, trust, and reputation
Group management	Group creation and management mechanisms address issues of scale and privacy.	Group dynamics
Group size	Optimal group size determination can improve efficiency of analysis.	Division and allocation of work, Group dynamics
Group diversity	Appropriate within-group diversity can result in more complete results.	Division and allocation of work, Group dynamics
Voting and ranking	Quantitative measures can be used for consensus and to lower integration costs.	Division and allocation of work, Identity, trust, and reputation, Consensus and decision making
Presentation	Support creation and export of presentations for telling analysis stories.	Common ground and awareness, Reference and deixis, Consensus and decision making

- How can the synthesis of individual contributions be improved? Can (semi-)automatic summarization or merging of separately developed data views (e.g.,[29]) be used to form aggregated contributions?
- How should selection and visual emphasis techniques be designed to provide nuanced pointing behaviors? Can referenced objects be unambiguously recognized by both human and machine collaborators?
- How can pointing and graphical annotation gracefully handle dynamic visualizations and changing data sets?
- How should social navigation cues be effectively added to visual analysis tools to unobtrusively improve awareness?
- Can automated techniques be used to help allocate effort? For example, mining past contributions, user profiles, and inferred social networks may enable systems to direct collaborators to tasks in need of attention.
- How can the fruits of collaborative visual analysis be more effectively exported, shared, and embedded in external media such as web pages, e-mail, and presentations?

To answer these questions, we envision future research projects of varying scopes. Researchers may focus on new visualization and interaction techniques for supporting collaboration. Such research should propose novel mechanisms and ideally evaluate them through comparative study with other approaches. As listed above, novel discussion models, pointing techniques, and story-telling interfaces are all candidates. Research into targeted techniques needs to be balanced with the design, deployment, and evaluation of holistic collaborative visual analysis environments. Such systems should enable real-world groups to conduct collaborative visual analysis. Studies of system usage can then measure the benefits of collaborative visual analytics in ecologically valid settings and inform best practices for combining collaboration mechanisms. A number of important experiments, such as those involving group management and incentives, may be best conducted in real-world settings (e.g., Ling *et al.*,[50] Resnick *et al.*[59] and Wu *et al.*[65]) and interfacing with the Internet is critical to understanding how findings are disseminated and how collaborative visual analytics can be more deeply weaved into the Web. These and other challenges present exciting opportunities for advancing visual analytics research.

Acknowledgments

We thank Martin Wattenberg, Fernanda Viégas, Stu Card, Coye Cheshire, and Andrew Fiore for their insightful comments.

References

1 Pirolli P, Card SK. Information foraging. *Psychological Review* 1999; **106**: 643–675.
2 Pirolli P, Card SK. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *Proceedings of International Conference on Intelligence Analysis* (McLean, VA) MITRE: 2005.
3 Thomas JJ, Cook, KA (Eds). *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Press: New York, 2005.
4 Dorling D, Barford A, Newman A. Worldmapper: The world as you've never seen it before. *IEEE Transactions on Visualization and Computer Graphics* 2006; **12**: 757–764.
5 Dietz PH, Leigh DL. DiamondTouch: a multi-user touch technology, *ACM Symposium on User Interface Software and Technology* (Orlando, FL) ACM: New York, NY, 2001; 219–226.
6 General Dynamics. Command post of the future. [WWW document] http://www.gdc4s.com/content/detail.cfm?item = 2a58f8e2-ef2b-4bb1-9251-42ee4961dd7f (accessed 28 November 2007).
7 Agrawala M, Beers A, Froehlich B, Hanrahan P, MacDowall P, Bolas M. The two-user responsive workbench: support for collaboration through individual views of a shared space. *ACM Conference on Computer Graphics and Interactive Techniques* (*SIGGRAPH'97*) (Los Angeles, CA) ACM: New York, NY, 1997; 27–332.
8 Anupam V, Bajaj CL, Schikore D, Schikore M. Distributed and collaborative visualization. *IEEE Computer* 1994; **27**: 37–43.
9 Brodlie KW, Duce DA, Gallop JR, Walton JPRB, Wood JD. Distributed and collaborative visualization. *Computer Graphics Forum* 2004; **23**: 223–251.
10 Chuah MC, Roth S. Visualizing common ground. *Information Visualization* (*IV*) (London, UK) IEEE Computer Society: Washington, DC, 2003; 365–372.
11 Benko H, Ishak EW, Feiner S. Collaborative mixed reality visualization of an archaeological excavation. *IEEE International Symposium on Mixed and Augmented Reality* (*ISMAR 2004*) (Arlington, VA) IEEE Computer Society: Washington, DC, 2004; 132–140.
12 Chui Y-P, Heng P-A. Enhancing view consistency in collaborative medical visualization systems using predictive-based attitude estimation. *First IEEE International Workshop on Medical Imaging and Augmented Reality* (*MIAR '01*) (Hong Kong) IEEE Computer Society: Washington, DC, 2001; 292.
13 Viégas FB, Wattenberg M. Communication-minded visualization: a call to action. *IBM Systems Journal* 2006; **45**: 801–812.
14 Benbunan-Fich R, Hiltz SR, Turoff M. A comparative content analysis of face-to-face vs. asynchronous group decision making. *Decision Support Systems Archive* 2003; **34**: 457–469.
15 Heer J, Viégas F, Wattenberg M. Voyagers and voyeurs: supporting asynchronous collaborative information visualization. *ACM Conference on Human Factors in Computing Systems* (*CHI'07*) (San Jose, CA) ACM: New York, NY, 2007.
16 Spotfire Decision Site Posters. [WWW document] http://spotfire.com/products/decisionsite_posters.cfm (accessed 28 November 2007).
17 Swivel. [WWW document] http://www.swivel.com (accessed 28 November 2007).
18 Viégas FB, Wattenberg M, van Ham F, Kriss J, McKeon M. Many eyes: a site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics (InfoVis'07)* 2007; **13**: 1121–1128.
19 Wikimapia. [WWW document] http://wikimapia.org (accessed 28 November 2007).
20 Benkler Y. Coase's penguin, or, Linux and The Nature of the Firm. *Yale Law Journal* 2002; **112**: 369–446.
21 Card SK, Mackinlay JD, Shneiderman B. *Readings in Information Visualization: Using Vision To Think*. Morgan-Kaufmann: San Francisco, CA, 1999.
22 Heer J, Agrawala M. Software design patterns for information visualization. *IEEE Transactions on Visualization and Computer Graphics* 2006; **12**: 853–860.
23 Russell DM, Stefik MJ, Pirolli P, Card SK. The cost structure of sensemaking. *ACM Conference on Human Factors in Computing Systems* (*CHI'93*) (Amsterdam, Netherlands) ACM: New York, NY, 1993.

24 Pirolli P. Social information foraging. *Information Foraging: Adaptive Interaction with Information*, Chapter 8. Oxford University Press: Oxford, 2007.
25 Dourish M, Chalmers M. Running out of space: models of information navigation. *Human Computer Interaction* (*HCI'94*) (Glasgow, UK) Cambridge University Press: Cambridge, UK, 1994.
26 Zhang J, Norman DA. Representations in distributed cognitive tasks. *Cognitive Science* 1994; **18**: 87–122.
27 von Ahn L. Games with a purpose. *IEEE Computer* 2006; **39**: 92–94.
28 Wright W, Schroh D, Proulx P, Skaburskis A, Cort B. The sandbox for analysis: concepts and evaluation. *ACM Conference on Human Factors in Computing Systems* (*CHI'06*) (Montreal, Quebec) ACM: New York, NY, 2006.
29 Brennan SE, Mueller K, Zelinsky G, Ramakrishnan IV, Warren DS, Kaufman A. Toward a multi-analyst, collaborative framework for visual analytics. *IEEE Symposium of Visual Analytics Science and Technology* (Baltimore, MD) IEEE Computer Society: Washington, DC, 2006.
30 Billman D, Convertino G, Shrager J, Pirolli P, Massar J. Collaborative intelligence analysis with CACHE and its effects on information gathering and cognitive bias. *Human Computer Interaction Consortium Workshop* (Frasier, CO) 2006.
31 Gordon TF, Karacapilidis N. The Zeno argumentation framework. *Sixth International Conference on Artificial Intelligence and Law* (Melbourne, Australia) ACM Press: New York, NY, 1997; 10–18.
32 Clark HH, Brennan SE. Grounding in communication. In: Resnick LB, Levine RM and Teasley SD (Eds). *Perspectives on Socially Shared Cognition*. APA Books: Washington, 1991; 127–149.
33 Clark HH, Wilkes-Gibbs D. Referring as a collaborative process. *Cognition* 1986; **22**: 1–39.
34 Brennan SE. How conversation is shaped by visual and spoken evidence. In: Trueswell J and Tanenhaus M (Eds). *Approaches to Studying World-Situated Language Use: Bridging the Language-as-Product and Language-as-Action Traditions*. MIT Press: Cambridge, 2005; 95–129.
35 Gergle D, Kraut RE, Fussell SR. Language efficiency and visual technology: minimizing collaborative effort with visual information. *Journal of Language & Social Psychology* 2004; **23**: 491–517.
36 Wattenberg M, Kriss J. Designing for social data analysis. *IEEE Transactions on Visualization and Computer Graphics* 2006; **12**: 549–557.
37 Carroll J, Rosson MB, Convertino G, Ganoe CH. Awareness and teamwork in computer-supported collaborations. *Interacting with Computers* 2005; **18**: 21–46.
38 Dourish P, Belotti V. Awareness and coordination in shared workspaces. *ACM Conference on Computer-Supported Cooperative Work* (Toronto, ON) ACM: New York, NY, 1992; 107–114.
39 Brush AJ, Bargeron D, Grudin J, Gupta A. Notification for shared annotation of digital documents. *ACM Conference on Human Factors in Computing Systems* (CHI'02) (Minneapolis, MN) ACM: New York, NY, 2002.
40 Hill WC, Hollan JD, Wroblewski D, McCandless T. Edit wear and read wear. *ACM Conference on Human Factors in Computing Systems* (CHI'92) (Monterey, CA) ACM: New York, NY, 1992; 3–9.
41 Willett W, Heer J, Agrawala M. Scented Widgets: improving navigation cues with embedded visualizations. *IEEE Transactions on Visualization and Computer Graphics* 2007; **13**: 1129–1136.
42 Clark HH. Pointing and placing. In: Kita S (Ed). *Pointing where Language, Culture, and Cognition Meet*. Erlbaum: London, 2003; 243–268.
43 Hill WC, Hollan JD. Deixis and the future of visualization excellence. *IEEE Visualization* (San Diego, CA) IEEE Computer Society: Washington, DC, 2007; 314–319.
44 Becker RA, Cleveland WS. Brushing scatterplots. *Technometrics* 1987; **29**: 127–142.
45 Martin AR, Ward MO. High dimensional brushing for interactive exploration of multivariate data. *IEEE Visualization* (Atlanta, GA) IEEE Computer Society: Washington, DC, 1995; 271–278.
46 Ahlberg C, Shneiderman B. Visual information seeking: tight coupling of dynamic query filters with starfield displays. *ACM Conference on Human Factors in Computing Systems* (*CHI'94*) (Boston, MA) ACM: New York, NY, 1994; 313–317.
47 Yang D, Rundensteiner EA, Ward MO. Analysis guided visual exploration to multivariate data. *IEEE Symposium on Visual Analytics Science and Technology* (Sacramento, CA) IEEE Computer Society, Washington, DC, 2007.
48 Clark HH, Schreuder R, Buttrick S. Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior* 1983; **22**: 245–258.
49 Heer J. Socializing visualization. *CHI 2006 Workshop on Social Visualization* (Montreal, Quebec) ACM: New York, NY, 2006.
50 Ling K, Beenen G, Ludford P, Wang X, Chang K, Cosley D, Frankowski D, Terveen L, Rashid AM, Resnick P, Kraut R. Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication* 2005; **10** (http://jcmc.indiana.edu/).
51 Cheshire C. Selective Incentives and Generalized Information Exchange. *Social Psychology Quarterly* 2007; **70**: 82–100.
52 Caillois R. *Man, Play, and Games*. Free Press of Glencoe: Glencoe, IL,1961.
53 Salen K, Zimmerman E. *Rules of Play: Fundamentals of Game Design*. MIT Press: Cambridge, MA, 2003.
54 Mohammed S. Toward an understanding of cognitive consensus in a group decision-making context. *The Journal of Applied Behavioral Science* 2001; **37**: 408–425.
55 Donath JS. Identity and deception in the virtual community. In: Smith M and Kollock P (Eds). *Communities in Cyberspace*. Routledge: London, 1998.
56 Goffman E. *The Presentation of Self in Everyday Life*. Anchor Books: New York, 1959.
57 Walther JB. Computer-mediated communication: impersonal, interpersonal and hyperpersonal interaction. *Communication Research* 1996; **23**: 3–43.
58 Sproull L, Kiesler S. Computers, networks, and work. *Scientific American* 1991; **265**: 84–91.
59 Resnick P, Zeckhauser R, Swanson J, Lockwood K. The value of reputation on eBay: a controlled experiment. *Experimental Economics* 2006; **9**: 79–101.
60 Brooks FP. *The Mythical Man-Month: Essays on Software Engineering*. Addison-Wesley: Reading, MA, 1975.
61 Hardin R. The free rider problem. In: Zalta EN (Ed). *The stanford encyclopedia of philosophy*. 2003.
62 Cummings J. Work groups, structural diversity, and knowledge sharing in a global organization. *Management Science* 2004; **50**: 352–364.
63 Schultz-Hart S, Frey D, Lüthgens C, Moscovici S. Biased information search in group decision making. *Journal of Personality and Social Psychology* 2000; **78**: 655–669.
64 Burt RS. Structural holes and good ideas. *American Journal of Sociology* 2004; **110**: 349–399.
65 Wu F, Huberman BA, Adamic LA, Tyler JR. Information flow in social groups. *Physica A: Statistical and Theoretical Physics* 2004; **337**: 327.
66 Golder SA, Huberman BA. Usage patterns of collaborative tagging systems. *Journal of Information Science* 2006; **32**: 198–208.
67 Scheff TJ. Toward a sociological model of consensus. *American Sociological Review* 1967; **32**: 32–46.
68 Hastie R. Experimental evidence on group accuracy. In: Grofman B and Owen G (eds). *Decision Research*, Vol. 2, JAI Press: Greenwich, CT 1986; 129–157.
69 Gigone D, Hastie R. The common knowledge effect: information sharing and group judgment. *Journal of Personality and Social Psychology* 1993; **65**: 959–974.
70 Surowiecki J. *The Wisdom of Crowds*. Random House: New York, 2004.
71 Stasser G, Titus W. Pooling of unshared information in group decision making: biased information sampling during discussion. *Journal of Personality and Social Psychology* 1985; **57**: 67–78.
72 Eccles R, Kapler T, Harper R, Wright W. Stories in geotime. *IEEE Symposium on Visual Analytics Science and Technology* (Sacramento, CA) IEEE Computer Society: Washington, DC, 2007.